

MESTO VELIKEGA SLOVARJA TUJK V SLOVENSKI LEKSIKOGRAFIJI

Članek obravnava *Veliki slovar tujk*, ki je leta 2002 izšel pri Cankarjevi založbi. Slovar tujk kot tip slovarja ni običajen v vseh leksikografskih tradicijah, za kar je več razlogov: besedotvorni, fonetični, puristični, brez dvoma pa je pomembno tudi število govorcev določenega jezika in njihovo pojmovanje »tujega« oz. »domačega«. Za ovrednotenje in umestitev slovarja znotraj slovenske leksikografije je bil vzorec gesel primerjalno pregledan z viri, ki so v slovarju navedeni kot primarni, in s korpusom *FIDA*. Analiza je pokazala, da je bilo zajetih veliko strokovnih izrazov, zaradi česar je to dejansko terminološki slovar velikega obsega. Nasploh je po pregledu gradiva mogoče reči, da takšen koncept slovarja tujk danes ni več aktualen; postavlja se tudi vprašanje, kakšen smisel imajo ponovne (prenovljene) izdaje starih leksikografskih priročnikov, saj so sodobni uporabniki vse zahtevnejši in pričakujejo kakovostne leksikografske izdelke.

1 Slovarji tujk in njihov namen

Mednarodni kulturni in gospodarski stiki se v jezikih odražajo z vnašanjem tujejezičnih prvin, kar večinoma izhaja iz kulturoloških posebnosti in posledično leksikalnih praznin v ciljnem jeziku (npr. poimenovanja za pojmovni in predmetni svet posameznih kultur). Drugi razlog za vnašanje tujih besed in besednih zvez je ustaljena mednarodna raba (takšni so npr. latinski izrazi v pravu in medicini). V slovenski, hrvaški, nemški in še nekaterih leksikografskih tradicijah takšne besede in besedne zveze obravnavajo t. i. slovarji tujk. Po eni strani je poznavanje tujk in njihove pravilne rabe večinoma pojmovano kot znak visoke izobrazbe in splošne razgledanosti, po drugi strani pa je motivacija za slovarje tujk jezikovni purizem – torej evidentiranje »tujega« in iskanje ustreznega, pravilnega in seveda »domačega«.

Na Slovenskem je najbolj znan, tudi zaradi številnih ponatisov, Verbinčev *Slovar tujk* iz leta 1968. V povojni slovenski leksikografiji ima ta slovar vsekakor pomembno mesto, saj je do izida celotnega *Slovarja slovenskega knjižnega jezika* (zadnja knjiga je izšla leta 1991) skupaj s slovarskim delom Slovenskega pravopisa iz 1950 oz. 1962 služil tudi kot splošni slovar.

Anglo-ameriška leksikografija ne pozna slovarjev tujk, kakršne poznamo v slovenskem in nemškem okolju, vseeno pa se pojavljajo vsebinsko podobni priročniki,

vendar bistveno manjšega obsega. Takšen je npr. *The Oxford Dictionary of Foreign Words and Phrases* iz leta 2000, ki navaja prek 8.000 pogostih besed in besednih zvez iz afrikanščine, francoščine, grščine, hebrejščine, ruščine, turščine itd. Priročnik poleg splošnega jezika pokriva tudi nekatera specializirana področja, npr. glasbo, umetnost, kulinariko, medicino, literarno teorijo. Iztočnice so opremljene z morebitnim variantnim zapisom, razlago, prevodom pomena v izvornem jeziku, naveden pa je tudi podatek, kdaj se je iztočnica prvič pojavila v angleščini.

2 Veliki slovar tujk Cankarjeve založbe

Nove in prenovljene oz. posodobljene izdaje slovarjev (eno-, dvo- ali večjezičnih, splošnih ali strokovnih) so v času hitrega razvoja in spreminjanja kulturnih, družbenih, gospodarskih in drugih okoliščin nujne tudi pri narodih z relativno majhnim številom govorcev določenega jezika. Majhno število govorcev seveda ne pomeni, da ti ne potrebujejo novih slovarskih del – novih tako v smislu posodobitve v slovarju predstavljenega gradiva kot tudi posodobitve slovarske obravnave in tiskovnih rešitev. Tako sta bila zaradi pravopisne reforme nemščine pri Založbi DZS izdana prenovljena *Veliki slovensko-nemški slovar* in *Veliki nemško-slovenski slovar*.¹

Veliki slovar tujk (dalje VST), ki je leta 2002 izšel pri Cankarjevi založbi, naj bi bil naslednik *Slovarja tujk* Franceta Verbinca, prvič izdanega leta 1968 in pozneje še večkrat ponatisnjenega. Zaradi delne zastarelости in relativno skromnega obsega so uredniki sprva načrtovali »le posodobitev in dopolnitev slovarja z gradivom iz tujih slovarjev«, po nekaj več kot treh letih zbiranja in urejanja gradiva pa se je pokazalo, da bo treba slovar bistveno prenoviti glede zasnove, besedišča in obravnave (*Veliki slovar tujk* 2002: VIII).²

Za boljšo preglednost slovarja je uporabljen dvobarvni tisk: v modri barvi so iztočnice, podiztočnice in številske oz. črkovne oznake za pomene, preostalo besedilo pa je v običajni črni barvi.³ Iztočnice so opremljene z naglasom in/ali izgovorjavo ter z izvorom (večinoma); področne oznake (kar 84 različnih) pri večpomenskih iztočnicah lahko pokrivajo ali vse ali le posamezne pomene, slovar pa opozarja tudi na figurativno rabo.

3 Primerjava VST z nekaterimi drugimi slovarji

V VST je naveden obsežen seznam virov – od domačih in tujih ozko strokovnih do običajnih, splošnih slovarjev in enciklopedij. Pri zajemanju gradiva so sestavljavci

¹ Ta slovarja sta po nastanku sicer dokaj nova, vsaj v primerjavi z Gradovim *Velikim angleško-slovenskim slovarjem*, Jurančičevim *Srbskohrvatsko-slovenskim slovarjem*, *Slovarjem slovenskega knjižnega jezika* in številnimi drugimi slovarji.

² Prenovljeni slovar prinaša Uvod (VIII–IX), Navodila za rabo (X–XI), Temelje *Velikega slovarja tujk* (XII), Vire (XIII–XVI), slovarski del z okrog 40.000 iztočnicami (1–1248), Reke, citate in fraze (1249–1288), Simbole, kratice in okrajšave (1289–1300) ter Kratice v slovarju (1301–1303).

³ Po navedbah ob izidu je bil prvi slovar z dvobarvnim tiskom v Sloveniji *Splošni angleško-slovenski in slovensko-angleški moderni slovar* Daše Komac, ki je leta 2001 izšel pri Cankarjevi založbi. Sicer pa tudi v Sloveniji dvobarvni tisk prehaja v standard, tako za dvojezične slovarje manjšega ali večjega obsega kot za enojezične enciklopedične in druge leksikografske izdelke.

želeli v čim večjem obsegu pokriti strokovna področja, ki so se na novo pojavila oz. se opazno razvila v obdobju od prve izdaje Verbinčevega slovarja, a so se hkrati nedvomno zavedali razhajanja med spremenjeno zasnovo slovarja in njegovim naslovom: »Izraz tujka v naslovu *Velikega slovarja tujk* je ostal po naslovu Verbinčevega *Slovarja tujk*. Dejansko v slovar niso uvrščene samo tujke, temveč v veliki meri izposojenke iz tujih jezikov.« (*Veliki slovar tujk* 2002: XII.)⁴

Kakšno besedje torej pokriva *VST* in kdo naj bi bili njegovi uporabniki? V čem se razlikuje od drugih slovarjev na trgu? Za primerjalno analizo *VST* so bili izbrani Verbinčev *Slovar tujk* (dalje *ST*; izhodišče), *Slovar slovenskega knjižnega jezika* (dalje *SSKJ*; splošni slovar, vendar že časovno oddaljen), *Slovenski etimološki slovar* (dalje *SES*; poljudna različica specialnega slovarja) in *Slovenski pravopis* iz 2001 (dalje *SP* 2001; novejši normativni priročnik s slovarskim delom, temelječim na *SSKJ*),⁵ ki so navedeni tudi med primarnimi viri za *VST*.⁶ Namen analize vzorca iz *VST* je bil ugotoviti, kolikšen del gradiva je glede na omenjene slovarje nov, katera področja pokriva *VST*, hkrati pa tudi ovrednotiti praktično uporabnost slovarja. Pri vrednotenju običajno ni niti časovno mogoče niti smiselno pregledovati slovarjev v celoti, zato je potrebno izbrati vzorec, ki ustreza kriterijema reprezentativnosti in naključnosti (Landau 1989: 306, 308).

3.1 Osnovni vzorec

Vzorec sestavlja prvih deset polnopomenskih iztočnic vsake tretje črke, tj. A, Č, F, I, L, O, R, T, W in Z. V vzorcu niso upoštevane enočrkovne iztočnice tipa A, a¹, a², niti iztočnice tipa 'v sestavljenkah', npr. a⁻¹ (ab-, abs-), a⁻² (ad-). Prav tako niso upoštevane kratične iztočnice tipa NA (*Network Adapter*). V vzorcu je tako 100 iztočnic, skupaj s podiztočnicami in kazalkami obsega vzorec 119 različnih leksemov, in sicer: aak, ab, *aba¹ (K), aba², abacija, abada, abaditi, *abadjeh (P), abaisse, abak, abakus (K), čačača, čador, čaire, čaj, čajen (P), čajevec (P), čajevke (P), čajka, čaka, čakan, čakavščina, čakra, čamara, fa, fabela, fabergé, fabianizem, fabijec, fable convenue, fabliau, fabricirati, fabriji, fabrikacija, fabrikacijski (P), iai-do, *iatrofizik (K), iatrofizika (P, K), iatrogen, iatrogenija, iatrokemija, iatrologija, *iatromehanik (P), iatromehanika (2 × P), ibadat, ibaditi (K – iztočnica v *VST* manjka), ibalide, iberis, kombuša (K), la, labamba, labanotacija, labarum, labdanum, label, la belle époque, labelum, labent, laberdan, ladanum (P), lavdanum (P), oaskop, oaza, obbligato, ob defectum, obducent, obducentka (P), obducirati, obducija, obduracija, obediencia, obelisk, obligato (K), rabarbara, rabat⁽¹⁾, rabat⁽²⁾, rabdit, rabditiformen, rabdofobija, rabdoiden, rabdom, rabdomantija, rabdomioliza, taal, tab, tabagija, tabakera, tabakizem, tabakoza, *tabasco (P), *tabasko (K), tabaskovka, tabatiera, vajang (K), valter (K), vampum (K), voki-toki (K), vokman (K),

⁴ Avtorji *VST* so verjetno izhajali iz Toporišičevih definicij izrazov *tujka* in *izposojenka* (Toporišič 1992: 299–300, 334).

⁵ *SP* 2001 je bil na začetku zastavljen kot kritika *SSKJ*, vendar je zlasti v slovarskem delu veliko prevzel iz *SSKJ* – tako glede obdelave gesel kot tudi samega geslovnika (*Slovenski pravopis* 2001: VII–IX; Vojnovič, 2001).

⁶ Bibliografski podatki so navedeni v razdelku Viri na koncu prispevka.

*wagireh, wagnerjanski, wajang, walkie-talkie, walkman, wallnut, walter, wampum, warp, wash-and-wear, zabaione, zabit, zahirit, zaibatsu, zaidit, zaidizem (P), zair, zajval, zakrament, zakramental, zakristan.*⁷

3.1.1 Področne oznake

V vzorcu je s področnimi oznakami natančneje določenih 60 enot (iztočnice, podiztočnice oz. posamezni pomeni). Največkrat – sedemkrat – se pojavi oznaka MED. (*iatrogenija, obducirati, obduracija, rabdofobija, rabdoiden, rabdomioliza, tabakoza*),⁸ po petkrat nastopata oznaki BOT. (*čajevce, čajevke, iberis, labelum, rabarbara*) in ZOOL. (*ibalide, rabdit, rabditiformen, tabaskovka, zajval*), po štiri enote imajo oznake GASTR. (*abaisse, laberdan, tabasko, zabaione*), REL. (*ibadat, obediencia, zahirit, zaidizem*), TEKST. (*abada, wagireh, warp, wash-and-wear*) in ZGOD. (*čaka, fabijec, fabriji, zabit*), po tri enote nosijo oznake GLASB. (*čakan, fa, la*), LITURG. (*labarum, zakrament, zakramental*) in PLES (*čaćača, labamba, labanotacija*), po dvakrat se pojavijo oznake ANAT. (*fabela, rabdom*), FNC. (*labent, zaibatsu*), LINGV. (*čakavščina, taal*) in TRG (*label, rabat*⁽²⁾). Devet oznak je zastopanih samo z eno enoto: ALP. (*wallnut*), GEORG. (*oaza*), INF.⁹ (*tab*), LIT. (*fabliau*), RAČ. (*label*), ŠP. (*iai-do*), UM. (*la belle epoque*), UM.ZGO. (*fabergé*) in VET. (*oaskop*). Enkrat se pojavi tudi kombinacija oznak ZGOD. in POLIT. (*fabianizem*).

V nekaterih primerih ni povsem jasno, do kod naj bi oznaka segala (npr. pri *labarum* verjetno velja samo za 2. pomen, pri *iberis* pa za oba pomena). Ponekod je ločevanje med področji preveč podrobno in zato marsikje nedosledno, npr. med religijo in liturgijo, medicino oz. veterino in anatomijo, informatiko in računalništvom, kar po eni strani izhaja iz različne obravnave v virih, na katere so se sodelavci slovarja opirali, po drugi strani pa iz nedorečene in neusklajene uredniške politike; v nekaterih primerih je prišlo tudi do podvajanja podatkov, npr. pri *zabit* je razlaga prvega pomena 'v Perziji nekdanji upravnik majhnega okrožja' opremljena še z oznako ZGOD.

3.1.2 Prekrivanje med VST ter ST, SSKJ, SES in SP 2001

Od skupno 119 leksemov jih je bilo 57 obdelanih že v vsaj enem od primerjanih slovarjev; šest leksemov je obdelanih v vseh slovarjih: *fabricirati, oaza, obelisk, rabarbara, rabat*⁽²⁾, *zakrament*, šest pa jih ima potrditev v *SSKJ, SES* in *SP 2001*: *čaj, čajjen, čajevce, la, voki-toki, walkie-talkie*.¹⁰

⁷ Označevanje vzorca: * = iztočnica, ki se v vzorcu pojavi še kot kazalka ali podiztočnica; K = kazalka; P = podiztočnica. Manjkajoče izpostavne številke so v zavitih oklepajih.

⁸ Na nesorazmerno velik delež medicinskega izrazja v *VST* opozarja tudi Klun in navaja, da glede na vzorec s 124 strani iz *VST* medicinska terminologija predstavlja približno 23 % celotnega slovarja; pri tem so upoštevane tudi povezane oznake, npr. ANAT., BIOKEM., FARM. (Klun 2003: 8). — V osnovnem vzorcu iz tega prispevka je delež medicinske terminologije glede na Klunove kriterije 7,6 %, v kontrolnem pa 8,3 %.

⁹ Razdelek *Kratice* v slovarju navaja vse področne oznake brez pik, na str. 1301 pa je navedena drugačna oznaka, kot je v vzorcu, tj. INFO. Za uporabnike je ta razdelek dokaj nepregleden, ker so področne oznake v abecednem seznamu skupaj s slovničnimi in jezikovnimi oznakami ter drugimi okrajšavami.

¹⁰ V drugi izdaji *SES* iz leta 2003 sta obdelana še leksema *obducirati* in *obdukcija*.

	<i>VST</i>	<i>ST</i>	<i>SSKJ</i>	<i>SES</i>	<i>SP 2001</i>
1.	abak	—	+	—	+
2.	abakus	+	+	—	+
3.	čačača	+	+	—	+
4.	čaj	—	+	+	+
5.	čajen	—	+	+	+
6.	čajevec	—	+	+	+
7.	čajevke	—	+	—	+
8.	čajka	—	+	—	+
9.	čaka	+	+	—	+
10.	čakan	+	—	—	—
11.	čakavščina	—	+	—	+
12.	čakra	—	—	—	+
13.	čamara	+	—	—	—
14.	fa	—	+	—	+
15.	fabianizem	—	+	—	+
16.	fabijec	+	+	—	+
17.	fable convenue	+	—	—	—
18.	fabliau	+	—	—	—
19.	fabricirati	+	+	+	+
20.	fabrikacija	+	+	—	+
21.	fabrikacijski	—	+	—	+
22.	iatrokemija	+	—	—	+
23.	kombuša	—	+	—	—
24.	la	—	+	+	+
25.	labarum	+	—	—	—
26.	labdanum	+	—	—	—
27.	label	+	—	—	—
28.	laberdan	+	—	—	—
29.	ladanum	+	—	—	—
30.	lavdanum	+	+	—	—
31.	oaza	+	+	+	+
32.	obligato	+	—	—	—
33.	obducent	+	+	—	+
34.	obducentka	—	—	—	+
35.	obducirati	—	+	—	+
36.	obdukciija	+	+	—	+
37.	obduracija	+	—	—	—
38.	obediencia	+	—	—	—
39.	obelisk	+	+	+	+
40.	rabarbara	+	+	+	+
41.	rabat ^[2]	+	+	+	+
42.	tabakoza	+	—	—	—
43.	tabasco	—	—	—	+
44.	tabasko	—	—	—	+
45.	tabatiera	—	+	—	—
46.	valter	—	+	—	+
47.	vampum	+	—	—	—
48.	voki-toki	—	+	+	+
49.	vokman	—	—	—	+
50.	wagnerjanski	—	+	—	+
51.	walkie-talkie	—	+	+	+
52.	walkman	—	—	—	+
53.	walter	—	+	—	+
54.	warp	+	—	—	—
55.	zakrament	+	+	+	+
56.	zakramental	—	+	—	+
57.	zakristan	+	+	—	+

3.2 Kontrolni vzorec

Postopkov za vzorčenje je seveda več: lahko se odločimo npr. za zajetje iz vseh črk ali pa samo iztočnic na določenih straneh; primerljivost vzorcev, pridobljenih z različnimi postopki, pa pokaže, ali vzorec dovolj dobro predstavlja slovar v celoti. Za preverjanje je bil v tem primeru pripravljen manjši vzorec tako, da je bila upoštevana prva polnopomenska iztočnica na vsaki 60. strani (začetek na str. 60), kar je 20 iztočnic, skupaj s kazalkami in podiztočnicami pa 24 leksemov: *apertura*, *bikomponenten*, *debalirati*, *dramaturg*, *dramaturgija* (P), *dramaturginja* (P), *epizootičen*, *epizootski* (K), *fosfat*, *hacendero*, *idiličen*, *kämpavisor*, *konformen*, *liberizem*, *mentol*, *nekropsija*, *panhaks*, *polacca*, *poloneza* (K), *radian*, *sadhak*, *spinel*, *taseografija*, *ultraflotacija*.¹¹

3.2.1 Področne oznake

V kontrolnem vzorcu je s področnimi oznakami natančneje določenih dvanajst enot (iztočnice, podiztočnice oz. posamezni pomeni). Dvakrat se pojavi oznaka MINERAL. (*fosfat*¹, *spinel*), po enkrat pa ANAT. (*apertura* – 1. pomen), EKON. (*liberizem*), FIZ. (*apertura* – 4. pomen), LIT. (*kämpavisor*), MAT. (*radian*), MED. (*nekropsija*), METAL. (*ultraflotacija*), REL. (*sadhak*), TEHN. (*taseografija*) in zool. (*panhaks*).

3.2.2 Prekrivanje med VST ter ST, SSKJ, SES in SP 2001

Od 24 leksemov je 13 obravnavanih v vsaj enem od primerjanih slovarjev, štirje pa so obdelani v vseh slovarjih: *dramaturg*, *fosfat*, *idiličen*, *poloneza*.

	VST	ST	SSKJ	SES	SP 2001
1.	<i>apertura</i>	+	+	–	+
2.	<i>dramaturg</i>	+	+	+	+
3.	<i>dramaturgija</i>	+	+	–	+
4.	<i>dramaturginja</i>	–	–	–	+
5.	<i>fosfat</i>	+	+	+	+
6.	<i>idiličen</i>	+	+	+	+
7.	<i>konformen</i>	+	+	–	+
8.	<i>mentol</i>	+	–	+	+
9.	<i>nekropsija</i>	+	–	–	–
10.	<i>polacca</i>	+	–	–	–
11.	<i>poloneza</i>	+	+	+	+
12.	<i>radian</i>	+	+	–	+
13.	<i>spinel</i>	+	+	–	+

¹¹ Označevanje je opisano v opombi 7.

3.3 Primerjava med osnovnim in kontrolnim vzorcem

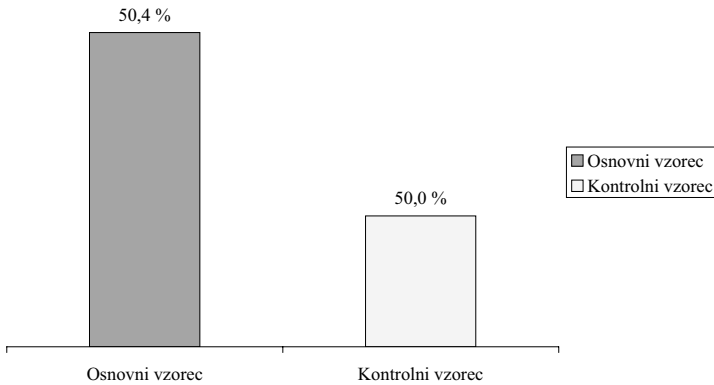


Diagram 1: *Delež leksemov s področnimi oznakami.*

Delež področnih oznak v obeh vzorcih iz *VST* je skoraj enak in tako – če sklepamo po rabi področnih oznak – kar polovica leksemov v vzorcu predstavlja strokovne izraze. V obeh vzorcih skupaj nastopa 30 različnih oznak, od tega se jih pet pojavi v obeh: ANAT., LIT., MED., REL. in ZOOL.

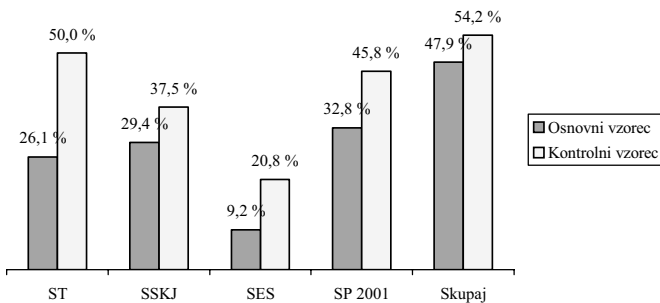


Diagram 2: *Pokritost VST s primerjanimi slovarji.*

Pri primerjavi slovarjev se je pokazalo, da kontrolni vzorec tako posamično kot skupno kaže višjo stopnjo prekrivanja leksemov kot osnovni. Največje razlike med vzorci so pri *ST* in *SES*, kjer je delež prekrivanja v osnovnem vzorcu skoraj dvakrat nižji.

4 *VST* in korpus *FIDA*

4.1 Korpusi v sodobnem jezikoslovju

Tradicionalni način sestavljanja slovarjev je temeljil na izpisovanju gradiva, urejenega v (večinoma abecedne) listkovne kartoteke. Iz takih zbirk so slovaropisci zajemali gesla, iskali potrditev pomenov in čim bolj tipične zglede, prepoznavali pogo-

ste oz. stalne zveze, ugotavljali terminološko rabo posameznih besed ali zvez. Podobno velja tudi za črpanje gradiva iz drugih slovarjev in leksikonov. Takšen način ročnega zbiranja in razvrščanja gradiva je zelo zamuden, zbrati in analizirati je mogoče le razmeroma majhno število besedil, presojanje o tem, kaj in v kakšni obliki bo sprejeto v slovar pa je omejeno na subjektivno presojo posameznika oz. skupine. Da bi lahko odpravili dejavnik subjektivnosti in hkrati pridobili čim več objektivnih, empirično preverljivih podatkov, so raziskovalci začeli razvijati elektronske besedilne korpuse (McEnery in Wilson 2001: 103), ki so že postali standardni vir podatkov za izdelavo slovarjev – sprva predvsem enojezičnih, v zadnjem času pa tudi dvojezičnih (Krek 1999: 8, McEnery in Wilson 2001: 144–145).¹²

Potrebe uporabnikov narekujejo oblikovanje različnih vrst korpusov, ki se razlikujejo po obsegu, zajetu zvrsti, časovnih okvirih, prenosniku, številu upoštevanih jezikov ipd. (Gorjanc 1999b: 51–52). Najbolj običajni so enojezikovni referenčni korpusi, ki so zaradi uravnoteženih razmerij med vrstami besedil in zajemanja žive, sodobne jezikovne rabe nujni za slovnične, pomenoslovne, besediloslovne, prevodoslovne in druge raziskave in s tem tudi za slovaropisce. Takšen tip korpusa v slovenskem prostoru predstavlja *Korpus slovenskega jezika FIDA* (<http://www.fida.net>); podobne ambicije ima tudi besedilni korpus *Nova beseda* (http://bos.zrc-sazu.si/s_beseda.html), ki nastaja na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU, a je v času priprave in izida *VST* zajemal le iz časnika *Delo* (Weiss 2001: 420).

4.2 Razmerje med *VST* in korpusom *FIDA*

Zasnova *VST* kaže, da je uredništvo želelo ustreči čim širšemu delu publike – učencem in študentom ter učiteljem in profesorjem, publicistom, prevajalcem in celo znanstvenikom – na to kažeta tako obseg slovarja kot pokrivanje velikega števila strokovnih področij. *VST* je bil sicer sestavljen še na tradicionalen način, tj. s črpanjem iz omejenega obsega pisnih virov, ki so jih uredniki in njihovi sodelavci pregledovali in gesla po osebni presoji uvrščali v slovar; zaradi čim bolj objektivne ocene slovarja pa je bilo smiselno preveriti, koliko leksemov in s kakšno pogostnostjo je mogoče najti v korpusu *FIDA*. Dobljene rezultate je treba vrednotiti z zavestjo, da referenčni korpus ni namenjen specializiranim jezikoslovnim raziskavam, npr. terminološkim (Stabej 1998: 104), zato pa daje preverljive (tudi številsko) podatke o sodobnem jeziku v splošni rabi. Če kot merilo za splošnost oz. terminološkost besedišča v *VST* vzamemo razmerje med deležema neoznačenih in označenih leksemov iz obeh vzorcev, lahko domnevamo, da je polovica *VST* splošnega in polovica terminološkega. Rezultati preverjanja po korpusu *FIDA* deloma potrjujejo to oceno, saj pri osnovnem vzorcu ni bilo potrditve za 66 leksemov, pri kontrolnem pa za 12. Deleži potrjenih in nepotrjenih leksemov so prikazani na diagramih 3 in 4.¹³

¹² Na Slovenskem je prvi slovar na korpusni osnovi sploh *Veliki angleško slovenski slovar Oxford*, ki je rezultat sodelovanja med založbama Oxford University Press in DZS (Grabnar in Šorli 2003: 127–128). Prva knjiga A–K je izšla februarja 2005.

¹³ Pri leksemih *aba¹*, *aba²*, *fa*, *la* in *tab* iz osnovnega vzorca ni bilo mogoče potrditi pomena oz. ugotoviti dejanskega števila enot, ker ti leksemi v korpusu *FIDA* niso lematizirani, ročna obdelava zadetkov pa bi zaradi homonimnosti zahtevala preveč časa.

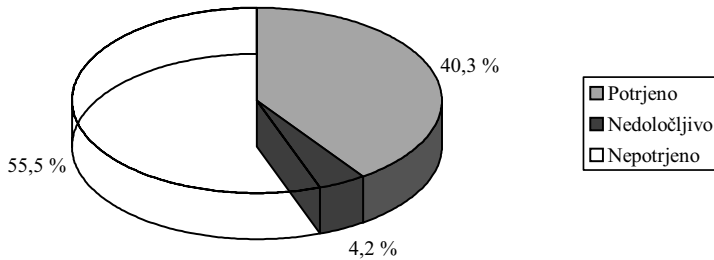


Diagram 3: Preverjanje VST po korpusu FIDA – osnovni vzorec.

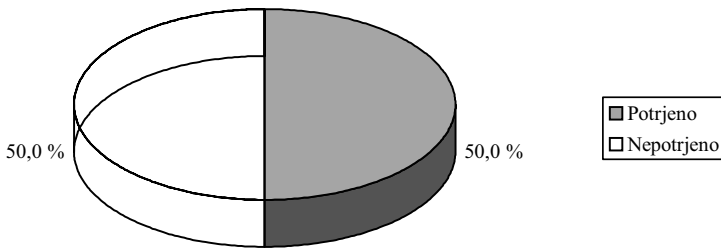


Diagram 4: Preverjanje VST po korpusu FIDA – kontrolni vzorec.

Izmed 57 leksemov iz osnovnega vzorca, ki so potrjeni vsaj v enem od primerjanih slovarjev, jih korpus FIDA potrjuje 36 (v oklepaju je navedeno število zadetkov): *abak* (10), *abakus* (47), *čačača* (6), *čaj* (3082), *čajen* (456), *čajevec* (12), *čajevke* (5), *čajka* (26),¹⁴ *čaka* (4), *čakra* (397), *fabliau* (3), *fabricirati* (16), *fabrikacija* (29), *labarum* (1), *label* (13), *lavdanum* (4), *oaza* (570), *obligato* (1), *obducent* (14), *obducirati* (7), *obdukcija* (459), *obelisk* (167), *rabarbara* (126), *rabat*⁽²⁾ (80), *tabasco* (15), *tabasko* (14), *vampum* (1), *voki-toki* (17), *vokman* (1), *wagnerjanski* (26), *walkie-talkie* (14), *walkman* (142), *walter* (10), *zakrament* (589), *zakramental* (3), *zakristan* (3). V kontrolnem vzorcu je potrjenih 11 od 13 leksemov: *dramaturg* (233), *dramaturgija* (337), *dramaturginja* (128), *fosfat* (202), *idiličen* (612), *konformen* (51), *mentol* (23), *polacca* (3), *poloneza* (14), *radian* (24), *spinel* (8).

¹⁴ Tu gre za primer korpusnega šuma – vsi zadetki so namreč iz enega vira, in sicer iz prevoda turističnega vodnika iz angleščine. Vzrok je ponesrečena prevajalkina izbira pri premoščanju leksikalne praznine na slovenski strani.

Področne oznake ima 16 od 47 slovarsko in korpusno potrjenih leksemov, poleg tega pa korpus *FIDA* potrjuje še kar nekaj leksemov, ki so v *VST* označeni, npr. *labanotacija*, *la belle epoque*, *tabaskovka*, *zajval*. To kaže, da je v *VST* vendarle več splošnega besedja kot terminov in da domneva o prekrivnosti označenega besedja in v korpusu *FIDA* nepotrjenih leksemov ne drži.

Kot merilo za določanje, ali leksem sodi med besedje, ki je običajno obravnavano v splošnih slovarjih, je bil postavljen pogoj vsaj 50 zadetkov v korpusu *FIDA*. Ti leksemi so predstavljeni v diagramih 5 in 6 – vsi imajo tudi potrditev v *SSKJ* in/ali *SP* 2001.

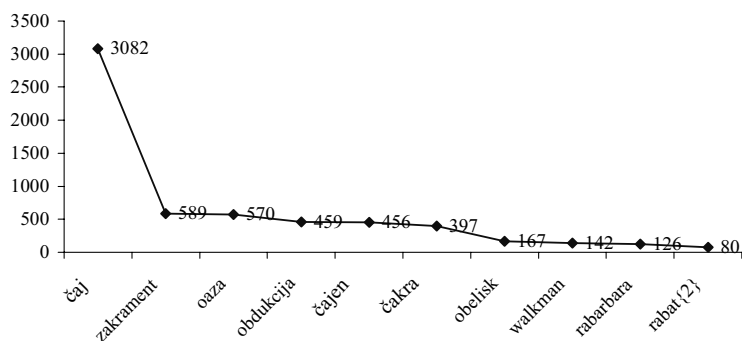


Diagram 5: Zadetki v korpusu *FIDA* s pogostnostjo nad 50 – osnovni vzorec.

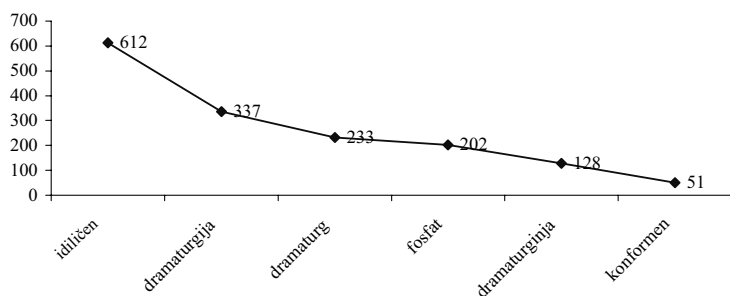


Diagram 6: Zadetki v korpusu *FIDA* s pogostnostjo nad 50 – kontrolni vzorec.

V osnovnem vzorcu ima 10 enot (20,8 %) pogostnost višjo od 50, v kontrolnem pa 6 (50 %). Diagrama 7 in 8 kažeta razmerja med frekvenčnimi skupinami v korpusu potrjenih leksemov iz obeh vzorcev. Osnovni vzorec ima izrazito veliko skupino leksemov, ki se v korpusu *FIDA* pojavijo od enkrat do desetkrat (50 %). V teh primerih gre običajno za citatno rabo, npr. *fabliau* (potrjena je tudi citatna raba množinske oblike *fabliaux*), *la belle epoque*, *obligato*, *zabaione*, *zaibatsu*. V kontrol-

nem vzorcu ima ta skupina za polovico manjši delež, kar lahko deloma pripišemo številčno manjšemu vzorcu, deloma pa drugačnemu načinu vzorčenja; prav tako pa je potrjena citatna raba (*polacca*). Vzorčenje je verjetno vzrok tudi za drugačna razmerja v skupinah z višjo pogostnostjo, kjer so deleži v kontrolnem vzorcu od dvakrat pa do skoraj petkrat višji od tistih v osnovnem vzorcu.

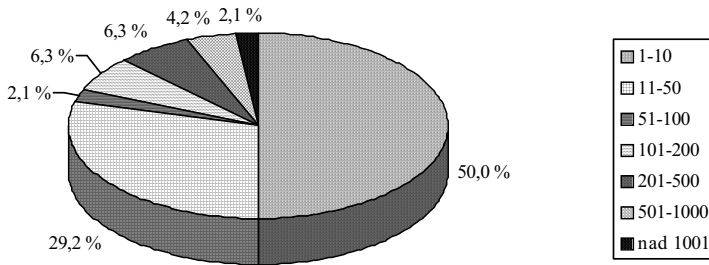


Diagram 7: Razmerje zadetkov v korpusu FIDA po skupinah – osnovni vzorec.

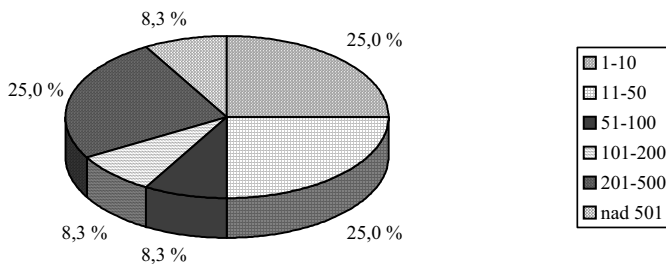


Diagram 8: Razmerje zadetkov v korpusu FIDA po skupinah – kontrolni vzorec.

5 Sklep

Želja po prikazu velikega števila področij pri t. i. prenovi *Velikega slovarja tujk* je prvotni načrt sestavljavcev spremenila v skoraj izključno zajemanje terminološkega gradiva, da bi ustregli čim večjemu številu potencialnih uporabnikov. Glede na 84 različnih področnih oznak, uporabljenih v *VST*, in 50-odstotno označenost obeh vzorcev bi bil naslov lahko kar *Veliki terminološki slovar*. Pri pregledu vzorca po korpusu *FIDA* se je sicer izkazalo, da so s področnimi oznakami opremljeni tudi nekateri pogosto rabljeni leksemi, ki so del splošnega jezika. Vse navedeno je po-

sledica črpanja iz različnih specializiranih (terminoloških) virov ter nedodelane in neusklajene uredniške politike, hkrati pa kaže na problem obravnavanja splošnega besedja v specializiranih slovarjih, kjer je opremljanje s področnimi oznakami za uporabnika zavajajoče. (Enako velja v obratni smeri, tj. pri navajanju terminologije v splošnih slovarjih.) Te probleme bi bilo mogoče rešiti oz. v znatni meri zmanjšati, če bi bili sestavljavcem slovarjev na Slovenskem poleg referenčnega korpusa na voljo tudi specializirani podkorpusi, npr. terminološki. Poleg možnosti leksikografske izrabe so taki korpusi uporabni tudi pri ustvarjanju novega izrazja v posameznih strokah, zelo koristni so za lektorje, večjezični (vzporedni) pa tudi za prevajalce. Zlasti pomembna je vloga korpusov pri izločanju oz. zmanjševanju subjektivnega elementa in uzaveščanju dejanske rabe pri omenjenih dejavnostih.

Na vprašanje, ali je v današnjem času koncept slovarja tujk, kakršen je obravnavani, še sploh aktualen, težko odgovorimo pritrdilno: terminologijo posameznih področij obravnavajo eno-, dvo- ali večjezični terminološki slovarji, etimologija je stvar specialnih slovarjev (v anglo-ameriškem jezikovnem okolju tudi splošnih slovarjev večjega obsega), splošno besedje pa pomensko obravnavajo splošni slovarji. Morda bi bil smiseln priročnik, ki bi vseboval le tiste tujke oz. citatne besede in besedne zveze, ki so v slovenščini res pogosto rabljene, a jih v splošnih enojezičnih slovarjih ni. Geslovnik za tak leksikografski priročnik bi moral nujno izhajati iz korpusa, zasnova bi morala biti skrbno premišljena, uredniška politika pa dosledna.

V razmislek založnikom, ki zaradi lagodnosti in relativno majhnih stroškov večinoma izdajajo ponatise ali pa »posodobljene« izdaje slovarjev, je potrebno povedati, da so v svetu že več kot desetletje v veljavi novi leksikografski standardi. Poleg tega so uporabniki – tudi zaradi uporabe interneta in dostopnosti leksikografskih virov v elektronski obliki – vedno bolj zahtevni in vprašanje je, koliko časa bodo še pripravljeni listati po obsežnih slovarjih, kjer morda niti ne bodo našli iskanega podatka.

Viri

Korpus slovenskega jezika FIDA <<http://www.fida.net>>.

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

Slovenski pravopis, 2001. Ljubljana: Založba ZRC.

Snoj, Marko, 1997: *Slovenski etimološki slovar*. Ljubljana: Mladinska knjiga.

Tavzes, Miloš in Adlešič, Gregor (ur.), 2002: *Veliki slovar tujk*. Ljubljana: Cankarjeva založba.

Verbinc, France, 1968: *Slovar tujk*. Ljubljana: Cankarjeva založba.

Literatura

- Anić, Vladimir in Goldstein, Ivo (ur.), 2000: *Rječnik stranih riječi*. Zagreb: Novi Liber.
- Dose, Maria (ur.), 1990: *Duden Fremdwörterbuch*. Mannheim [etc.]: Dudenverlag.
- Gorjanc, Vojo, 1999a: Kaj in kako v korpus *FIDA*? *Razgledi* 13. 7–8.
- Gorjanc, Vojo, 1999b: Korpusi v jezikoslovju in korpus slovenskega jezika *FIDA*. Kržišnik, Erika in Lokar, Meta (ur.): *Slovenski jezik, literatura in kultura tukaj in zdaj. 35. seminar slovenskega jezika, literature in kulture: zbornik predavanj*. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovanske jezike in književnosti Filozofske fakultete. 47–59.
- Gorjanc, Vojo, 2000: Nekatere možnosti jezikoslovne izrabe enojezikovnih korpusov. Orel, Irena (ur.): *Čas v slovenskem jeziku, literaturi in kulturi. 36. seminar slovenskega jezika, literature in kulture: zbornik predavanj*. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovanske jezike in književnosti Filozofske fakultete. 335–348.
- Grabnar, Katarina in Šorli, Mojca, 2003: Novi veliki angleško-slovenski slovar Oxford-DZS. *Jezik in slovstvo* 48/3–4 (maj–avg. 2003). 126–133.
- Jakopin, Primož, 2001: Slovenski nacionalni korpus – idejni osnutek projekta. *Jezikoslovni zapiski* 7. 411–417.
- Klun, Boris, 2003: V naglici ne more nastati dober slovar. *Delo* 45/74. 8.
- Krek, Simon, 1999: Računalniški korpusi v slovaropisju. *Razgledi* 13. 8.
- Krek, Simon, 2003: Jezikovni priročniki in novi mediji. *Jezik in slovstvo* 48/3–4 (maj–avg. 2003). 29–46.
- Landau, Sydney I., 1989: *Dictionaries: The Art and Craft of Lexicography*. Cambridge University Press.
- McEnery, Tony in Wilson, Andrew, 2001: *Corpus Linguistics*. Second Edition. Edinburgh University Press.
- Stabej, Marko, 1998: Besedilnovrstna sestava korpusa *FIDA*. *Uporabno jezikoslovje* 6. 96–106.
- Stabej, Marko, 1999: Korpus *FIDA*: nam računalnik lahko pokaže (slovenski) jezik? *Razgledi* 13. 6–7.
- Toporišič, Jože, 1992: *Enciklopedija slovenskega jezika*. Ljubljana: Cankarjeva založba.
- Vojnovič, Nastja, 2001: Geslovník. *Slovarska delavnica*. <<http://www.zrc-sazu.si/pravopis/delavnica.htm>>.
- Weiss, Peter, 2001: Slovenski nacionalni korpus Maks na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU: utemeljitev. *Jezikoslovni zapiski* 7. 419–428.