

Nekaj podatkov iz Besedotvornega slovarja slovenskega jezika za črko B po SSKJ

1 Besedotvorni slovar in besedna družina

Besedotvorni slovar slovenskega jezika je zasnovan kot tip enojezičnega specialnega slovarja, ki prinaša podatke o tvorbeni povezanosti besed, kar pomeni, da so ob nemotivirani iztočnici v ustreznih tvorbenih povezavah prikazane vse tvorjenke s skupnim korenem, ki so dosledno členjene na podstavni in obrazilni del. Za tako povezane besede velja v slovenščini poimenovanje *besedna družina*, po definiciji pa jo sestavljajo besede istega korena (J. Toporišič, 1992: 7; 1996: 89; 2000: 108, 116; A. Vidovič Muha, 2000: 108). Opozoriti velja, da je v ustrezni ruski, češki in slovaški literaturi mogoče najti poimenovanja *slovoobrazovatel'noe gnezdo* (Tihonov, 1985: 4), *slovotvorné hnízdo* (MČ, 1: 207–208) in *slovotvorné hniedzdo* (J. Mistrík, 1993: 410). Tako bi kazalo razmisliti, ali za enoto slovenskega besedotvornega slovarja uporabiti poimenovanje *besedotvorno gnezdo* ali *besedotvorna družina*. S prvim izrazom neposredno določimo, da gre za povezavo besed na sinhroni tvorbeni ravni, kar pri izrazu *besedna družina* ni tako enoumno, saj je ta lahko zasnovana tudi na diahronih, etimoloških merilih, kot kaže npr. naslov monografije *Besedna družina iz korena *god- v slovanskih jezikih* (A. Šivic Dular, 1999). Da pa je mogoče izraz *besedna družina* uporabiti tudi v primerih, ko gre za druženje tvorjenk, prepoznavnih na sinhronih pretvorbenih povezavah med motivirajočo in motivirano besedo, kaže izraz *die Wortfamilie*, uporabljen v naslovu besednodružinskega slovarja sodobnega nemškega jezika (G. Augst, 1998), zato je morda smiselno ostati pri tem terminu, ustaljenem tudi v slovenskem jezikoslovju.

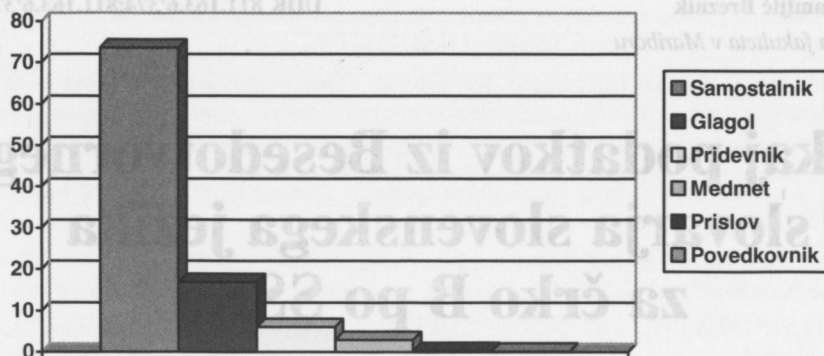
2 Število iztočnic in njihova besednovrstnost

Trenutno je obdelanih 594 iztočnic. Podatek še ni dokončen, vendar z gotovostjo lahko trdim, da se bo njihovo število gibalo v bližini 600 enot, pri tem pa je treba poudariti, da kot iztočnice nastopajo samo besede iz SSKJ z vsaj eno tvorjenko. Prvo besedno družino tako predstavlja medmet *ba*, zadnjo glagol *bzikati*.

Zanimivo je pogledati odstotno zastopanost posameznih besednih vrst v iztočnicah, ki pa podatkovno ni presenetljiva, saj je bilo mogoče pričakovati, da bo največ samostalnikov (73,6 %), sledijo glagoli (17 %), pridevniki (6,1 %), medmeti (2,9 %), prislovi (0,3 %) in povedkovniki (0,2 %).¹ Podatki prav gotovo ne odražajo stanja, kot bi ga dobili z obdelavo vsega besedja SSKJ od A do Ž, ker je tukaj omejitveni kriterij črka B. Verjetno pa bo držal vsaj podatek, da je med iztočnicami največ samostalnikov in glagolov (Augst, 2000: 5).

¹ Tukaj gre samo za povedkovnik bot. Pri določanju te besedne vrste upoštevam oznake v SP 2001.

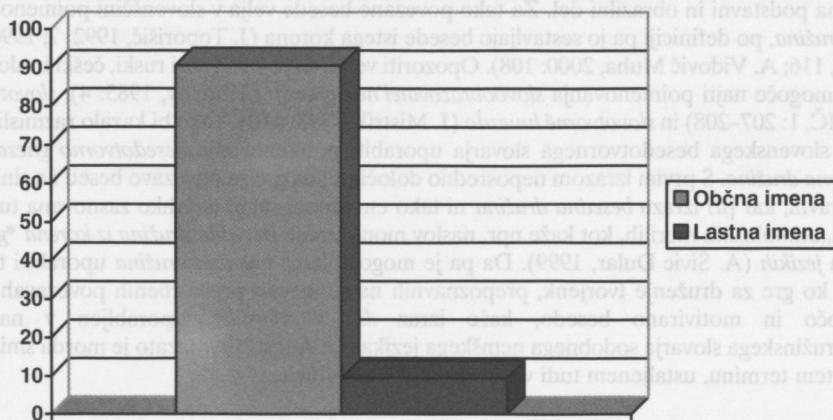
Diagram 1: Odstotni prikaz iztočnic po besednih vrstah.



3 Število obdelanih besed in obseg besednih družin

Po besednih družinah je razvrščenih približno 7720 besed. Pri tem je 90,6 % občnih in 9,4 % lastnih imen.

Diagram 2: Odstotni prikaz zastopanosti občno- in lastnoimenskih iztočnic.



Podatek, da v povprečju vsaka besedna družina vsebuje trinajst tvorjenk, je malo poveden, saj je v veliki meri opazna ekstremnost navzgor ali navzdol. Tako je besednih družin z eno tvorjenko 18,4 %, z dvema 10,9 %, s tremi 3,9 %, s štirimi ali več tvorjenkami pa kar 66,2 %.

Na videz se zdi, da je kar veliko majhnih besednih družin, toda njihove tvorjenke v celotnem obsegu predstavljajo le 4-odstotni delež, kar 96 % tvorjenk pripada večjim besednim družinam.

Samo v ponazorilo še o ekstremu druge vrste, kot so iztočnice *bog*, ki ima 372 tvorjenk ali 4,8 % vseh obdelanih besed, *biti*, ki ima približno 250 tvorjenk ali 3,2 % vseh obdelanih besed, *beseda* 240 tvorjenk ali 3,1 % in *bel* 117 tvorjenk ali 1,5 % vsega obdelanega besedja črke B.

Trenutno še nisem preverila podatka o razmerju med deležem tvorjenk, zajetih iz SSKJ, in deležem tvorjenk iz dodatnega gradivskega vira, ki ga predstavlja Besedišče slovenskega jezika (dalje BSJ). Na ta način bi lahko za črko B natančno določila število osamelih besed, tj. takih, ki ob sebi v SSKJ nimajo niti ene tvorjenke.

Diagram 3: Odstotni prikaz iztočnic po številu tvorjenk.

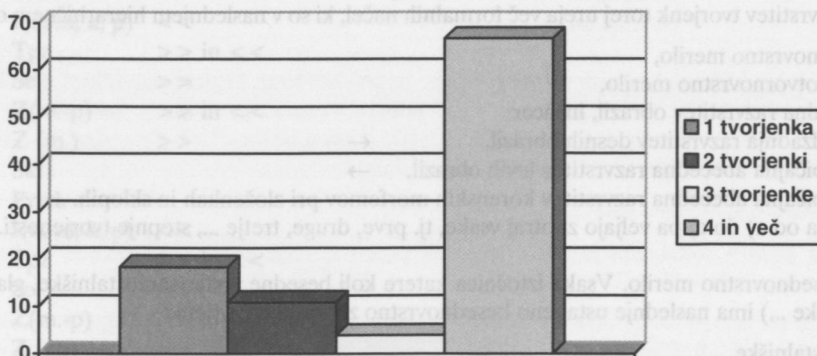


Diagram 4: Odstotni prikaz tvorjenk, ki pripadajo eno-, dvo- ali tričlenski besedni družini, v razmerju do tvorjenk, ki pripadajo štiri- ali veččlenski besedni družini.

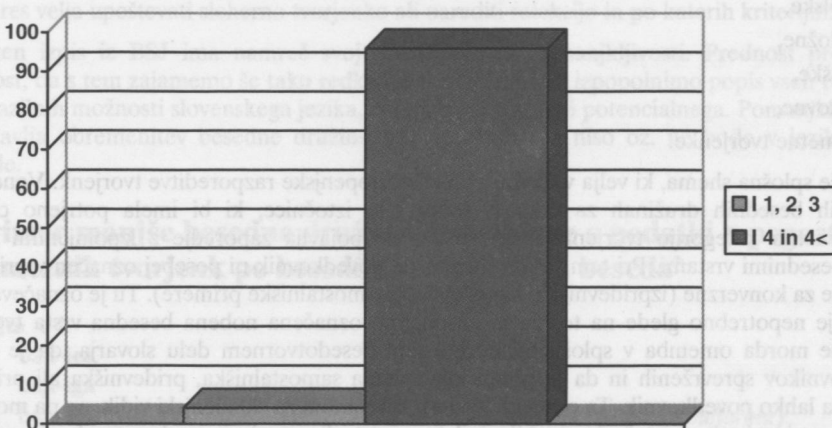
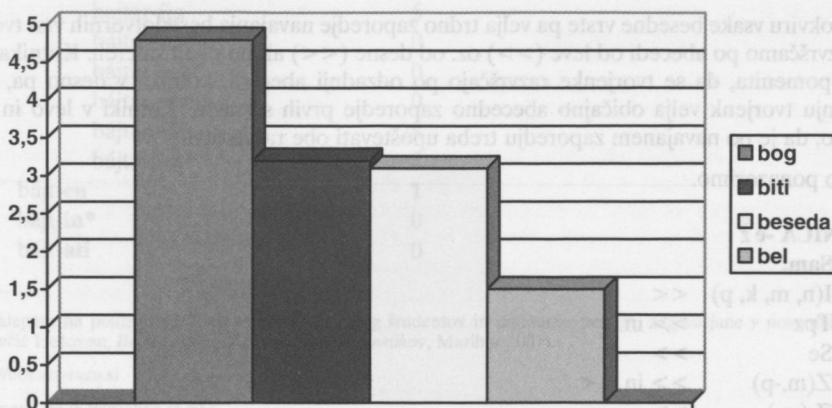


Diagram 5: Odstotni prikaz tvorjenk štirih večjih besednih družin glede na vse tvorjenke iztočnic na B.



4 Struktura besedne družine

4.1 Razvrstitev tvorjenk torej ureja več formalnih načel, ki so v naslednjem hierarhičnem odnosu:

- a) besednovrstno merilo,
 - b) besedotvornovrstno merilo,
 - c) abecedna razvrstitev obrazil, in sicer:
 - odzadnja razvrstitev desnih obrazil, ←
 - običajna abecedna razvrstitev levih obrazil, →
 - običajna abecedna razvrstitev korenskih morfemov pri zloženkah in sklopih.
- Merila od a) do c) pa veljajo znotraj vsake, tj. prve, druge, tretje ..., stopnje tvorjenosti.

4.1.1 Besednovrstno merilo. Vsaka iztočnica katere koli besedne vrste (samostalniške, glagolske, pridevniške ...) ima naslednje ustaljeno besednovrstno zaporedje tvorjenk:

- samostalniške,
- pridevniške,
- prislovne,
- povedkovniške,
- glagolske,
- predložne,
- vezniške,
- členkovne,
- medmetne tvorjenke.

Takšna je splošna shema, ki velja v okviru vsake istostopenjske razporeditve tvorjenk. Vendar pri obdelanih besednih družinah za črko B še ni bilo iztočnice, ki bi imela potrjeno celotno besednovrstno kategorijo tvorjenk. Najpogosteje se pojavlja zaporedje z izpolnjenimi prvimi petimi besednimi vrstami. Pri tem velja omeniti, da povedkovnik ni posebej označen v primerih, kadar gre za konverzne (izpridevniške, izprislovne, izsamostalniške primere). Tu je označevanje te kategorije nepotrebno glede na to, da ni eksplicitno označena nobena besedna vrsta tvorjenk. Dovolj je morda omemba v splošnem teoretičnem besedotvornem delu slovarja, da je večina povedkovnikov sprevrženih in da je potencialno vsaka samostalniška, pridevniška ali prislovna tvorjenka lahko povedkovnik. Ta podatek je zlasti relevanten za skladijski vidik, ne pa morda za besedotvorni, saj tako nastajanje povedkovnikov predstavlja manj pomemben tvorbeni postopek, t. i. konverzijsko izpeljavo, pri kateri ne gre za v besedotvorju temeljno morfološko spremembo, marveč le za spremembo oblikoslovne kategorije.

4.1.2 V okviru vsake besedne vrste pa velja trdno zaporedje navajanja besedotvornih vrst tvorjenk, ki jih razvrščamo po abecedi od leve (>>) oz. od desne (<<) ali po obeh smereh. Kotnika v levo namreč pomenita, da se tvorjenke razvrščajo po odzadnji abecedi, kotnika v desno pa, da pri razvrščanju tvorjenk velja običajno abecedno zaporedje prvih sestavin. Kotniki v levo in desno pomenijo, da je po navajanem zaporedju treba upoštevati obe razvrstitvi.

To lahko ponazorimo.

IZTOČNICA -e ž

1. stopnja **Sam.**

I(n, m, k, p)	<<
Tpz	>> in <<
Se	>>
Z(m.-p)	>> in <<
Z(m.)	>>

Skl	
Prid.	
I(n, m, k, p)	<<
Tpz	>> in <<
Se	>>
Z(m.-p)	>> in <<
Z (m.)	>>
Skl	
Prisl.	
I(n, m, k, p)	<<
Tpz	>> in <<
Se	>>
Z(m.-p)	>> in <<
Z (m.)	>>
Skl	
itn.	

Formalna ureditev besedne družine je uporabna in dovolj pregledna.² Zastavlja pa se vprašanje, ali v BSJ res velja upoštevati sleherno tvorjenko ali narediti selekcijo in po katerih kriterijih.

Dodaten izpis iz BSJ ima namreč svoje prednosti in pomanjkljivosti. Prednost predstavlja možnost, da s tem zajamemo še tako redke tvorjenke in s tem izpopolnimo popis vseh tvorbenih in obrazilnih možnosti slovenskega jezika, četudi zgolj na ravni potencialnega. Pomanjkljivost pa predstavlja obremenitev besedne družine s tvorjenkami, ki niso oz. ne bodo v jeziku nikoli zaživele.

5 Primer manjše besedne družine, dopolnjene s podatki o pogostnosti posameznih tvorjenk po besedilni zbirki Nova beseda³

16. bájta -e ž	192
bájt-ica	6
bájt-iga	0
bajt-úra* ⁴	0 (potrjena pa je tvorjenka <i>bajturina</i>)
bajt-išče*	0
bájt-ar	31
bájt-ar-ica	3
bajtar-tja	5
bájt-ar-ka	0
bájt-ar-ček*	0
bájt-ar-stvo	0
bájt-ar-ski	28
bájt-arj-ev*	2
bájt-en	1
bájt-in*	0
bájt-ati	0

² To sklepam na podlagi izdelanih seminarskih nalog študentov in diplomske naloge, zagovarjane v novembru 2001 (I. Kavrečič Holcman, *Besedne družine prvotnih povedkovnikov*, Maribor 2001).

³ <http://bos.zrc-sazu.si>

⁴ * pomeni, da je tvorjenka iz BSJ.

bájt-anje*	0
pod-bájtati*	0
podbájt-anje*	0

V celotnem besedilnem korpusu se oblika *-bajt-* pojavi 266-krat, in sicer največkrat vse sklonske oblike iztočnice (192), po pogostnosti sledi pomen opravkarja (*bajtar*), nato povezanost s samostalniško podstavo, izraženo s *-ski* (*bajtarski*), pomen lastnosti ali popredmetene lastnosti (*bajtarija*), manjšalnica (*bajtica*) in feminativ (*bajtarica*).

Katere podatke je mogoče izluščiti?

1. Beseda *bajta* v sodobnem jeziku ni posebej pogosta, verjetno je najpogostejše rabljena kot pogovorna vzporednica k *hiši*.
2. S tem je utemeljen obseg besedne družine, ki velja za manjšo, saj ima skupno 18 tvorjenk, od tega le 10 potrjenih tudi po SSKJ.
3. Podatki besedilnega korpusa potrjujejo le osem tvorjenk. Zato bi morda v tem primeru zadoščal naslednji obseg besedne družine, primeren morda za šolsko izvedbo slovarja:

16. bájta -e ž

bájt-ica	6
bajt-úra* NB: bajtúrina	1
bájt-ar	31
bájtár-ica	3
bajtar-íja	5
bájtár-ski	28
bájtárj-ev*	2
bájt-en	1

4. Tvorjenka *bajtarica* je pogostejša od *bajtarke*.⁵

6 Uporabnost slovarja

Slovar besednih družin je v prvi vrsti uporaben za jezikoslovno stroko, ki na ta način lahko dobi a) natančen pregled nad deležem netvorjenk in tvorjenk; b) povprečen obseg besednih družin s podatki o tovrstnih ekstremih navzgor ali navzdol; c) popis vseh možnih besedotvornih obrazil in njihovo pogostnost. Tak slovar je mogoče dopolniti tudi s pomeni tvorjenk in njihovo besedilno rabo, s čimer se po uporabnosti približa enojezičnemu nacionalnemu slovarju, le da je abecedni ustroj zamenjan z besednodružinskim.

Iz njega bi lahko izhajala izdelava šolskega besedotvornega slovarja, ki bi selektivno ožal besedne družine glede na pogostnost tvorjenk, preverjenih po besedilnem korpusu.

Na podlagi tega bi se lahko ugotovili najpogostejši ustroji besednih družin za posamezne besedne vrste. Ta podatek bi lahko vodil k sestavi takega slovarja za tujce, v katerem bi ti lahko dobili pregled najtipičnejših vzorcev za nastajanje novih tvorjenk, s pomočjo katerih bi po analogiji pridobivali in razumeli novo besedje. Pri tem bi bila razmisleka vredena možnost, da bi tak slovar izdelali po načelu tipičnih pomenskih polj. Obstaja namreč velika verjetnost, da imajo besede, pripadajoče pomenskemu polju, ki ga določa skupna nadpomenka, tudi precej podobno strukturo besednih družin. S tem bi učečim se slovenskega jezika ponudili več podatkov hkrati, in sicer

⁵ Novi pravopis sploh prinaša le obliko *bajtarica* in dodatno še svojilni pridevnik *bajtaričin*. *Slovenski pravopis*, Ljubljana 2001, 347.

možnost vzporejanja pojmovne, besednodružinske in pomenske členjenosti svojega maternega in slovenskega jezika (I. Stramlijič Breznik 2001b).

Literatura

G. Augst (1998). Wortfamilienwörterbuch der deutschen Gegenwartssprache. Tübingen.

G. Augst (2000). Die Mächtigkeit der Wortfamilien — Quantitative Auswertungen zum 'Wortfamilienwörterbuch der deutschen Gegenwartssprache'. Praxis- und Integrationsfelder der Wortbildungsforschung. Heidelberg, 1–18.

J. Mistrík (1993). Encyklopédia jazykovedy. Bratislava, 410.

Mluvenice češtiny 1 (1986). Praga, 207–208.

I. Stramlijič Breznik (1998). Elementi besedotvornega slovarja v strukturi samostalniških gesel Murkovega slovensko-nemškega slovarja. Murkov zbornik. Maribor (Zora 9), 249–258.

— — (1999a). Raziskovalne perspektive slovenskega besedotvorja. Zbornik Slavističnega društva Slovenije 10: Slovensko jezikoslovje danes in jutri — Slovenski slavistični kongres, Celje, 112–118.

— — (1999b). Prispevki iz slovenskega besedoslovja, Maribor.

— — (2000). Besedna družina besede. Jezikoslovni zapiski 6, 45–55.

— — (2001a). Prikaz besedne družine v besedotvornem slovarju. Studia Slavica Savariensia 1–2, 18–30.

— — (2001b). Kateri podatki iz besedotvornega slovarja so uporabni v dvo- ali večjezični leksikografiji? (Referat je bil predstavljen na 3. mednarodnem leksikološko-leksikografskem znanstvenom skupu, Zagreb 16. 11. 2001, v tisku).

A. Šivic Dular (1999). Besedna družina iz korena *god- v slovanskih jezikih. Ljubljana.

A. N. Tihonov (1985). Slovoobrazovatenyj slovar russkogo jazyka. Moskva.

J. Toporišič (1992). Enciklopedija slovenskega jezika. Ljubljana.

— — (1996). Slovenski jezik in sporočanje 2. Maribor, 89.

— — (2000). Slovenska slovnica. Maribor, 108; 116.

A. Vidovič Muha (2000). Slovensko leksikalno pomenoslovje. Govorica slovarja. Ljubljana, 108.

Vir

Slovar slovenskega knjižnega jezika z Odzadnjim slovarjem slovenskega jezika in Besediščem slovenskega jezika z oblikoslovnimi podatki: Elektronska izdaja na plošči CD-ROM. Ljubljana 1998.

UDK 811.163.6'374:811.163.6'373.611

**SOME DATA FROM THE SLOVENE LANGUAGE WORD-FORMATION
DICTIONARY FOR LETTER B IN THE DICTIONARY OF THE STANDARD
SLOVENE LITERARY LANGUAGE**

Corpus analysis could also be used to determine the most common structure of word families by individual parts of speech. This data could then be used to produce a dictionary for non-native users which would provide an insight into the most typical patterns of formation of new words and help the users to acquire and understand new words on the basis of analogy. An alternative might be to organise such a dictionary on the principle of semantic fields, as it is highly probable that words belonging to a semantic field determined by a common hypernym also have a very similar structure of their word families. Such a dictionary would provide learners of Slovene with several kinds of data at the same time, a possibility to compare the conceptual, word-family and semantic structure of their mother tongue and Slovene.