

UDK 811.163.6'374

Izvorni znanstveni rad

Rukopis primljen 1. III. 2021.

Prihvačen za tisk 21. V. 2021.

doi.org/10.31724/rihjj.47.2.13

**Mija Michelizza**

ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša

Novi trg 2, SI-1000 Ljubljana

orcid.org/0000-0001-9741-600X

mija.michelizza@zrc-sazu.si

## MED TIPIČNIM IN STEREOTIPNIM: DRUŽBENA OBČUTLJIVOST PRI PRIPRAVI ILUSTRATIVNEGA GRADIVA V eSSKJ

V prispevku *Med tipičnim in stereotipnim: družbena občutljivost pri pripravi ilustrativnega gradiva v eSSKJ* obravnavamo tipične kolokacije, ki so bile v leksikografskem procesu zaradi stereotipnosti ali politične nekorektnosti iz Slovarja slovenskega knjižnega jezika, tretja izdaja (eSSKJ) izločene, pri čemer so ostale ohranjene v slovarki bazi, ob izvozu slovarja za objavo na portalu Fran pa se ne izpišejo. Izpostavljamo štiri skupine tovrstnih tipičnih kolokacij: prve se pojavljajo pri leksemih, ki poimenujejo osebe, zlasti ranljive skupine, druge so kolokacije z zaznanimi nacionalnimi stereotipi, tretja skupina so nekorektne kolokacije pri leksemih, ki poimenujejo živali, in četrta nekorektne kolokacije pri žaljivi, vulgarni leksiki ipd. V prispevku tudi pojasnimo, zakaj se nekateri stereotipni elementi v slovar vseeno uvrščajo, ter opozorimo na pomen beleženja tovrstnih elementov, ki se je pokazal skozi proces dela, po začetnih redakcijah slovarja eSSKJ.

### 1. Uvod<sup>1</sup>

Stereotipni elementi v splošnih razlagalnih slovarjih so bili zaradi specifične situacije v začetku 21. stoletja v slovenski znanstveni in strokovni literaturi že razmeroma dobro raziskani, posledično se je v istem obdobju začelo tudi obširneje opozarjati na pomen družbene občutljivosti v slovaropisju. Glavni aktualni slovarski vir je takrat še vedno predstavljala prva izdaja Slovarja slovenskega

---

<sup>1</sup> Prispevek je nastal v okviru programa P6-0038, ki ga financira ARRS.

knjižnega jezika (dalje SSKJ); ta je – gledano z današnjega vidika – slovar z mnogimi stereotipnimi in politično nekorektnimi slovarskimi elementi (npr. Gorjanc 2004 in 2005, Kern 2015), ki pa so bili v splošnem značilni za večino slovarjev, ki so nastali v istem času kot omenjeni slovar. Kot je pokazala npr. primerjava slovenskega SSKJ in poljskega slovarja *Inny Słownik Języka Polskiego* (Bednarska 2013), tovrstne elemente zasledimo tudi v poljskem enojezičnem razlagalnem slovarju, kar pa seveda ne pomeni, da na take neustreznosti s sodobnega vidika ni (bilo) treba opozarjati. Na pomen družbene občutljivosti v leksikografiji opozarjajo tudi mnogi novejši priročniki (npr. Béjoint 2000: 120–124, Atkins in Rundell 2008: 425–426 idr.). Leta 2007 je bil stereotipom v slovenskem jeziku, literaturi in kulturi posvečen 43. seminar slovenskega jezika, literature in kulture, kjer so bili med drugim na primeru SSKJ predstavljeni tudi nekateri stereotipi v slovarju in o njem (Rozman in Holz 2007). Gre torej za obdobje, v katerem je prišlo do razmeroma velike diskrepance med duhom časa in duhom časa nastanka takrat edinega aktualnega, a z vidika politične korektnosti, kot jo pojmuje danes, na več mestih neustreznega slovarja SSKJ. Politična korektnost je postajala vedno bolj aktualna tema, zato se je tudi o družbeno nesprejemljivih vidikih v jeziku in slovarju – slednji je bil in je pogosto še vedno dojet kot avtoriteta – začelo množično govoriti in pisati, tako v znanstvenih krogih kot tudi v publicistiki. Opozorila jezikovnih uporabnikov, da je treba družbeno ranljive skupine in vse, kar je z njimi v povezavi, v slovarjih prikazovati korektno, še vedno zelo pogosto prihajajo na Inštitut za slovenski jezik Frana Ramovša ZRC SAZU. Vse to izkazuje, da je družbena občutljivost v slovarjih pomemben vidik, ki ga je ob pripravi novih jezikovnih priročnikov treba temeljito premisliti in s posebno občutljivostjo tudi udejanjati.

Za razliko od obdobja po izidu prve izdaje SSKJ, ko so bile na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU aktivnosti v zvezi s splošnim razlagalnim slovarjem v veliki meri preusmerjene v pripravo pravopisnega slovarja, pa je bilo zadnje desetletje z vidika slovenske splošne leksikografije zelo plodno. Leta 2012 je namreč izšel Slovar novejšega besedja slovenskega jezika (dalje SNB), leta 2014 je prišlo do dopolnitve in delne prenove slovarja SSKJ – tako imenovana druga izdaja Slovarja slovenskega knjižnega jezika (dalje SSKJ2), istega leta je nastal slovarski portal Fran, na katerem je začel redno letno izhajati Sprotni slovar slovenskega jezika (dalje Sprotni slovar), z letom 2016 pa še tretja izdaja Slovarja slovenskega knjižnega jezika (dalje eSSKJ).

Kern (2015) je z vidika politične korektnosti analiziral slovarja SNB in SSKJ2 ter prikazal, da so novejši slovarji za slovenščino ta vidik pomembno izboljšali, dejstvo pa je, gre za zelo občutljivo temo, ki se je moramo uredniki v najširšem smislu zavedati v vseh fazah nastajanja slovarja – od sestave geslovnika, začetnih analiz gradiva in redakcije do končnih pregledov. S podobnimi težavami, ki si jih bomo v nadaljevanju tudi podrobneje ogledali, se ukvarjajo in jih navajajo tudi pri pripravi hrvaškega spletnega slovarja Hrvatski mrežni rječnik (dalje Mrežnik) (Hudeček in Mihaljević 2018). Tendenco k družbeno odgovornim leksikografskim rešitvam z vidika stereotipov o spolih v eSSKJ je predstavila Ledinek (2019, prim. še Ledinek in Michelizza 2021), na vidik družbene občutljivosti pa velja opozoriti ne samo pri splošnih razlagalnih slovarjih, ampak tudi pri drugih vrstah slovarjev. O tej problematiki sta v kontekstu terminoloških slovarjev za slovenščino pisala Trojar in Žagar Karer (2013).

V pričujočem prispevku se bomo na stereotipne slovarske elemente osredotočili z vidika leksikografskega dela pri eSSKJ. S pomočjo analize kolokacij, ki so bile zaradi različnih razlogov izločene, vendar za potrebe arhiva v slovarski bazi ohranjene, bomo preverili, kako se kažejo presojanja leksikografov o primernosti uslovarjanja sicer tipičnih, a hkrati z vidika družbene občutljivosti in politične korektnosti neprimernih kolokacij. Dotaknili se bomo tudi potrebe po občutljivosti leksikografov pri uslovarjanju kolokacij, ki se ob analizi besedilnega korpusa in sorodnih orodij prikazujejo kot tipične, torej statistično relevantne, vendar pa so dejansko slovarsko neustrezne, zlasti z vidika politične nekorektnosti.

## 2. O slovarju eSSKJ

Analiza prispevka temelji na redakcijskem delu pri novem enojezičnem splošnem razlagalnem slovarju za slovenščino – tretji izdaji Slovarja slovenskega knjižnega jezika, imenovanem tudi eSSKJ, ki nastaja na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU in je dostopen na portalu Fran (Ahačič, Perdih in Ledinek 2015, Perdih 2020). Slovar eSSKJ je izdelan povsem na novo, jezikovni opis knjižnega jezika izhaja iz analize sodobnega jezikovnega gradiva in torej ne nastaja kot prenova katerega od obstoječih slovarjev (Gliha Komac idr.

2015: 2). Posebej opozarjamo, da ima tretja izdaja navkljub imenu slovarja, ki je enako kot pri prvih dveh izdajah, povsem drugačen koncept kot njeni predhodnici. Z vidika nastajanja gre za *rastoči slovar*, kar pomeni, da se deli slovarja (tj. določeno število razmeroma<sup>2</sup> končanih slovarskih sestavkov) v rednih časovnih intervalih objavljajo na spletu, čeprav slovar še ni v celoti končan.<sup>3</sup> Enkrat letno je objavljen t. i. *prirastek slovarja*, pri čemer gre za slovarske sestavke, ki so bili redakcijsko končani in pregledani v preteklem letu od zadnje objave (Ledinek in Michelizza 2020).

Podlaga za vsak leksikografski opis je dobro strukturiran gradivski vir. To je veljalo že pri nastajanju prvega splošnega razlagalnega slovarja za slovenščino SSKJ, ki se je pri nastajanju slovarja zgledoval po praški leksikografski šoli in ki je imel za gradivo tedaj zavidljivih 3155 enot, ki so jih ekscerpirali takratni sodelavci Inštituta za slovenski jezik.<sup>4</sup> Gradivo je bilo urejeno kot listkovna kartoteka z več kot 6 milijoni listkov, kar v korpusni terminologiji pomeni 6 milijonov današnjih konkordanc, za katere pa je treba poudariti, da so bile premišljeno izbrane in urejene, šum je bil torej že izločen. Gradivski vir pri slovarju eSSKJ je besedilni korpus Gigafida, ki je bil v zadnjih letih nadgrajen v Gigafido 2.0 (Krek idr. 2016). Pri redakcijskem delu sta uporabljana oba korpusa; prvi je bil vir za pripravo besednih skic, drugi pa je v uporabi ob pregledu naključnih konkordanc.

## 2.1. Ilustrativno gradivo v eSSKJ

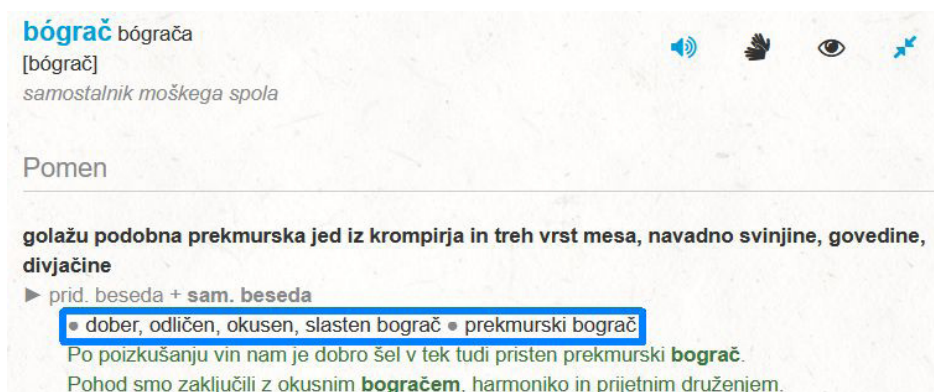
Ilustrativno gradivo v slovarju eSSKJ je predstavljeno s tipičnimi kolokacijami in besedilnimi zgledi. Tipične kolokacije so razvrščene glede na skladijske strukture, sledijo jim besedilni zgledi, ki so zgledi v obliki celotne povedi. Slednji so sicer lahko navedeni tudi samostojno, če analiza gradiva ne pokaže za iztočnico tipičnih struktur. Ilustrativno gradivo je v slovarju navedeno pri vsakem

<sup>2</sup> Zaradi kasnejših usklajitvenih pregledov in poenotenj v prvih letih po objavi še lahko prihaja do manjših sprememb znotraj že objavljenih slovarskih sestavkov (prim. Gliha Komac idr. 2015: 4).

<sup>3</sup> Ko govorimo o rastočih slovarjih, je treba poudariti, da včasih končanost niti ni cilj rastočega slovarja. Tak primer je Sprotni slovar slovenskega jezika, ki je enako kot eSSKJ objavljen na portalu Fran, prav tako se enkrat letno posodablja, nadgrajuje zlasti z novejšim, še neustaljenim besedjem slovenskega jezika, vendar pa nima predvidenega končnega obsega slovarja (več o Sprotnem slovarju prim. Krvina 2021).

<sup>4</sup> Šele leta 1986 se je Inštitut za slovenski jezik po svojem prvem upravniku akademiku dr. Franu Ramovšu preimenoval v Inštitut za slovenski jezik Frana Ramovša (Hajnšek-Holz 1992: 125).

pomenu in podpomenu, prav tako pa tudi pri pomenih in podpomenu frazeoloških in drugih stalnih besednih zvez. Pri analizi se bomo osredotočili na tipične kolokacije, kot je prikazano in označeno na Sliki 1 pri iztočnici *bograč*.



Slika 1. Prikaz tipičnih kolokacij v slovarju eSSKJ

Kolokacije so glede na Koncept (Gliha Komac idr. 2015: 52) v slovarju eSSKJ razumljene kot »statistično prepoznavne sopojavitve najmanj dveh besed, ki predstavljajo vmesni člen med (najbolj) prostimi besednimi zvezami na eni strani in stalnimi zvezami različnih tipov na drugi strani. V slovarju izkazujejo tipične in za govorce sprejemljive sopojavitve najmanj dveh besed – nasproti neobičajnim oz. manj sprejemljivim sopojavitvam«. Pomemben poudarek pri razumevanju pojmovanja kolokacij v slovarju eSSKJ z vidika obravnavane teme je, da gre za statistično sopojavitev, pri čemer pa mora redaktor v gradivu vedno potrditi, da gre za tipično rabo, ki mora biti za govorce tudi sprejemljiva. S sprejemljivostjo seveda lahko zaobjamemo različne vidike, gotovo pa vključuje tudi družbeno korektnost.

## 2.2. Slovarska baza slovarja eSSKJ in atributi za izpis

Prva redakcija leksema v eSSKJ se vedno začne z natančnim pregledom korpusnega gradiva – naključnih konkordanc in besednih skic. Verjetno vsak leksikograf, ki sodeluje pri tako kompleksnih slovarjih, kot je npr. eSSKJ, slej ali prej opazi, da si načela, ki so bila postavljena v začetku, do določene mere lahko nasprotujejo in tako začnejo nastajati izjeme; od teh so nekatere lahko tudi

regularne, spet druge pa se lahko v slovarju pojavijo zgolj pri enem primeru ali dveh. Takšen trk načel pri prikazu tipičnega in stereotipnega je, kot smo videli, v Konceptu slovarja (Gliha Komac idr. 2015) deloma predviden, saj veleva sprejemljivost uslovarjenih kolokacij, vendar pa vemo, da se stereotipni elementi trdovratno ohranjajo prav v jeziku in da nekatere stereotipne elemente težko označimo za družbeno popolnoma nesprejemljive, spet druge včasih kot take sploh težko prepoznamo, tretji pa se morda kot tipični kažejo zaradi neuravnoveženosti deležev določenih besedil v besedilnem korpusu. Kot zlasti izrazito opazamo prevlado publicistike v korpusih Gigafida in Gigafida 2.0. Glede na podatke o besedilnotipski sestavi korpusa, ki jih povzemamo s spletne strani korpusa Gigafida 2.0 (<https://viri.cjvt.si/gigafida/System/About>, dostop 25. februar 2021) je besedil časopisov 47,8 % in revij 16,5 %, kar je skupaj 64,3 %, pri čemer lahko med publicistiko uvrstimo tudi del internetnih besedil. Ta so sicer glede na predstavljeno besedilnotipsko sestavo korpusa obravnavana posebej, dejstvo pa je, da so v skupino internetnih besedil uvrščeni tudi novičarski portali, v sklopu katerih prav tako prevladujejo publicistična besedila. Internetnih besedil je v Gigafidi 2.0 28 %, leposlovja pa 3,5 %, 3,8 % predstavljajo stvarna besedila in 0,3 % drugo.

Slovarska baza slovarja eSSKJ nastaja v programu iLex in je zapisana v označevalnem jeziku XML, strukturirana pa je v skladu z internim standardom in na način, da je lahko pretvorjena v katerikoli standardni format za zapis leksikalnih podatkovnih zbirk, npr. TEI (Perdih in Ledinek 2019: 125). Redaktorji smo že zelo kmalu opazili, da bomo v slovarski bazi potrebovali možnost, da se nekatere enote, ki so sicer jezikovno relevantne, v slovarju iz različnih razlogov ne izpišejo; nekatere so (pretirano) stereotipne, druge so kot neustrezne označili terminološki svetovalci, spet tretje lahko odlično potrjujejo kakšno slovnično lastnost, vendar so z vsebinskega vidika popolnoma neustrezne ipd., zato smo že pred prvo objavo slovarskega prirastka leta 2016 pri različnih elementih strukture XML v slovarski bazi uvedli atribut za izpis *ne*. Kasneje se je izkazalo, da zaradi politično nekorektnih vnosov potrebujemo še nekoliko močnejši atribut, ki smo ga poimenovali kot *neprimerno*, funkcijonira pa na enak način kot atribut za izpis *ne*, kar pomeni, da se pri elementih strukture XML z omenjenima atributoma pri izvozu slovarske baze za prikaz na portalu Fran vsebina elementa ne izpiše.

### 3. Analiza izločanja kolokacij

Analiza izločenih kolokacij je bila izvedena in se nanaša na stanje v interni bazi slovarja eSSKJ v programu iLex 9. in 10. decembra 2020. Pregledanih je bilo 412 že objavljenih slovarskih sestavkov, pri čemer seznam vključuje enobesedne iztočnice, stalne besedne zveze in frazeološke enote, pri katerih je bila vsaj ena zabeležena kolokacija označena z atributoma za izpis *ne* ali *neprimerno*. Atribut za izpis *neprimerno* smo uvedli pri pregledih tik pred zadnjo objavo leta 2019, zato smo v analiziranem gradivu našli le manjše število takih elementov (zlasti pri besedilnih zgledih), zagotovo pa bo ta atribut zelo koristen v prihodnje, tudi za nadaljnje analize, podobne pričujoči, saj je bilo pri pregledovanju precej dela z izločevanjem za uslovarjenje drugače neprimernih kolokacij, ki so bile v veliki večini.

Analiza je pokazala, da lahko zaradi stereotipnosti in politične nekorektnosti izločene kolokacije razvrstimo v 4 skupine, in sicer (1) nekorektne kolokacije pri leksemih, ki poimenujejo osebe, zlasti ranljive skupine (spolni, socialni, ekonomski, družbeni stereotipi), (2) kolokacije z zaznanimi nacionalnimi stereotipi, (3) nekorektne kolokacije pri leksemih, ki poimenujejo živali, in (4) nekorektne kolokacije pri žaljivi, vulgarni leksiki ipd.

#### 3.1. Nekorektne kolokacije pri leksemih, ki poimenujejo osebe

##### Mladoletnica

Pri analizi besedne skice leksema *mladoletnica* (Slika 2) smo uredniki opazili, da so tipične kolokacije in večina besedilnih zgledov do te mere neprimerni, da smo sprejeli konceptualno odločitev in v tovrstnih primerih povsem izločili slovarsko navajanje tipičnih kolokacij, ki pa jih v bazo vseeno vnesemo in elemente strukture XML za kolokacije označimo z atributom za izpis *neprimerno*. Smernice za redakcijo v takih primerih so, da z največjo mero družbene občutljivosti izberemo nekaj sprejemljivih besedilnih zgledov (Slika 3), pri čemer se trudimo zaobjeti vse vidike – pozitivne, nevtralne in hkrati tudi negativne; za ustrezno ponazoritev pa lahko izjemoma posežemo tudi po drugih gradivskih virih. Odločitev je zabeležena tudi v internem Redakcijskem priročniku.

mladoletnica <small>(samostalnik)</small>			Gigafida frekvenca = 1,778 (1.26 na milijon)											
WBZ_gbz	1,615	0.80	sbz_SBZ2	307	4.60	sbz_z_SBZ	193	18.40	SBZ2_gbz	80	4.00	sbz_pri_SBZ	15	6.90
posiliti	26	4.20	posilstvo	36	10.58	seks	36	8.90	potekati	4	1.95	nosečnost	6	10.88
splaviti	6	3.51	prostitucija	14	10.28	občevanje	5	8.08	gbz_z_SBZ	71	6.80	splav	4	10.71
otipavati	7	3.49	ličinka	27	9.72	škandal	7	7.50	seksati	8	7.64	sbz_nad_SBZ	9	18.10
rojiti	4	2.64	nadlegovanje	9	8.29	odnos	77	5.57	dobivati	4	5.78	nasilje	5	4.70
nadlegovati	11	2.60	splav	4	7.81	afera	5	5.10	spati	4	5.34			
posiljevati	5	2.19	zlorabljanje	6	7.70	razmerje	8	4.63	družiti	9	5.31			
izsiljevati	9	1.91	zloraba	28	5.73	gbz_SBZ4	167	3.40						
seksati	9	1.86	pbz_SBZ	220	1.00	posiliti	24	9.25	pbz2_SBZ2	46	0.90			
zlorabljati	8	1.40	zlorabljen	4	5.75	zahajati	4	7.55	mlad	4	0.49			
zlorabiti	7	1.28	noseč	6	5.51	izsiljevati	8	6.99						
zanositi	5	1.24												

Slika 2. *Mladoletnica* v besednih skicah korpusa Gigafida

**mladolétnica** mladolétnice  
[mladolétnica]  
samostalnik ženskega spola

Pomen

**1. ženska, ki ne dosega zakonsko določene starosti, navadno osemnajst let, in zato še nima polnih pravno določenih pravic ter dolžnosti**  
Po prvih zbranih obvestilih naj bi **mladoletnica** od doma pobegnila s fantom, ker njeni starši njune zveze niso odobraval.

Novi kazenski zakonik bo pomembno prispeval k učinkovitemu boju zoper spolno zlorabljanje in izkoriščanje **mladoletnic** in mladoletnikov.

Le kako je dobro vzgojena, pridna hčerka in učenka osnovne šole postala tako problematična **mladoletnica**?

Slika 3. *Mladoletnica* v slovarju eSSKJ

Poleg temeljnega argumenta, da skušamo pri tovrstni leksiki slovarske sestavke pripraviti z veliko mero družbene občutljivosti, odločitev za prilagojen slovarski prikaz pri takih leksemih utemeljujemo tudi z dejstvom, da so zaradi velike količine publicističnih besedil v korpusih Gigafida in Gigafida 2.0 besedila črne kronike (v katerih se, sodeč po analizi, leksem *mladoletnica* najpogosteje pojavlja) zastopana nesorazmerno pogosto, kar smo že ponazorili tudi s predstavitvijo besedilnotipske sestave korpusa Gigafida 2.0. Poleg tega povsem drugačno sliko v gradivu pri leksemu *mladoletnica* kaže tudi pregled interne listkovne kartoteke Inštituta za slovenski jezik Frana Ramovša ZRC SAZU, na podlagi katere je nastajala prva izdaja SSKJ, čeprav se zavedamo časovne distance. Od 12 zgledov je samo eden izrazito negativen in povezan z zlorabo, dva zgleda po-



udarjata problematični vidik teh oseb (problemi in poboljševanje), sedem zgledov je z današnjega vidika povsem korektnih, dva zгледа pa sta zaradi zapisa z roko težko berljiva, zaradi česar nismo mogli z gotovostjo opredeliti konteksta.

Podoben primer je leksem *mladoletnik*, ki pa se od leksema *mladoletnica* razlikuje v tem, da je besedna skica pokazala nekoliko bolj pestro sliko, zato smo kot redakcijsko dopolnili tudi v tovrstnih primerih sklenili v veliki meri omejiti navajanje kolokacij (Slika 4), ne pa povsem izločiti, kot to velja v primeru *mladoletnice*. Ob analizi namreč ne smemo pozabiti tudi na to, da se leksem ob osebah moškega spola zlasti v množini pogosto rabi tudi spolno nevtrarno. Tudi ta konceptualna odločitev je zapisana v internem Redakcijskem priročniku.

The screenshot shows the dictionary entry for 'mladolétnik' in the eSSKJ. The entry includes the word 'mladolétnik', its grammatical information '[mladolétnik]' and 'samostalnik moškega spola'. Below this is the section 'Pomen' (Meaning), which states: 'kdor ne dosega zakonsko določene starosti, navadno osemnajst let, in zato še nima polnih pravno določenih pravic ter dolžnosti'. It also lists related terms like 'n-letni mladoletnik', 'nasilen mladoletnik', 'obsojen, osumljen mladoletnik', 'pijan, vinjen mladoletnik', and 'pogrešan, poškodovan, ranjen mladoletnik'. A green note mentions a case where a severely injured teenager was transported to a hospital. Another green note states that among delinquent teenagers, many do not return from the system.

Slika 4. *Mladoletnik* v slovarju eSSKJ

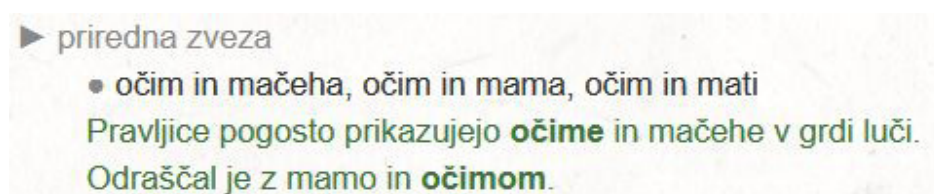
## Očim

Pri slovarskem sestavku leksema *očim* sta v bazi kot neprimerni zabeleženi kolokaciji *očim zlorablja koga* in *nasilen očim*, ki kasneje v slovar nista bili uvrščeni. Ob pregledu besedne skice opazimo, da je bilo z vidika korektnosti gradivo že ob prvi redakciji prečiščeno (Slika 5), ker pa pri redakciji tega spornega vidika, ki se kaže v gradivu, nismo želeli povsem prikriti, smo kot stavčni zgled vključili tudi primer, kjer je stereotip, da očimi slabo ravnajo z otroki (navadno) nove partnerice, ki izhaja tudi iz nekaterih pravljičnih besedil, postavljen v kontekst (prim. prvi besedilni zgled na Sliki 6).

**očim** *(samostalnik)*  
Gigafida frekvenca = 2,708 (1.92 na milijon)

WBZ_gbz	2,689	0.90	SBZ1_gbz	727	8.20	priredje	540	4.40	pbz_SBZ	420	1.40	SBZ2_gbz	213	7.60
zlorabljeni	78	4.67	zlorabljeni	43	8.34	mačeha	88	11.33	sadističen	10	7.45	umoriti	12	7.09
pretepati	27	3.37	posiliti	10	7.67	polbrat	28	9.96	neobčutljiv	4	7.20	ovaditi	5	5.89
umoriti	38	3.13	pretepati	9	6.82	pastorek	13	9.28	brezvesten	5	6.71	prijaviti	11	5.62
posiliti	11	2.93	posvojiti	6	6.17	polsekstra	6	7.87	dekletov	4	6.10	oprostiti	8	5.57
trpinčiti	4	2.80	mamiti	4	5.98	mama	108	7.54	nasilen	26	5.29	ustreliti	6	5.55
posiljevati	6	2.41	zagrešiti	5	5.15	mati	85	7.26	speč	4	5.15	ubiti	7	5.04
posvojiti	11	2.17	obnesti	6	4.90	luka	7	6.01	brutalen	5	4.66	zagledati	4	4.77

Slika 5. *Očim* v besednih skicah korpusa Gigafida



Slika 6. *Očim* v slovarju eSSKJ

## Pešec

Nekorektne kolokacije smo našli tudi pri leksemu *pešec*, ki sam po sebi sicer ne sodi med družbeno občutljivo leksiko, a raba vseeno kaže na določeno nekorektnost, ki se v besedilnem korpusu izkazuje zaradi problema neuravnoteženosti gradivskih virov, ki smo ga že omenjali pri primeru *mladoletnice*. Vidimo, da izločene kolokacije (Slika 7) pomensko niso zelo drugačne od uslovarjenih (Slika 8), vendar pa je bilo v besedni skici tipičnih kolokacij, ki so si bile med seboj zelo podobne, enostavno preveč, da bi jih na tak način predstavili v slovarju. Podobna primera sta še leksema *žleb* in *vetrovka*, kjer prva asociacija pri človeku zagotovo ni v povezavi s črno kroniko, vendar pa se je pokazalo, da se ta dva leksema dovolj pogosto pojavljata v besedilih črne kronike, da korpusna analiza to prikaže kot tipično. *Ukrasti žleb* je ena najbolj pogostih kolokacij v korpusu pri leksemu *žleb*, *vetrovka* pa je pogosto uporabljena za opis roparja ali pogrešane osebe. Zaradi tovrstne neuravnoteženosti in prevelike zastopanosti besedil črne kronike v analiziranem besedilnem korpusu je redatorsko delo zagotovo oteženo in posledično dolgotrajnejše.

kolokacija `lepis="ne":` **mimoidoči pešec**  
 kolokacija `lepis="ne":` **nesrečni pešec**  
 kolokacija `lepis="ne":` **neznani pešec**  
 kolokacija `lepis="ne":` **odrasli pešec**  
 kolokacija `lepis="ne":` **povožen pešec**  
 kolokacija `lepis="ne":` **vinjen pešec**

Slika 7. *Pešec* v slovarski bazi eSSKJ

► prid. beseda + sam. beseda  
 • *n*-letni pešec • na čem ležeči pešec • mrtvi pešci • nepreviden pešec • poškodovan, ranjen pešec • star pešec  
 Strojevodja je neprevidnega **pešca**, ki je hodil tik ob progi, opozarjal z zvočnim signalom.  
 Ne da bi se zmenil za mimoidoče **pešce** in avtomobile, hiti skozi gnečo.

Slika 8. *Pešec* v slovarju eSSKJ

### 3.2. Kolokacije z zaznanimi nacionalnimi stereotipi

#### Česen

Pri slovarskem sestavku za leksem *česen* smo našli izločeno kolokacijo *kitajski česen* (Slika 9). Na prvi pogled bi lahko mislili, da gre za posebno sorto česna, vendar pa ob pregledu gradiva vidimo, da je bila kolokacija ustrezno izločena, saj se pridevnik nanaša na izvor česna, besedilni zgledi pa so v kontekstu povečane rabe pesticidov, nezakonitega uvoza in radiokativnosti (Slika 10).

kolokacija: **jesenski, spomladanski, zimski česen**  
 kolokacija: **mlad/mladi česen**  
 kolokacija `lepis="ne":` **kitajski česen**

Slika 9. *Česen* v slovarski bazi eSSKJ

cenim, sicer pa francoski krompir, **kitajski česen** špansko čebulo in italjansko zelenjavo. Na žič in grozdje spomladi, skupaj s **kitajskim česnom**. A kdor ima za te stvari nekaj občutka, ovorimo, in seveda - zloglasnem **kitajskem česnu**. Menda naredimo več zase, če ga sploh. Italijanski krompir, madžarska moka, **kitajski česen**, pakistanski fižol, španski paradižnik, grah. Danes boli čisto drugače od **kitajskega česna**, čudno zrelih banan, zgornjih oreškov piščanic. Foto: EPA Tihotapljenje poceni **kitajskega česna** v EU je v porastu in prav prek Norveške.

Slika 10. Konkordance zveze *kitajski česen* v korpusu Gigafida

## Žajfnica

Izpostavimo še primer kolokacije *mehiška žajfnica*, ki je bila v procesu leksikografske obravnave leksema *žajfnica* prav tako izločena (Slika 11). Ker pa smo se odločili, da ne želimo prikrivati dejstva, da telenovele pogosto nastajajo prav v Mehiki, smo informacijo ohranili, vendar na manj izpostavljenem mestu, in sicer v besedilnem zgledu (Slika 12).

kolokacija: televizijska žajfnica  
kolokacija **epi="ne"**: mehiška žajfnica

Slika 11. *Žajfnica* v slovarski bazi eSSKJ

► sam. beseda + v + sam. beseda v mestniku  
• vloga v žajfnici  
Sprejela je glavno vlogo v novi **žajfnici** v slogu mehiških nadaljevanj, ki bo vsak dan na programu nacionalne televizije.

Slika 12. *Žajfnica* v slovarju eSSKJ

### 3.3. Nekorektne kolokacije pri leksemih, ki poimenujejo živali

#### Maček

Pri slovarskem sestavku leksema *maček* je bila za izpis odstranjena kolokacija *mučiti mačke* (Slika 13). Naj poudarimo, da z neuvrščanjem takšnih kolokacij, ki opisujejo družbeno nesprejemljiva dejanja, ne želimo prikrivati ali izkrivljati realnosti. Pravzaprav je vprašanje, ki se zastavlja na tem mestu, do kakšne mere

prikazano realnost izkrivlja korpus s tem, ko to prikazuje kot tipično, kar se je že pokazalo tudi v nekaterih predhodnih analiziranih primerih. Obenem menimo, da je treba biti v slovarju – zlasti pri iztočnicah, ki z družbeno nesprejemljivimi dejanji niso neposredno povezane – s tovrstnimi prikazi zelo previden, saj se je v raziskavah že večkrat pokazalo, da uporabniki slovar pojmujejo kot avtoriteto, hkrati pa pričakujejo, da bo opis odražal nekakšen ideal, h kateremu kot družba stremimo (Gorjanc 2005: 197–199; Kern 2015: 146).

kolokacija: imeti mačka  
 kolokacija: najti mačka  
 kolokacija: videti mačka  
 kolokacija: eps="ne" mučiti mačke

Slika 13. *Maček* v slovarski bazi eSSKJ

### 3.4. Nekorektne kolokacije pri žaljivi, vulgarni leksiki ipd.

#### Hinavski, hlapec in gobezdalo

S pojavom spleta in vključenostjo spletnih besedil v besedilne korpuse (Michelizza 2015: 47–48) se zaradi anonimnosti zelo širi sovražni govor, nastrojen proti različnim družbenim skupinam. Da so tovrstna besedila problematična, so zaznali tudi sestavljavci korpusa Gigafida in v Gigafidi 2.0 izločili besedila z mnogimi nestandardnimi elementi, med drugim tudi uporabniške komentarje z novičarskih portalov (Krek idr. 2020: 3342). Običajno politično korektnost povezujemo z ranljivimi skupinami, ki navadno nimajo velikega družbenega vpliva ali finančnih sredstev, da bi sami lahko učinkovito ščitili svoj položaj. Med analizo izločenih kolokacij v bazi slovarja eSSKJ pa se je izkazalo, da do statistične zaznavnosti v besednih skicah zelo pogosto pridejo tudi politično nekorektne kolokacije, povezane s politiki in politikom. Čeprav nas v duhu časa lahko hitro zavede, da bi se s tovrstnimi kolokacijami ali celo z žaljivimi zgledi, ki so namenjeni politiki ali politikom, poistovetili, pa moramo biti tudi pri uslovarjanju tovrstnih – širše družbeno morda celo sprejemljivih – stereotipov zelo previdni in družbeno občutljivi. Take primere kolokacij najdemo pri redakcijah

leksemov *hinavski* – kolokacija *hinavska politika* (Slika 14), *hlapec* – kolokacija *režimski hlapec* (Slika 15) in *gobezdalo* – zlasti kolokacija *politično gobezdalo* (Slika 16).

kolokacija: **hinavski način**  
kolokacija: **hinavski nasmeh, nasmešek**  
kolokacija: **hinavski obraz**  
kolokacija: **hinavska morala**  
kolokacija **izpis="ne": hinavska politika**  
kolokacija: **hinavsko sprenevedanje**

Slika 14. *Hinavski* v slovarski bazi eSSKJ

kolokacija: **okupatorjev hlapec**  
kolokacija: **ponižen hlapec**  
kolokacija **izpis="ne": režimski hlapec**

Slika 15. *Hlapec* v slovarski bazi eSSKJ

kolokacija **izpis="ne": napihnjeno, veliko gobezdalo**  
kolokacija **izpis="ne": navadno gobezdalo**  
kolokacija **izpis="ne": politično gobezdalo**  
kolokacija **izpis="ne": znano gobezdalo**

Slika 16. *Gobezdalo* v slovarski bazi eSSKJ

### 3.5. Uslovarjene stereotipne kolokacije

#### Baba

Med analizo izločenih nekorektnih kolokacij smo že nakazali, da poskušamo v nekaterih primerih nekorektno in stereotipno realnosti prikazati na drugačen, običajno manj vpadljiv način, saj vseeno gre za tipično rabo. So pa v slovarju seveda tudi mejni primeri, ko uredniki in pregledovalci zelo tehtamo, kako prikazati ilustrativno gradivo. Tak primer so kolokacije pri leksemu *baba* (Slika 17), ki je v večini primerov zagotovo stilno zaznamovan, pogosto do stopnje slabšalnosti, v gradivu pa se najdejo tudi primeri, kjer je leksem uporabljen šaljivo, npr.: *Saj se moramo babe tudi kdaj pa kdaj malo sprostiti in poveseliti*. Zdi se, da gre za specifične primere, kjer lahko leksem s prevladujoče slabšalnimi pomeni oseba, na katero se leksem nanaša, uporabi v šaljivem tonu. Težava nastane, ko je treba takšne in podobne specifične značilnosti leksemov prikazati znotraj predvidene slovarske strukture. V podobnih primerih se leksikografi vedno posvetujemo v večji skupini, upoštevamo širši kontekst, pozorno analiziramo besedilne zglede in tehtno premislimo o številu tovrstnih uslovarjenih tipičnih kolokacij in besedilnih zgledov. Pri leksemu *baba* in kolokacijah, ki smo jih uredniki in pregledovalci izločili, vidimo, da so morda za stopnjo bolj žaljivi od uslovarjenih (npr. *nora baba* in *zmešana baba*, ki sta uslovarjeni, nasproti kolokacijam *histerična baba* in *hudobna baba*, ki sta izločeni), vendar tudi uslovarjene kolokacije ostajajo na stopnji žaljivosti, saj se je kot takšna pokazala tipična raba leksema *baba*.

kolokacija: blond baba  
kolokacija: grda baba  
kolokacija: nage babe  
kolokacija: nora, zmešana baba  
kolokacija: sitna, tečna, zoprna baba  
kolokacija: stara baba  
kolokacija izpis="ne": debela baba  
kolokacija izpis="ne": histerična baba  
kolokacija izpis="ne": hudobna baba  
kolokacija izpis="ne": pijana baba

Slika 17. *Baba* v slovarski bazi eSSKJ

### Črne mačke prinašajo nesrečo

Stereotipnost v večji meri ohranjamo npr. v frazeologiji. Navajamo primer frazeološke enote *Črne mačke prinašajo nesrečo*, kjer je že v sami razlagi pojasnjeno, da gre za vraževerje, ki z realnostjo nima nobene zveze (Slika 18).

#### Črne mačke prinašajo nesrečo.

*Črna mačka prinaša nesrečo.*

**izraža vraževerno prepričanje, da je srečanje s črno mačko slab znak, da ima lahko negativne posledice**

Ali je res, da **črne mačke prinašajo nesrečo**, kadar prečkajo cesto pred vami? Zagotovo jo prinašajo. Same sebi.

Slika 18. Frazeološka enota *Črne mačke prinašajo nesrečo* v slovarju eSSKJ



## 4. Razprava

V analizi smo pod drobnogled vzeli primere tipičnih kolokacij slovarja eSSKJ, za katere smo leksikografi ob redakciji ali pregledih pred objavo opazili, da je njihovo sprejetje v slovar (lahko) sporno. Pokazalo se je, da je bilo treba navkljub načelnemu opozorilu v Konceptu slovarja (Gliha Komac idr. 2015: 52) ta vidik glede na korpusno gradivo pri posameznih leksemih obravnavati še bolj podrobno, saj smo si leksikografi med seboj zelo različni, zato tudi tovrstne elemente občutimo in posledično obravnavamo vsak nekoliko drugače. Treba je bilo sprejeti dodatna načela obravnave te leksike, ki smo jih opredelili v internem Redakcijskem priročniku, in v slovarski bazi urediti vnose z uvedbo atributa za izpis *ne* in *neustrezno*. Kot smo že v uvodu omenili, so se z zelo podobnimi težavami ukvarjali tudi pri hrvaškem enojezičnem razlagalnem slovarju Mrežnik. Slovarja Mrežnik in eSSKJ imata tudi sicer precej podobnosti v smislu problematičnosti gradivskih virov, terminski začetki obeh slovarjev so skoraj sovpadli, nenazadnje pa gre tudi za sorodna jezika, zaradi česar so težave uslovarjanja podobnih leksemov podobne, seveda pa se koncepta obeh slovarjev tudi razlikujeta v predvidenem obsegu, tipu razlag, načinu navajanja tipičnih kolokacij idr. (Hudeček in Mihaljević 2018).

V raziskavi izločenih tipičnih kolokacij v slovarski bazi eSSKJ se je pokazalo, da sicer leksiko, pri kateri opazamo zadržke leksikografov, lahko razdelimo v 4 skupine, pri katerih sta pomembni politična korektnost in družbena občutljivost, in sicer (1) pri leksemih, ki poimenujejo osebe, zlasti ranljive skupine; (2) pri leksemi, pri katerih smo zaznali nacionalne stereotipe; (3) pri leksemih, ki poimenujejo živali ter (4) pri žaljivi in vulgarni leksiki. Pereča problematika prikaza spolskih parov,<sup>5</sup> pri čemer moški par pogosto označuje osebo ne glede na spol, in prevelikega števila stereotipnih elementov, zaznanih z analizo gradiva, zlasti pri feminativih se kaže tako v eSSKJ (Ledinek 2019, Ledinek in Michelizza 2021) kot tudi v Mrežniku (Hudeček in Mihaljević 2018). Pri slovarju Mrežnik poleg družbeno odgovorne izbire ilustrativnega gradiva problematiko rešujejo tudi s posebno pragmatično opombo (Hudeček in Mihaljević 2018: 108, Hudeček 2021). Slovar eSSKJ pa je postavljen v kontekst slovarskega portala Fran, kjer sobiva

<sup>5</sup> O jezikoslovni terminologiji, povezani z besedotvorjem ženskih in moških parov v hrvaščini, prim. Hudeček in Mihaljević 2019.

z drugimi slovarji ter Jezikovno svetovalnico (Dobrovoljc, Lengar Verovnik in Bizjak Končar 2018) in tudi Terminološko svetovalnico (Žagar Karer 2017). V slednjih dveh je pogosto mesto tudi za tovrstne daljše pojasnitve rabe, seveda pa so pojasnitve odvisne od vprašanj, ki jih postavijo jezikovni uporabniki.

Kot se je torej pokazalo v analizi izločenih tipičnih kolokacij, gre v veliki meri za pričakovane problematične skupine leksemov, vendar pa se je potrdilo tudi to, da moramo poleg družbeno ranljivih skupin korektno predstavljati tudi družbene skupine, ki jim ranljivosti ne pripisujemo – celo nasprotno, največ nekorektnih kolokacij se je pokazalo pri skupini, ki ji pripisujemo velik družbeni vpliv; skrajno negativna podoba se je namreč pokazala pri tipičnih kolokacijah, povezanih s politiki in politiko. Hkrati je treba še opozoriti, da lahko pride do nekorektnih elementov pri kateremkoli leksemu, kot se je pokazalo na primeru *česna* in tipične kolokacije *kitajski česen*.

Mestoma so v eSSKJ zavestno uslovarjeni nekateri stereotipni elementi, zlasti pri slabšalni leksiki tudi med tipičnimi kolokacijami (tak leksem je seveda s kvalifikatorjem tudi ustrezno označen), pri nevtralni leksiki pa je tendenca, da se tovrstnim elementom v slovarju v večji meri izognemo, pri leksiki, povezani z najranjlivejšimi skupinami, včasih celo do te mere, da tipičnih kolokacij ne navajamo in za ilustrativno gradivo izberemo le nekaj korektnih besedilnih zgledov. V določenih primerih so stereotipi nakazani v besedilnem zgledu, kjer so lahko hkrati postavljeni v širši kontekst. Stereotipnost pa je v večji meri ohranjena pri frazeoloških enotah, kjer je ta element pojasnjen v razlagi, kot smo prikazali na primeru frazema *Črne mačke prinašajo nesrečo*. Glede na obravnavane primere lahko povzamemo, da je najtežji del redakcijskega procesa zagotovo ugotavljanje ločnice med tipično in/ali stereotipno rabo leksemov ter iskanje ustreznega prikaza tovrstnih elementov v slovarskem sestavku, zlasti zato ker je pri taki obravnavi leksike prisoten tudi subjektivni faktor.

Vsekakor je nujno, da je vidik družbene občutljivosti upoštevan pri vseh fazah leksikografskega dela, pri čemer se zavedamo, da gre za večplastno problematiko in zato morda v slovarju eSSKJ ostajajo primeri, ki se komu zdijo politično nekorektni in pri katerih s pričujočo analizo tega nismo opazili. S sprotno objavo slovarskih sestavkov na portalu Fran spodbujamo tudi uporabnike slovarja k povratnim informacijam, čeprav so ob spletni objavi vsi slovarski sestavki poleg več uredniških pregledov deležni še terminološkega pregleda pri izter-

minološki leksiki in korektur ter so zato, kot je zapisano v Konceptu slovarja (Gliha Komac idr. 2015: 4), »referenčni in razmeroma dokončni«.

## 5. Zaključek

Odgovorna izbira ilustrativnega gradiva zlasti pri družbeno občutljivi leksiki in tudi sicer upravičeno postaja pomemben segment leksikografskega dela pri veliko evropskih slovarskih projektih, kar je pokazala tudi okrogla miza *Rječnik i stereotipi* v Zagrebu decembra 2020. Pri izkušnjah ob pripravi prvih prirastkov slovarja eSSKJ bi zlasti opozorili na koristnost beleženja stereotipnih, tudi politično nekorektnih elementov v slovarskih bazah, ki sicer običajno ostajajo le za interno uporabo in se v slovarju kasneje ne izpišejo, lahko pa služijo za raziskave, podobne pričujoči, in odstirajo pretekle odločitve v zvezi z uslovarjanjem družbeno občutljive leksike, ki so lahko v pomoč pri prihodnjih podobnih zadregah.

## Literatura

AHAČIČ, KOZMA; LEDINEK, NINA; PERDIH, ANDREJ. 2015. Portal Fran – nastanek in trenutno stanje. *Simpozij Obdobja 34, Slovnica in slovar – aktualni jezikovni opis*. Ur. Smolej, Mojca. Znanstvena založba Filozofske fakultete. Ljubljana. 57–66.

ATKINS, B. T. SUE; RUNDELL, MICHAEL. 2008. *The Oxford Guide to Practical Lexicography*. Oxford University Press. New York.

BEDNARSKA, KATARZYNA. 2013. Kdo naj dela, kdo pa skrbi za otroke? Spolni stereotipi v slovarjih. *Družbena funkcijnost jezika (vidiki, merila, opredelitve)*. Ur. Žele, Andreja. Znanstvena založba Filozofske fakultete (Obdobja, 32). Ljubljana. 35–42.

BÉJOINT, HENRI. 2000. *Modern Lexicography. An Introduction*. Oxford University Press. Oxford.

DOBROVOLJC, HELENA; LENGAR VEROVNIK, TINA; BIZJAK KONČAR, ALEKSANDRA. 2018. Spletna jezikovna svetovalnica za slovenski jezik ter normativne zadrege uporabnikov in jezikoslovcev. *Zbornik prispevkov s simpozija 2017. Škrabčevi dnevi 10*. Nova Gorica. 52–69.

GLIHA KOMAC, NATAŠA; JAKOP, NATAŠA; JEŽOVNIK, JANOŠ; KLEMENČIČ, SIMONA; KR-VINA, DOMEN; LEDINEK, NINA; MIRTIC, TANJA; PERDIH, ANDREJ; PETRIC, ŠPELA; SNOJ,

- MARKO; ŽELE, ANDREJA. 2015. *Koncept novega razlagalnega slovarja slovenskega knjižnega jezika*. ISJFR ZRC SAZU. Ljubljana. doi.org/10.3986/9789612548391.
- GORJANC, VOJKO. 2004. Politična korektnost in slovarski opisi slovenščine – zgolj modna muha? *Moderno v slovenskem jeziku, literaturi in kulturi: zbornik predavanj. Seminar slovenskega jezika, literature in kulture 40*. Ur. Stabej, Marko. Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete. Ljubljana. 153–161.
- GORJANC, VOJKO. 2005. Neposredno in posredno žaljiv govor v jezikovnih priročnikih: diskurz slovarjev slovenskega jezika. *Družboslovne razprave XXI/48*. 197–209.
- HAJNŠEK-HOLZ, MILENA. 1992. Inštitut za slovenski jezik Frana Ramovša Znanstveno-raziskovalnega centra SAZU. *Jezik in slovstvo 37/5*. 125–130.
- HUDEČEK, LANA; MIHALJEVIĆ, MILICA. 2018. Croatian web dictionary Mrežnik: One year later – What is different? *Jezikovne tehnologije in digitalna humanistika 2018*. Ur. Fišer, Darja; Pančur, Andrej. Znanstvena založba Filozofske fakultete. Ljubljana. 106–113.
- HUDEČEK, LANA; MIHALJEVIĆ, MILICA. 2019. Hrvatsko mocijsko nazivlje. *Jezikoslovni zapiski 25/1*. 107–126. doi.org/10.3986/jz.v25i1.7569.
- HUDEČEK, LANA. 2021. Rječnik i stereotipi – Kakva je plavuša u Mrežniku? *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje 47/2*. doi.org/10.31724/rihjj.47.2.9.
- KERN, BORIS. 2015. Politična korektnost v slovaropisju. *Škrabčevi dnevi 8*. Ur. Zuljan Kumar, Danila; Dobrovoljc, Helena. Založba Univerze v Novi Gorici. Nova Gorica. 144–154.
- KREK, SIMON; GANTAR, POLONA; ARHAR HOLDT, ŠPELA; GORJANC, VOJKO. 2016. Nadgradnja korpusov Gigafida, Kres, ccGigafida in ccKres. *Zbornik konference Jezikovne tehnologije in digitalna humanistika*. Ur. Erjavec, Tomaž; Fišer, Darja. Znanstvena založba Filozofske fakultete. Ljubljana. 200–203.
- KREK, SIMON; ARHAR HOLDT, ŠPELA; ERJAVEC, TOMAŽ; ČIBEJ, JAKA; REPAR, ANDRAŽ; GANTAR, POLONA; LJUBEŠIĆ, NIKOLA; KOSEM, IZTOK; DOBROVOLJC, KAJA. 2020. Gigafida 2.0: The Reference Corpus of Written Standard Slovene. *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*. Marseille, 11–16 May 2020. 3340–3345.
- KRVINA, DOMEN. 2021. Sprotni slovar slovenskega jezika, covid-19 in z njim povezano (novejše) besedje. *Zbornik Leksikografija i leksikologija u svetlu aktuelnih problema*. Ur. Ristić, Stana; Lazić Konjik, Ivana; Ivanović, Nenad. Institut za srpski jezik SANU. 660–680.
- LEDINEK, NINA. 2019. Spol v slovenskih enojezičnih splošnih razlagalnih slovarjih kot slovnici in družbena kategorija. *Slavistična revija*, letnik 67/2019, št. 2, april–junij. 189–201.

LEDINEK, NINA; MICHELIZZA, MIJA. 2020. Enojezični splošni razlagalni slovarji slovenščine in spodbude za njihovo rabo v srednješolskih učbenikih za slovenski jezik. *Slovenski jezik in književnost v srednjeevropskem prostoru. Zbornik Slavističnega društva Slovenije 30*. Ur. Šekli, Matej; Rezoničnik, Lidija. Zveza društev Slavistično društvo Slovenije. Ljubljana. 255–266.

Ledinek, Nina; Michelizza, Mija. 2021. Gender in Slovenian Monolingual General Explanatory Dictionaries. *Collegium Antropologicum 45/4*. 329–339. doi.org/10.5671/ca.45.4.5.

MICHELIZZA, MIJA. 2015. *Spletna besedila in jezik na spletu*. Založba ZRC. Ljubljana. doi.org/10.3986/9789610504108.

PERDIH, ANDREJ; LEDINEK, NINA. 2019. Multi-word Lexical Units in General Monolingual Explanatory Dictionaries of Slavic languages. *Slovenski jezik – Slovene Linguistic Studies 12*. 113–134. doi.org/10.3986/sjsls.12.1.07.

PERDIH, ANDREJ. 2020. Portal Fran: Od začetkov do danes. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovje 46/2*. 997–1018. doi.org/10.31724/rihjj.46.2.28.

ROZMAN, TADEJA; HOLZ, NANIKA. 2007. Slovar – jezika vladar? *Stereotipi v slovenskem jeziku, literaturi in kulturi: zbornik predavanj*. 43. seminar slovenskega jezika, literature in kulture. Ur. Novak Popov, Irena. Univerza v Ljubljani, Filozofska fakulteta, Oddelek za slovensistiko, Center za slovenščino kot drugi/tuji jezik. Ljubljana. 25–35.

TROJAR, MITJA; ŽAGAR KARER, MOJCA. 2013. Družbena občutljivost v terminoloških slovarjih. *Družbena funkcijnost jezika: (vidiki, merila, opredelitve). Simpozij Obdobja 32*. Ur. Žele, Andreja. Znanstvena založba Filozofske fakultete. Ljubljana. 457–463.

ŽAGAR KARER, MOJCA. 2017. Terminološko svetovanje: izkušnje in perspektive. *Slovenska terminologija danas: primljeno na V. skupu Odeljenja jezika i književnosti od 30. maja 2017. godine, na osnovu referata akademika Predraga Pipera i prof. dr. Rajne Dragičević*. Ur. Piper, Predrag; Jovanović, Vladan. Beograd: Srpska akademija nauka i umetnosti: Institut za srpski jezik SANU. 361–370.

## **Between the Typical and the Stereotypical: Social Sensitivity in the Preparation of Examples of Use in eSSKJ**

### *Abstract*

In this paper, we discuss the social sensitivity of lexicographers in the third edition of the *Dictionary of the Slovenian Standard Language – eSSKJ* in the case of excluding typical collocations that lexicographers perceived as politically incorrect or otherwise potentially controversial during their editorial work. The analysis has shown that in most cases these are expected collocations which are especially related to the most vulnerable

social groups; however, it has turned out that in principle non-vulnerable social groups should also be portrayed in a correct way, given that the corpus material actually shows a very negative picture in typical collocations related to politics and politicians. Even in this case, care must be taken that the lexicographic presentation is correct. In addition, controversial collocations can also occur in vocabulary where one would not necessarily think of incorrectness, e.g. *kitajski česen* (Eng. Chinese garlic). Some stereotypical elements are consciously included in the eSSKJ dictionary, especially when it comes to very typical uses – but these examples are particularly sensitively selected and usually presented in less exposed places, e.g. in textual examples with a broader context; if presented in typical collocations, those that are less offensive are selected, and in the case of phraseological units, the stereotypical aspect is usually explained in the definition itself. Despite the general warning in the Dictionary Concept that the typical collocations must be acceptable for speakers, it has turned out that additional principles of treatment should be adopted when dealing with this type of vocabulary, and the attributes *ne* (Eng. no) and *neprimerno* (Eng. inappropriate) were added to the dictionary database soon after the first dictionary entries were made. Those two attributes mark elements (in our case for typical collocations) which are not exported to the Fran portal, but remain in the database and allow research such as the present one.

### **Između tipičnog i stereotipnog: društvena osjetljivost u pripremi primjera upotrebe u eSSKJ-u**

#### *Sažetak*

U ovom radu razmatramo društvenu osjetljivost leksikografa u trećem izdanju *Rječnika slovenskog književnog jezika* – eSSKJ u slučaju izuzimanja tipičnih kolokacija koje su leksikografi tijekom uredničkog rada doživljavali kao politički nekorektne ili na neki drugi način potencijalno kontroverzne. Analiza je pokazala da je u većini slučajeva riječ o očekivanim kolokacijama koje su posebno povezane s najranjivijim društvenim skupinama, ali se pokazalo i da društvene skupine koje nisu ranjive također treba prikazati na ispravan način; u slučaju korpusnog materijala pokazuje se vrlo negativna slika u tipičnim kolokacijama koje su povezane s politikom i političarima. U ovom slučaju također treba paziti da su rječnički članci urađeni korektno. Nadalje, kontroverzne kolokacije mogu se u rječniku pojaviti i kod natuknica kod kojih se to ne bi očekivalo, npr. kod natuknice *kitajski česen* (hrv. kineski češnjak). Neki stereotipni elementi svjesno su uključeni u rječnik eSSKJ, posebno kada je riječ o vrlo tipičnim upotrebama, ali ti su primjeri osobito pozorno odabrani i obično predstavljeni na manje istaknutim mjestima, npr. u tekstualnim primjerima sa širim kontekstom; ako su

predstavljeni u tipičnim kolokacijama, odabiru se oni koji se mogu označiti kao manje uvredljivi, a u slučaju frazeoloških jedinica, stereotipni aspekt obično se objašnjava već u definiciji. Unatoč općenitom upozorenju u koncepciji rječnika da tipične kolokacije moraju biti prihvatljive za govornike, ispostavilo se da bi se u ovoj vrsti rječnika trebali usvojiti dodatni principi postupanja: atributi *ne* i *neprimereno* (hrv. neprikladno) dodani su u bazu rječnika ubrzo nakon prvog izdanja rječnika. Elementi označeni tima dvama atributima (u našem slučaju za tipične kolokacije) ne izvoze se na portal Fran, već ostaju u bazi podataka kao arhiva i omogućuju istraživanje poput ovoga.

**Ključne besede:** slovar, leksikografija, eSSKJ, stereotipi, kolokacije, družbena občutljivost

**Keywords:** Dictionary, Lexicography, eSSKJ, Sterotypes, Collocations, Social Sensitivity

**Ključne riječi:** rječnik, leksikografija, eSSKJ, stereotipi, kolokacije, društvena osjetljivost