

(SLOVENSKA) LEKSIKA MED LEKSIKONOM IN SLOVNICO

V prispevku obravnavamo področje, ki smo ga znotraj jezikoslovja označili kot »sivo polje« med leksikalno oz. pomensko in skladenjsko problematiko jezika. Pri tem izhajamo iz spoznanja, da leksikalnega vedenja ni mogoče ločiti od slovničnega in da so v tem smislu leksikalne enote v enaki meri obremenjene s pomenskim in skladenjskim potencialom jezika. Na teh izhodiščih ter z izdelavo besednih skic za slovenščino, ki delujejo znotraj sistema *SketchEngine*, predlagamo postopke za izrazno in pomensko prepoznavanje večbesednih leksikalnih enot v obsegu besedilnega okolja, ki je za to odločilno.

1 Uvod

1.1 Leksikon in slovnica kot ločeni komponenti jezika

Čprav je z današnjega vidika leksikon¹ (leksika, besedišče) razumljen kot celoten leksikalni inventar, ki obsega najrazličnejše informacije o jeziku, je v preteklosti izraz v prvi vrsti pomenil knjigo, slovar, ki vsebuje le najpomembnejše podatke o besedah (Johnson 1755). Izhodišče, po katerem je leksika v nasprotju s slovnico kot zbirom pravil za oblikovanje smiselnih in slovnično pravilnih stavkov, razumljena kot zbir leksikalnih enot (LE) in njihovih lastnosti, se je uveljavilo z jezikoslovno teorijo Noama Chomskega (1965). Na tej podlagi generativna slovnica razume

¹ Izraz *leksikon* uporabljamo tu še v generativnem smislu kot zbir leksikalnih enot in njihovih pomenov, ki obstaja ob slovnici kot samostojni del jezika. Sicer pa danes navadno s tem izrazom označujemo zbir besednih oblik v jeziku, ki so jim po možnosti pripisane tudi oblikoslovne vrednosti.

leksiko(n) kot neodvisno komponento jezika: kot zbir neregularnih enot, ki jih ni mogoče generirati z obstoječimi slovničnimi pravili. Po mnenju Chomskega po slovničnih pravilih tvorjene zveze ne morejo biti predmet leksikona, saj naj bi bil ta namenjen zgolj shranjevanju idiosinkratičnih enot (Chomsky in Halle 1968: 12). Kot taka je bila leksika za preučevanje človeškega jezika manj zanimiva in pojmovana kot obrobni jezikovni pojav.

1.2 Jezikoslovje med leksiko in skladnjo

Na podlagi spoznanja, da ločnice med slovničnimi pravili in besedami, ki jih s temi pravili oblikujemo, ni mogoče jasno in dokončno potegniti, je bil predmet leksikološkega preučevanja, tj. besede, njihovi pomeni, razmerja med besedami, njihove povezovalne lastnosti itd., nadgrajen s skladenjskimi razmerji, ki se ustvarjajo med LE, zlasti z vidika vloge, ki jo opravljajo v stavku. Pojmovanje leksike ne samo kot celotnega inventarja besed v jeziku, pač pa kot kompleksnega skladenjsko-pomenskega pojava, je sedaj vsaj že 25 let v ospredju ne samo leksikoloških in leksikografskih, pač pa jezikovnih analiz nasploh. Povečanje zanimanja za to področje je v veliki meri tudi odgovor na obravnavanje leksemov kot neke vrste polnilcev neodvisnih skladenjskih modelov, ki jih predvideva slovnica.² Dejstvo namreč je, da učenje besed in znanje, kako jih učinkovito rabiti, vedno vključuje tudi učenje razumevanja in rabe njihove fonologije, pomenskega potenciala, njihove slovnice in morfologije, pregibnostnega in sintagmatskega potenciala ter v večini jezikov tudi njihove skladnje (De Mauro 2006: 19). Sodobne jezikoslovne teorije, zlasti tiste, ki se naslanjajo na računalniške potenciala, kot bomo na kratko predstavili v nadaljevanju, vse bolj dojemajo leksiko kot centralno skladišče jezikovnega vedenja (Ooi 1998: 2).

1.2.1 Besedna povezovalnost (kolokabilnost)

S pritegnitvijo vidika besedne povezovalnosti v analizo medbesednih razmerij se je postavilo vprašanje obsega pomensko zadostne besede kot LE. Že ob koncu petdesetih let 20. stoletja so se nasledniki J. R. Firtha, McIntosh, Halliday, Sinclair idr., v nasprotju z leksikološkimi pristopi, ki so temeljili na (ne)razstavljalnosti, komponentni analizi in pomenskih razmerjih med LE, posvetili preučevanju besedne povezovalnosti, ki ni bila izključno vezana na ugotavljanje razstavljalnosti besednega pomena. Pozornost je bila v prvi vrsti namenjena distribuciji besed v besedilu in ugotavljanju predvidljivosti pogostejšega sopojavljanja nekaterih besed z enimi kot z drugimi. S tem pa so bila dana tudi izhodišča za pojmovanje večbesednih leksikalnih enot (VLE) kot statistično predvidljivega in hkrati semantično pogojenega pojava.

² Tako izhodišče dobro ponazarja Bloomfieldova izjava (1933: 274) o leksikonu, ki je le seznam osnovnih posebnosti/nepravilnosti in kot tak slovnici le dodan.

1.2.2 Korpusi in leksikologija

V vsej razsežnosti je preučevanje besedne povezovalnosti prišlo do izraza z razvojem korpusnega jezikoslovja, kjer je mogoče s statistikami izluščiti to, kar izstopa glede na regularnost vzorca, pomena, besedilnega okolja ipd. Eno bistvenih izhodišč korpusnega jezikoslovja je, če si izposodimo besede Wolfganga Teuberta (1999: 298), da so naravni jeziki na enak način zaznamovani z analogijo kot z anomalijo in da iskanje univerzalnih struktur slovnice in leksikona pod vplivom generativne slovnice in kognitivne semantike ne more zadostiti omenjenima vidikoma. Ob tem se še posebej izpostavlja dejstvo, da se jezik na dokazljiv način manifestira samo v besedilih in da nam od jezika ločena interpretacija ali formalne izpeljave ne morejo osvetliti njegovega naravnega bistva. Korpusnemu jezikoslovju tako prej ustreza koncept evropskega strukturalizma Ferdinanda de Saussurja v smislu, da je pomen neločljivo povezan z izrazom, tj. z besedo, zvezo in besedilom, s čimer usmerja jezikoslovčev pozornost na vzorce in strukture pomenske kohezije, ki se ustvarja v razponu med posamezno besedo in stavčno zgradbo, definirano s skladijskimi pravili (Teubert 1999: 298–300). Kot posebej zanimiva vprašanja na ravni že najzgodnejših korpusnih analiz so se tako razumljivo oblikovala okoli pojma LE kot pomenskega in hkrati skladijskega jezikovnega pojava. Najbolj temeljito so se ta izhodišča udejanila v Collins COBUILD-ovih na korpusu temeljitejših slovarskih projektih, ne po naključju v slovarjih, ki so namenjeni učenju angleščine kot tujega jezika. Po prihodu teh slovarjev v slovaropisju nič več ni bilo tako kot prej: noben splošni, dvojezični ali specializirani slovar, ki je želel slediti toku časa in dostojno opravljati svoj namen, ni mogel več mimo korpusne metodologije. Naslednje korake v smeri povezovanja pomenskih in skladijskih lastnosti besed, ki temeljijo na računalniških jezikovnih bazah in se s pridom uporabljajo tako za leksikografske namene kot za procesiranje naravnih jezikov, so naredili prav Sinclairjevi učenci. Med pomembnejšimi projekti, na katerih sloni sodobna leksikografija, moramo omeniti vsaj tri: teorijo semantičnih shem oz. projekt *FrameNet Semantics*, korpusni iskalni sistem *SketchEngine* z možnostjo izdelave t. i. besednih skic (*Word Sketches*) ter opisovanje LE na podlagi analize korpusnih vzorcev (*Corpus Pattern Analysis*).³

1.2.2.1 Projekt FrameNet ali teorija semantičnih shem⁴ sega že v čas s konca šestdesetih let 20. stoletja, ko je Charles Fillmore teorijo povezovanja pomenskih in skladijskih jezikovnih vidikov predstavil v slovnici sklonov (1968: *The case for case*). V devetdesetih letih je bila teorija nadgrajena z izhodišči korpusnega in računalniškega jezikoslovja ter z možnostjo uporabe v leksikografske namene. Bistvo skladijsko-semantičnega pristopa, kot ga v opisu LE predvideva *FrameNet*, je iz obsežnih elektronskih besedilnih korpusov izločiti podatke, ki

³ Internetne strani, na katerih je mogoče izvedeti podrobnosti o omenjenih projektih in o prenosih na nekatere evropske jezike, so navedene na koncu prispevka med literaturo.

⁴ Celoten koncept je z vidika slovenščine predstavljen v prispevku Simona Kreka *FrameNet in slovenščina* v pričujoči številki *Jezika in slovtva*.

odražajo povezavo med pomenskimi in skladenjskimi lastnostmi LE ter jih formalizirati znotraj sheme (ang. *frame*). To pomeni, da je informacija o semantični valenci zajeta v udeleženski strukturi posamezne LE, skladenjska valenca pa je določena s t. i. *fraznimi tipi* (npr. samostalniška, predložna ipd. besedna zveza), ki zasedajo potencialna določila, ter s funkcijo, ki jo opravljajo v okviru stavka. Poleg povezovanja pomenskih in skladenjskih lastnosti v opisu LE so temeljne prednosti framenetovske logike v oddaljevanju od jezikoslovnih interpretacij, ki skušajo potegniti strogo ločnico med jezikom kot sistemom in jezikom kot diskurzom, in v približevanju interpretacijam, ki poudarjajo funkcijsko plat jezika (Krek 2008).

1.2.2.2 Besedne skice (*Word Sketches*) so avtomatski na korpusu temelječi povzetki slovnicega in kolokacijskega vedenja neke besede (Krek in Kilgarriff 2006). Programski modul *Sketch Engine* pa lahko sprejme korpus in predvidene slovnice vzorce kateregakoli jezika, iz katerih procesira besedne skice. Omenjeno programsko orodje deluje tudi na korpusu *FidaPLUS*, ki je oblikoslovno-skladenjsko označen, kako je mogoče besedne skice za slovenščino uporabiti za leksikografske in druge jezikoslovne namene, pa bomo prikazali v nadaljevanju.

1.2.2.3 Analiza korpusnih vzorcev (CPA) temelji na projiciranju pomena iz sobesedila na posamezno besedo in izhaja iz teorije jezikovnih konvencij in možnosti njihove izrabe (ang. *Theory of Norms and Exploitations*, Hanks 1994; Hanks in Pustejovsky 2004, 2005). Analiza korpusnih vzorcev se v veliki meri spogleduje s projektom *FrameNet*, pri čemer ponuja FN v prvi vrsti poglobljeno analizo semantičnih shem, ki se osredotoča na možnosti združevanja posameznih glagolov v sheme na podlagi njihovega podobnega skladenjskega in semantičnega vedenja v sobesedilu, končni cilj pa je povezati različne glagole v eno shemo na podlagi enakih shemskih elementov. CPA se osredotoča na sistematično analizo tipičnih pomenskih vzorcev posameznega glagola, manj pa je za ta model zanimiva možnost njihovega medsebojnega povezovanja. V postopku analize so posamezne besede (poleg glagolov so v igri še samostalniki)⁵ na podlagi primerov iz realne rabe povezane v semantično motivirane sintagmatske vzorce. Tem vzorcem se nato pripiše pomen, ki je izražen s t. i. implikaturami (gl. spodaj), na njihovi podlagi pa je mogoče postopoma oblikovati pomensko ontologijo.⁶ Identifikacija semantičnega vzorca nikakor ni preprost in avtomatiziran postopek, pač pa zahteva veliko leksikografske spretnosti in prakse: med najtežjimi leksikografskimi

⁵ Logika *FrameNeta* in CPA je v prvi vrsti usmerjena na besedne vrste, ki so znotraj stavčne strukture organizatorji pomena, tj. predvidevajo, napovedujejo in določajo lastnosti drugih elementov stavka.

⁶ Opise LE v podatkovnih zbirkah je mogoče deliti na ontološke, ki opisujejo predvsem pomenska razmerja, kot so so-, nad- in podpomenskost. Najbolj znana ontološka podatkovna zbirka je prav gotovo *WordNet*, so pa primarno te zbirke namenjene računalniškim aplikacijam. *FrameNet* in CPA sta z združevanjem pomena in skladnje bolj pisana na kožo podatkovnim bazam, ki so namenjene slovarskim in slovnicearskim projektom.

odločitvami je prav gotovo določitev ustrezne stopnje posplošitve, znotraj katere je mogoče razločevati posamezne (pod)pomene.

Za ponazoritev smo za glagol *vložiti* vzeli primer iz korpusa *FidaPLUS*: *Pivovarna Laško bo vložila pritožbo*. Pomen glagola v izkazanem kontekstu, tj. s predmetnim dopolnilom *pritožba*, vsebuje naslednje implikature: [Pivovarna Laško = institucija] bo vložila [pritožbo = dokument], ki jih lahko izrazimo takole: »če institucija vloži pritožbo, potem posreduje dokument, s katerim sproži pravni postopek«. Isti glagol v drugačnem kontekstu, npr. *V novo dejavnost smo vložili veliko denarja*, seveda zahteva drugačne implikature.

2 Metodologija in namen analize

Pri obravnavanju LE, v prispevku se omejujemo na večbesedne, smo pozornost usmerili v (so)besedilo, kjer je mogoče v medsebojni odvisnosti opazovati tako skladenjska razmerja kot pomenske lastnosti. Z izhodišči v realnih tekstih⁷ smo se želeli izogniti polnjenju sistemsko predvidljivih vzorcev s primeri, ki tem modelom ustrezajo, hkrati pa se posvetiti tistim segmentom leksike, ki z vidika slovarjev doslej niso ponudili zadovoljivih rešitev. Leksikografska praksa in leksikalne analize so namreč pokazale, da je nekatere v jeziku običajne funkcijske besede in nekatere pogoste glagole zelo težko leksikalno opisati kot izolirane enote, saj imajo le redko status neodvisnega leksikalnega elementa. Tako izhodišče širi področje raziskovanj besednozvezne problematike, ki je bilo še do nedavnega vezano predvsem na metaforične in pomensko nerazstavljive frazeološke enote (FE), tudi na nefrazeološke in neidiomatične besede in zveze, ki navadno niso deležne posebnega zanimanja, zlasti ne z vidika slovarskega opisa, ki bi ponudil uporabniku koristno in zanesljivo informacijo. V nasprotju s tem je izhodišče obravnavanja večbesedne problematike v pričujočem prispevku vezano na upoštevanje povezovalnih lastnosti besed, kot jih določajo sintagmatska in paradigmatska razmerja. Z izhodišči v omenjenih vidikih besedne povezovalnosti postavljamo v središče besedo oz. njeno osnovno obliko in ugotavljamo lastnosti, ki se kažejo v odnosu do drugih besed v njenem besedilnem okolju. Pri tem postane jasno, da sta pomen in skladenjski potencial besed in zvez odločilno porazdeljena med elemente besedilnega okolja in da teh povezav z analizami, ki temeljijo zgolj na pomenski (ne)razstavljivosti LE, ni mogoče nedvoumno opredeliti. Namen analize je v prvi vrsti leksikografski, saj želimo opozoriti na tiste besedne zveze, ki predstavljajo LE, oz. izločiti tiste elemente pomensko-skladenjskih povezav, ki so nujni za celostno slovarsko informacijo.

⁷ Analiza in vsi zgledi, ki jih navajamo v nadaljevanju, temeljijo na korpusu *FidaPLUS*.

3 Leksikalne enote in razmerja med njimi

3.1 Beseda

Slovenska slovnica (Toporišič 2000: 103–104) določa *besede* kot »najmanjše pomensko samostojne enote stavčne povedi«. Kot take so nerazdružene enote glasov oziroma črk, zaporedja njihovih zlogovnih ali pomenskih sestavin ni mogoče zamenjati, pač pa lahko kot celote zamenjujejo položaj v stavku. »Sivo polje« med eno- in večbesednimi LE je pri Toporišiču obravnavano le z vidika posebnosti pisanja skupaj oz. narazen, pri čemer gre predvsem za ohranjanje kontinuitete konkretnega pravopisnega pravila, manj pa je upoštevana dejanska jezikovna raba, ki ob posameznih primerih favorizira različne možnosti zapisa.⁸ Številčni podatek, ki ga navajamo ob spodnjih zgledih, pomeni število pojavitev v korpusu *FidaPLUS*.

<i>na vse zgodaj</i> 22	<i>kratko malo</i> 3446	<i>CD plošča</i> 465	<i>avdio kaseta</i> 110	<i>brez zveze</i> 446
<i>na vsezgodaj</i> 3	<i>kratkomalo</i> 993	<i>CD-plošča</i> 58	<i>avdiokaseta</i> 198	<i>brezzveze</i> 15
<i>navsezgodaj</i> 1806		<i>cd plošča</i> 22		<i>brezveze</i> 101
		<i>plošča CD</i> 1980		

Ob različnih možnostih, ki se nanašajo na pisanje skupaj in narazen, je problem kompleksen predvsem s pomenskega vidika: posamezne besede, ki tvorijo v kombinaciji z drugimi besedami LE, npr. *stara mama*, navadno ohranjajo v besedni zvezi tudi nekatere pomenske lastnosti, ki jih imajo sicer zunaj konkretne kombinacije. Npr. za pridevnik *star* lahko rečemo, da ohranja v zvezi *stara mama* več pomenskih lastnosti, ki jih ima (po SSKJ) v pomenu 'ki ima razmeroma veliko let' in morda še 'ki je v zadnjem obdobju življenja', kot pa pomenskih lastnosti, ki jih vsebujejo drugi pomeni istega pridevnika, izkazani v SSKJ, npr. 'ki nima več potrebne, ustrezne kakovosti za uživanje' (*star kruh*) ali 'ki je dalj časa v uporabi' (*star avtomobil*). Čeprav je razberljivost pomena sestavin pri nekaterih zvezah očitnejša kot pri drugih, se nanjo – v smislu ugotavljanja leksikaliziranosti – ni dobro zanašati: nenazadnje pomeni *stara mama* 'mamo nekogaršnjega očeta', njena dejanska starost pa je relativna.

3.2 Prosta in stalna besedna zveza ter frazeološka enota

V slovenistični teoriji je za naše izhodišče zanimiva opredelitev, po kateri so (proste) *besedne zveze* »odvisnostno ali logično povezane besede kake besedilne enotex« (Toporišič 1992: 7). Gre torej za zveze, ki jih »v govoru delamo po skladenjskih vzorcih, ne da bi jih jemali gotove iz spomina, npr. *oblačno nebo* po vzorcu pridevnik + samostalnik« (Toporišič 1992: 230). Z razumevanjem besednih zvez kot gradbenih

⁸ Problem zapisa skupaj (z vezajem ali brez njega) kot merilo besednosti postavlja tudi Logar (2005: 178) v razpravi o besedotvornih sklopih.

enot višjih skladenjskih zvez je tesno povezan tudi pojem *stavčnih fraz* (Toporišič 1982: 39, 271; 2000: 557), za katere velja, da so večbesedni ustrezniki besednih vrst. Zanje se pričakuje, da imajo iste lastnosti kot samostalniki, glagoli ipd. in opravljajo v stavku iste stavčnočlenske vloge (Toporišič 1982: 39–41). Hkrati se v jezikoslovni literaturi postavlja potreba po razmejitvi glede na tiste zveze, ki v jeziku obstajajo kot del njegovega leksikalnega fonda. Na tej podlagi se je vzpostavilo razmerje med stalnimi besednimi zvezami (SBZ) in frazeološkimi enotami (FE). Bistvena lastnost, ki naj bi FE ločevala od SBZ, je nujna prisotnosti vsaj ene sestavine, ki v FE ne nastopa v svojem slovarskem pomenu, čeprav smo zgoraj pokazali, da je identifikacija pomenske samostojnosti besednozveznih sestavin relativna in odvisna od govorčevih asociacij (*stara mama*). Kot bistvena se za vse FE v nasprotju s SBZ postavlja še ekspresivnost⁹ kot lastnost, ki vzpostavlja razmerje do nezaznamovane leksike v obliki parafraz ali enobesednih leksemov, npr. *brcniti v prazno/v temo* – *zgrešiti, spregledati bistvo; streljati kozle – delati, govoriti neumnosti; delati iz muhe slona* – *pretiravati*, čeprav je bilo že večkrat opozorjeno, da etikete ekspresivnosti nikakor ni mogoče prilepiti vsem FE ali njihovim rabam (prim. Čermák 1985: 168; Gantar 2007: 138–141). V slovenski leksikalni teoriji (Vidovič Muha 1988; 2000: 26–28) je kljub temu pojem SBZ rezerviran za tiste VLE, katerih pomen je načeloma prepoznaven iz njihovih sestavin, npr. *sončna očala, nizek pritisk, bančna kartica, bencinski servis, brezplačna telefonska številka, programski jezik* itd., hkrati pa tudi za tiste večbesedne termine, ki imajo – razen ekspresivnosti – vse značilnosti FE, npr. *črna skrinjica, taščin jezik, črna luknja, veliki pok*¹⁰ itd. Na drugi strani je frazeologija, ki si prizadeva ohraniti jasne meje glede predmeta svojega preučevanja, v raziskavah usmerjena le na tiste zveze, katerih sestavine izkazujejo t. i. *frazeološki* ali *idiomatični pomen*. Poleg tega sta za FE bistveni še večbesednost in oblikovna ter pomenska ustaljenost zveze kot celote (Moon 1998: 6). Znotraj ožje frazeološke problematike je treba omeniti še idiome, ki v odnosu do FE označujejo skrajno točko pomensko osamosvojene leksikalne in ne vedno nujno večbesedne enote. Kot taki predstavljajo idiomi pomensko samostojne besede oziroma zveze brez očitne pomenske povezave s katero koli od sestavin, ki jo tvorijo, npr. *kolumbovo jajce, (biti/iti se) gnilo jajce* ipd.

3.3 Kolokacija

Pojem *kolokacije* kot v jeziku opazne zveze dveh ali več besed, ki se pojavljajo v predvidenih slovnicih vzorcih, sopojavljanje v njih pa je pričakovano in ne naključno (Vrbinc 2001: 51), je bil v slovensko jezikoslovno teorijo sprejet ob koncu

⁹ Na tem mestu se ne bomo spuščali v vprašanje zaznamovanosti/ekspresivnosti FE, je pa v zvezi s tem potrebno opozoriti na obnašanje FE v besedilih, kjer njihovo rabo pogostokrat določajo tudi drugi elementi, kot je npr. izogibanje ponavljanju istega leksema (*ovirati – metati polena pod noge*), ustvarjanje besedilne kohezije ipd.

¹⁰ Na tem mestu puščamo ob strani vprašanje realne terminološkosti omenjenih SBZ. Distribucija zveze *veliki pok* v naravoslovnih in tehničnih besedilih oz. publikacijah predstavlja v *FidiPLUS* le slabo tretjino vseh pojavitev.

devetdesetih let 20. stoletja prek kontrastivne angleško-slovenske leksikografske problematike (Gabrovšek 1991; Vrbinc 1995; Jurko 1997). Nanj se je v tem času naslonila tudi slovenistična frazeološka teorija (Kržišnik-Kolšek 1988: 51) na podlagi Čermákovega pojmovanja kolokabilnosti, ki določa obseg povezovalnega potenciala besede glede na njene možnosti povezovanja z drugimi besedami na linearni (sintagmatski, skladijski) in vertikalni (paradigmatski, pomenski) osi (Čermák 1985: 173). Neomejenost na eni in preštevnost prvin na drugi strani, ki se znotraj povezovalnega (kolokacijskega) niza povezujejo v smiselno celoto, ponuja po Čermákovem mnenju dve temeljni množici besednih zvez: take, katerih povezovalni niz je odprt, število prvin znotraj njega pa je neomejeno, npr. *star avto, zgodba, vprašanje, peč, ura, mestno jedro, prijatelj* ... in take, katerih kolokacijski niz je praviloma zaprt, število prvin pa je preštevno oz. skrajno omejeno, npr. *stara zaveza/testament; stara celina/svet/kontinent; star maček/lisjak; star kozel/prdec; stara sablja* itd. Brisanje ostrih razlik med pomensko predvidljivimi in pomensko nepredvidljivimi kolokacijami pa je omogočilo razumevanje besedne povezovalnosti v smislu stopnjevalnosti in ne različnih tipov leksikalnih enot (Bollinger 1976).

4 Izraba korpusa pri prepoznavanju VLE in formalizaciji leksikalnega védenja

Obravnavanje LE na posameznih ravniinah jezikovne zgradbe je izpostavilo zelo različna merila, ki ne dajejo enotnih zaključkov o tem, kdaj imamo opraviti s prosto besedno zvezo in kdaj z leksikalizirano večbesedno enoto, ki hkrati ni idiomatična, ekspresivna in bolj ali manj podlega jezikovno predvidenim skladijskim procesom. Če je idiome razmeroma lahko prepoznati na podlagi skrajne pomenske neprozornosti, na lastnosti, ki bi zanesljivo in dokončno ločevale proste od stalnih in frazeoloških besednih zvez, ni mogoče računati, saj pogosto ne moremo, izhajajoč zgolj iz lastnega jezikovnega občutka, zanesljivo trditi, kdaj gre za samostojni pomen besede in kdaj je zanj odgovorna konkretna besedna kombinacija. Če vzamemo za primer pridevnik *svetel* v zvezah *svetla prihodnost* ('obetavna, perspektivna') in *svetla izjema* (SSKJ: 'pozitivna; posnemanja vredna'), se zdita pomena vsaj na prvi pogled razmeroma samostojna, hkrati pa sta si blizu glede na pozitivno vrednostno konotacijo, še posebno, če ju primerjamo z istim pridevnikom v zvezah kot: *svetla soba* ('z veliko svetlobe') ali *svetli lasje* (SSKJ: 'po stopnji barvne izrazitosti podoben dnevni svetlobi').

Védenje, da se besede povezujejo na podlagi slovnične in pomenske kompatibilnosti, daje v korpusnem okolju dobra izhodišča za prepoznavanje VLE, saj s tem, ko pravimo, da se besede med seboj povezujejo, mislimo predvsem na pogostnost njihovega sopojavljanja. Čeprav je v večini primerov najtrdnjši del VLE s primernimi korpusnimi orodji razmeroma lahko identificirati, je hkrati očitno, da je tako pomenska kot zgradbena trdnost zveze, zlasti na njenih skrajnih mejah, zrahljana, in da najtrdnjše in hkrati pogoste sopojavitve še niso

zadostni pogoj za leksikalizacijo.¹¹ Povezovalno najtrdnejše zveze besed smo določili kot *frazna jedra* (Gantar 2007: 174), zanje pa predvidevamo, da so lahko bodisi zgolj opazne kombinacije besed v jeziku, npr. t. i. najpogostejši bi- tri- in večgrami (Gorjanc 2005; Gantar 2007: 41), da lahko tvorijo leksikalno enoto bodisi v okviru svoje minimalne dvobesedne zgradbe, npr. *zgoraj brez, pasti notri*, bodisi nastopajo kot deli obsežnejših leksikalnih enot, kar velja zlasti za zveze s slovničnopomenskimi elementi, npr. *spustiti (koga/kaj) iz rok, prehajati iz rok v roke, živeti iz rok v usta, jesti (komu) iz roke* itd. Tako pojmovana frazna jedra so hkrati tudi organizatorji neposrednega besedilnega okolja in napovedujejo nadaljnje sestavine, ki oblikujejo celostno zgradbo potencialne VLE in njen pomen.

4.1 Obseg kolokacijske paradigme

Pojem *kolokacijska paradigma* povzemamo po Čermáku (1985: 171), z njim pa smo označili besede (kolokatorje), ki se na podlagi pomenskih in skladenjskih lastnosti tipično povezujejo z drugimi besedami. Pri tem izhajamo iz predpostavke, da imajo večpomenske besede, ki se znotraj posameznega pomena povezujejo z omejenim številom kolokatorjev, več možnosti za oblikovanje LE kot besede, ki imajo obsežen kolokacijski potencial. Pri odkrivanju fraznih jeder smo si pomagali s sistemom za izdelovanje besednih skic (WS). Statistični izračuni v *FidiPLUS* v kombinaciji z besednimi skicami ponujajo kompleksno informacijo o najtipičnejših soopojavitvah, stavčnem okolju in tipičnih zapolnitvah skladenjskih mest, kar predstavlja neke vrste skladenjsko-pomenski profil preučevane besede. V spodnjih seznamih so kolokatorji urejeni besednovrstno, upoštevana pa je tudi zapolnitev skladenjskega vzorca, kot se izkazuje v WS, npr. predloge *pod, pred, izpred, skozi, iz, na, za, v* ipd. smo povezali s tipičnimi glagoli. Hkrati smo na tej stopnji kolokatorje razvrstili v pomenske skupine (ang. *clusters*),¹² znotraj njih pa smo upoštevali tudi oblikoslovne variante, npr. *moleti/moliti, krvavitev/krvavenje* itd. Pri uvrstitvi v skupno pomensko polje smo upoštevali možnost metaforične povezave med besednimi pomeni, čeprav se ta v resnici udejanja šele v konkretnem sobesedilu, npr. pridevniki *dober/občutljiv/oster* tvorijo sinonimni niz samo v kombinaciji z *imeti* in *nos* (*imeti dober/občutljiv/oster nos za kaj*), zunaj te kombinacije med njimi ne obstaja nikakršna pomenska povezava. Kolokatorje, ki jih ni bilo mogoče vključiti v katero od pomenskih skupin, smo umestili v samostojni stolpec (prim. *rdeč, umeten; napudrati* itd.).

¹¹ Na splošno je leksikalizacija pojmovana kot proces, v katerem se niz besed in/ali morfemov sčasoma ustali in razvije lastni pomen in/ali funkcijo (Moon 1998: 36). Drugače oz. z leksikografskega vidika bi bilo mogoče reči, da je leksikalizacija pojav v dejavni jezikovni kompetenci oz. produkciji, ki je post festum prepoznana in zabeležena z leksikografskim dejanjem.

¹² To možnost predvideva tudi sistem WS, rezultate, ki jih navaja, pa smo deloma preuredili glede na naknadno preverjanje konkordanc in z upoštevanjem sinonimnih variant tipa: *orlovski – kljunast* (*nos*) ipd.

Prid_nos

dolg velik širok ploščat kratek okrogel	zamašen/ zamašenost smrkav	zlomljen polomljen razbit krvav krvaveč	kljukast ukrivljen koničast prifrknjen privihan kriv povešen	rdeč	orlovski ovnov kljunast	dober občutljiv	umeten
--	----------------------------------	---	--	------	-------------------------------	--------------------	--------

Glag_nos

povesiti/ pobesiti vihati/ zavihati/ vihanje dvigniti	zlomiti razbiti poškodovati polomiti	zamašiti zatisniti/ zatiskati tiščati	obrisati/brisati	napudrati
--	---	--	------------------	-----------

Glag_Predl¹³_nos

obrisati/ brisati pod	pomoliti/ moleti tiščati pod podržati pod potisniti/ potiskati riniti pod	vreči pod metati pod poriniti pod prinesti pod	pokaditi/kaditi pod	pomahati/ mahati pod	
zaloputniti pred zapreti/ zapiirati zakleniti pred	pomoliti/moliti podržati pred postaviti pred poriniti pred	odnesti pred izmakniti pobrati ukrasti	speljati/odpeljati pred uiti pred izmuzniti pred zbežati pred pobegniti pred	mahati/ pomahati migati pred bingljati pred	imeti pred
		odnesti izpred izmakniti ukrasti	speljati/odpeljati izginiti izpred		
vleči/ povleči za potegniti za voditi za	prijeti za držati za				
vtikati/ vtakniti v	iti v	udariti v			

¹³ Variantnost predloga ob pomensko sorodnih glagolih je v preglednici označena z obarvanimi polji

vrtati/ vrtanje po bezati po drezati po brskati po šariti po	dobiti po	udariti po kresniti po mahniti po treščiti po počiti po lopniti po obrisati po	popraskati po požgečkati po podrgniti po poščegetati po		
dihati skozi pihniti skozi vdihniti/ vdihavati/ vdihovati skozi					

Sam_Predl_nos

kapljice za pršilo za razpršilo za	ščitnik za	mazilo za	
puh pod brk pod dlaka pod	mikrofon pod	pest pod prst pod	koža pod
izcedek iz izloček iz	krvavitev iz kri iz		

Rezultati predstavljajo vsaj dva različna tipa besednih zvez: (1) (proste) besedne zveze, med katerimi WS izločijo le najtipičnejše, zato jih upravičeno lahko imenujemo kolokacije, npr. *dolg, velik, širok, ploščat; zamašen; zlomljen; kljukast* ipd. *nos; zlomiti, razbiti, poškodovati; obrisati, napudrati* ipd. *nos*, ter (2) ustaljene in hkrati leksikalizirane zveze, za katere je – z analizo njihovega besedilnega okolja ali s pregledom konkordanc – mogoče ugotoviti samostojni pomen, npr. *povesiti/pobesiti nos* ‘kazati razočaranje, užaljenost’; *vihati/zavihati/dvigniti nos* ‘izražati nestrinjanje; vzvišenost’; *potegniti za nos* ‘prevarati, ukaniti’. Natančnejši pregled konkordanc v zvezi s posamezno zvezo odkriva tudi predvidljivost širšega besedilnega okolja in slovnično obnašanje besed na predvidenih mestih skladijskega vzorca, npr. *(kdo) povesi nos* in *nos se povesi (komu)*; *(kdo) viha nos (nad kom/čim)*. Posebno očitno ustvarjajo samostojne pomenne in s tem leksikalni status predložne glagolske zveze, zlasti ko so vezave s konkretnim predlogom razmeroma osamljene, npr. *(imeti, povesiti, spustiti) nos do tal* ‘biti razočaran, užaljen’, ali pa med glagolom in predlogom stoji še ena sestavina, npr. *(oditi, ostati) z dolgim nosom/dolgega nosu*. V nekaterih primerih obstajajo dobesedni in idiomatični pari, ki vključujejo iste ali pomensko sorodne glagole, npr. neleksikalizirano: *(o)brisati (si) nos* : leksikalizirano: *obrisati se pod nosom* ‘ne dobiti česa pričakovanega’. Se pa v takih primerih potrjuje pravilo, da leksikalizirana zveza glede na neleksikalizirano vedno izkazuje

določeno spremembo ali omejitev v skladenjskem vzorcu (Gantar 2007: 246–247). Poleg obveznega predloga (*pod*) se leksikalizirana različica za pomen ‘ne dobiti česa pričakovanega’ očitno ne rabi prehodno, tj. ni v navadi reči: *obrisati koga pod nosom (za kaj)* – npr. v smislu ‘prikrajšati ga’, hkrati pa v tem pomenu v predložni zvezi nastopa le dovršni povratnosvojljni glagol: *obrisati se*. Posebej zanimive so s samostalnikom *nos* ustaljene predložne zveze, ki delujejo kot vmesni člen med pomensko samostojnimi LE in pogostimi prostimi zvezami. Zanje je značilno, da v kombinaciji z glagolom določene pomenske skupine opisujejo konkretno gesto, npr. (*zaloputniti, zapreti*) *pred nosom*; (*pomoliti, moleti, tiščati, poriniti*) ipd. *pod nos*; (*odnesti, izmakniti, ukrasti*) *pred nosom/izpred nosa*; (*speljati, odpeljati, izginiti, uiti*) *pred nosom/izpred nosa/nosu*; (*pamahati, mahati, migati*) *pred/pod nosom*. Polja, ki izkazujejo variantnost predlogov znotraj pomensko sorodnih glagolov, smo, kot rečeno, sivo obarvali.

4.2 Formalni kazalci leksikaliziranosti

Čeprav se v primeru ugotavljanja leksikaliziranosti zvez v prvi vrsti zanašamo na jezikovni občutek, obstajajo določeni objektivni kazalci, ki napovedujejo leksikalizacijo. Predvidevali smo namreč, da ima beseda, ki ima v določenem pomenu obsežen povezovalni potencial, manj možnosti za leksikalizacijo, kot beseda, katere povezovalni potencial je omejen oz. pripadajo kolokatorji znotraj posameznega pomena sorodnemu pomenskemu polju, kar pomeni, da so v določenem pomenskem razmerju, npr. *ukrasti/izmakniti; mahati/migati/bingljati*. Tega merila se sicer ni mogoče držati za vsako ceno, saj skromen kolokacijski niz (tj. majhno število kolokatorjev, ki se sopojevajo ob besedi za konkretni pomen) sam na sebi še ni dovolj za proglašitev VLE. Zvez kot *obrisati (si) nos, napudrati si nos, imeti pred nosom* ni mogoče določiti za leksikalizirane samo zato, ker se tipično sopojevajo s skrajno omejenim številom kolokatorjev. Do neke mere je njihovo pomensko samostojnost mogoče upravičiti z enkratnostjo/konkretnostjo predmetnosti, na katero se nanašajo, npr. *umetni nos, obrisati (si) nos* (‘usekniti se’) ali z obstoječim konotativnim pomenom, kot v primeru *rdeč nos* (= ‘znamenje/simbol alkoholiziranosti’) in *napudrati si nos* v zgledih:

... ljudje z naluknjanimi jetri in **rdečimi nosovi** na zdravniških pregledih trdijo, da ne pijejo alkohola, ampak samo domače vino (kar po njihovem mnenju ni alkohol, ker je domače)

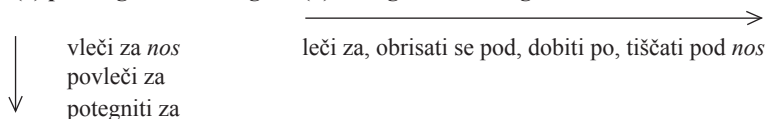
Ko si je Maja šla končno ‘**napudrat nos**’, sem ga vprašal: ‘Ti, veliki osvajalec, zakaj si toliko časa čistil očala in se tako popolnoma odkril pred njo?’

V skladu s predpostavko pa je v zgornjih preglednicah dejansko mogoče evidentirati zveze, kjer skrajno omejen povezovalni niz napoveduje leksikalni status, npr. *dober/občutljiv nos; pokaditi pod nos; iti/gre v nos; vleči/voditi za*

nos; *dobiti po nosu*; *vtikati nos* itd. V nekaterih primerih se število kolokatorjev povečuje na paradigatski ravni, vendar ostajajo znotraj istega pomenskega polja, kar kaže prej na pestro variantnost znotraj istega pomena, kot kaže shema (a) spodaj, in ne na poimenovalni sintagmatski proces, ki odloča o več samostojnih LE ali zgolj o konkretni zapolnitvi v jeziku predvidenega skladenjskega vzorca, kot prikazujemo v shemi (b).

(a) paradigatski obseg

(b) sintagmatski obseg



Nasprotno so omejitve kolokatorjev na sintagmatski osi, ki jih spremlja tudi oblikovna ustaljenost, npr. *na vrat na nos*, pokazatelj skrajno omejene povezovalnosti, ki je navadno odločujoča za prepoznavnost FE.

Pri odločitvi o leksikalnem statusu zveze ima zadnjo besedo njen besedilni kontekst. Za zveze z omejeno povezovalno paradigmo smo ugotovili tipične skladenjske vzorce, znotraj katerih lahko uresničujejo svoj nedobesedni leksikalni pomen: *imeti dober/občutljiv nos za (kaj)*; *(kaj) gre v nos (komu)*; *vleči/voditi (koga) za nos*; *pokaditi (komu) pod nos*; *vtikati nos v (kaj)*, pri čemer denimo za zadnji primer velja, da pomen leksikalizirane zveze povsem brez težav nosi glagol sam: prim. *vtikati se*.

Hkratno upoštevanje skladenjskih in pomenskih povezovalnih procesov, na podlagi katerega je mogoče izdelati pregledno sliko kolokatorjev, ki zasedajo posamezna mesta znotraj skladenjsko-pomenskega vzorca preučevane besede/zveze oz. njenega pomena, pride še posebno do izraza pri glagolih. Za ponazoritev smo izbrali glagol *namazati*, zanima pa nas besedilno okolje v okviru ene besede pred in za preiskovano besedo. Kolokatorje smo, kot že pri samostalniku, razvrstili glede na besedno vrsto, znotraj tega pa glede na prepoznavne pomenske skupine (označuje jih rimska številka). Kot že v prejšnjem primeru, smo upoštevali besedno skico omenjenega glagola.

Prisl	Glag	Sam I	Sam II	S a m III	P	Sam IV	Sam V
debelo	namazati	kruh	koža	pekač	z	maslo	krema
tanko		testo	ustnica	posoda		nadev	mazilo
enakomerno		palačinka	obraz			olja	vazelin
		rezina	trepalnica			marmelada	losjon
		biskvit	roka			gorčica	žavba
			noga			mast	
						margarina	
						beljak	

Pri oblikovanju skladijsko-pomenskega okvira je bilo treba rešiti problem samostalniških kolokatorjev, ki nastopajo v skladijskem vzorcu glagola s t. i. dvopredmetno zgradbo, saj se ti razvrščajo glede na vezljivostni vzorec in oblikoskladijske lastnosti določil različno, medtem ko udeleženske vloge ostajajo enake, kot kažeta zgleda iz *FidePLUS* za primera (c) in (d) spodaj.

(c) namazati kruh z maslom

... če si zjutraj na debelo **namazemo kruh z maslom** ali margarino, zaužijemo skoraj dvakrat več kalorij, kot če bi pojedli dve rezini polnozrnatoga kruha.

(d) namazati maslo na kruh

... če ne bi imeli žlic, vilic in noža, s čim bi **namazali maslo na kruh**, pomešali kavo ali čaj?

Kompleksnejšo informacijo zagotavlja sistem besednih skic, saj posebej izloča tipične predložne zveze, npr. *namazati z (maslo, krema, marmelada, olje, vazelin, skuta, mast, lepilo ...)*, *namazati na (kruh, pekač ipd.)*, nedoslednosti pa ostajajo pri razvrstitvi določil glede na skladijske funkcije, zlasti v primeru predmetnih tožilniških določil, ki so prekrivna z imenovalnikom, saj jih program avtomatično pripiše osebkovemu mestu: *(pekač, kruh, mesto) namazati*, čeprav gre za predmetna določila. Pomenske skupine kolokatorjev, kot je razvidno iz preglednic **Sam I** : **SamIV** in **SamV**, so nam tu v pomoč, saj je (po logiki *FrameNeta*) mogoče ločiti samostalniška določila z različnimi udeleženskimi vlogami, npr. *kruh, testo, palačinke, rezina, biskvit, pekač, koža, ustnica, obraz* itd. in *krema, mazilo, vazelin; maslo, nadev, olje* itd.

5 Sklep

Pojmovanje leksike kot celotnega inventarja besed v jeziku in hkrati kot kompleksnega skladijsko-pomenskega pojava je usmerilo pozornost leksikalnih analiz na vzorce pomenske kohezije, ki se ustvarja v razponu med posamezno besedo in celostno stavčno zgradbo, katere del je. Takšno izhodišče, ki je lahko v vsej svoji razsežnosti prišlo do izraza s korpusnim jezikoslovjem, kjer je mogoče s statističnimi izračuni izluščiti to, kar izstopa glede na regularnost vzorca in pomena, širi področje večbesedne leksike, ki je bilo še do nedavnega vezano predvsem na frazeološke enote, na področje besedne povezovalnosti nasploh. Izhodišče preučevanja se s posamezne jezikovne ravnine premakne v besedilo, kjer se jezikovni elementi povezujejo na podlagi slovnične kompatibilnosti in pomenske soodvisnosti. Prvi korak k odkrivanju leksikalnega potenciala konkretne besedne kombinacije vidimo v določitvi obsega povezovalnega niza. Pri tem si je mogoče pomagati z oblikovanjem besednih skic za slovenščino. Rezultati, ki jih dobimo na ta način, ponujajo kompleksno informacijo o najpogostejših sopojavitvah, stavčnem okolju in tipičnih zapolnitvah skladijskih mest, kar predstavlja neke vrste skladijsko-pomenski profil preučevane besede. Ugotovili smo, da ima beseda, ki ima v določenem pomenu obsežen povezovalni potencial, manj

možnosti za leksikalizacijo, kot zveza, katere povezovalni potencial je omejen. V nekaterih primerih se število kolokatorjev povečuje na paradigmatski osi, vendar ostajajo znotraj istega pomenskega polja, kar kaže prej na bogato variantnost znotraj istega pomena besede ali zveze kot pa na sintagmatski proces, ki odloča o več samostojnih LE. Pri odločitvi o leksikalnem statusu zveze ima vedno zadnjo besedo njen besedilni kontekst, ki določa obseg in elemente besedilnega okolja, znotraj katerega lahko preučevana beseda uresničuje svoj pomen.

Literatura

Bloomfield, Leonard, 1933: *Language*. New York: Holt & Co.

Bollinger, Dwight. L., 1976: Meaning and memory. *Forum Linguisticum* 1 (1). 1–14.

Chomsky, Noam, 1965: *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam, in Halle, Morris, 1968: *The Sound Pattern of English*. New York: Harper and Row.

Cowie, Antony (ur.), 1998: *Phraseology: Theory, Analysis, and Applications. Introduction*. Oxford: Oxford University Press. 1–20.

Čermák, František, 1985: Frazeologie a idiomatika. Čermák, František in Filipec, Josef: *Česká lexikologie*. Praha: Academia. 166–248.

De Mauro, Tullio, 2006: On Lexicon and Grammar. Corino, E., Marellò, C., in Onesti, C. (ur.): *Proceedings XII. EURALEX International Congress*. Congresso internazionale di Lessicografia, Atti Volume I. Torino, Italia, September 6th– 9th: Università degli Studi di Torino.

FidaPLUS: Korpus slovenskega jezika FidaPLUS: <<http://www.fidaplus.net/>>.

Gabrovšek, Dušan, 1991: Leksikološki in leksikografski vidiki nekaterih besednih zvez: »dajanje predloga« in sorodne zgradbe v slovenščini in angleščini. *Vestnik* 25/1–2. 105–135.

Gabrovšek, Dušan, 2000: Phraseology galore: Words and their combinations. *Vestnik* 34/1–2. 183–238.

Gantar, Polona, 2006: Korpusni pristop v frazeologiji in slovanske aplikacije. Vidovič Muha, Ada (ur.): *Slovensko jezikoslovje danes. Slavistična revija* 54 (posebna številka). 151–163, 519–532.

Gantar, Polona, 2007: *Stalne besedne zveze v slovenščini: korpusni pristop*. Ljubljana: Založba ZRC, ZRC SAZU (Zbirka Lingua Slovenica 3).

Gorjanc, Vojko, 2005: *Uvod v korpusno jezikoslovje*. Domžale: Izolit.

Hanks, Patrick, 1994: Linguistic Norms and Pragmatic Exploitations, or Why Lexicographers need Prototype Theory and Vice Versa. Kiefer, F., Kiss, G., in Pajzs, J. (ur.): *Papers in Computational Lexicography: Complex '94*. Research Institute for Linguistics, Hungarian Academy of Sciences.

Hanks, Patrick, in Pustejovsky, James, 2004: Common Sense About Word Meaning: Sense in Context. *TSD 2004*. Brno.

Hanks, Patrick, in Pustejovsky, James, 2005: A Pattern Dictionary for Natural Language Processing. *Revue Francaise de linguistique appliquée* 10:2.

Johnson, Samuel, 1755: *A Dictionary of the English Language*. London: Knapton and Others.

Jurko, Primož, 1997: *Leksikalne kolokacije s kontrastivnega stališča: poskus razvrstitve po težavnosti*. Magistrsko delo. Mentor: Dušan Gabrovšek. Ljubljana: Filozofska fakulteta.

Krek, Simon, 2008: *FrameNet* in slovenščina. *Jezik in slovstvo* 53/5. 37–54.

Krek, Simon, in Kilgarriff, Adam, 2006: Slovene Word Sketches. *Jezikovne tehnologije: zbornik 9. mednarodne multikonference Informacijska družba IS 2006, 9. do 10. oktober 2006. Language technologies: proceedings of the 9th International Multiconference Information Society IS 2006, 9th-10th October 2006*.

Kržišnik-Kolšek, Erika, 1988: *Frazeologija v moderni*. Magistrsko delo. Mentorica: Breda Pogorelec. Ljubljana: Filozofska fakulteta.

Kržišnik, Erika, 1987: Prenovitve stalnih besednih zvez v Kosmačevi prozi 30. let. Zadavec, Franc, Koruza, Jože, Skaza, Aleksander, in Toporišič, Jože (ur.): *Socialni realizem v slovenskem jeziku, književnosti in kulturi: mednarodni simpozij v Ljubljani od 26. do 28. junija 1985* (Obdobja 7). Ljubljana: Filozofska fakulteta. 517–532.

Logar, Nataša, 2005: Besedotvorni sklopi. *Slavistična revija* 53/2. 171–192.

Moon, Rosamund, 1998: *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon Press (Oxford Studies in Lexicography and Lexicology).

OCD: *Oxford Collocation dictionary for students of English*. Crowther, J., Dignen, S., in Phillips, J., (ur.), 2003 (peta izdaja). Oxford: Oxford University Press.

OOI, B. Y. Vincent, 1998: *Computer Corpus Lexicography*. Edinburgh: Edinburgh University Press.

Sinclair, John M., 1991: *Corpus, Concordance and Collocation*. Oxford: Oxford University Press.

SSKJ: *Slovar slovenskega knjižnega jezika na CD-romu z Odzadnjim slovarjem slovenskega jezika in Besediščem slovenskega jezika z oblikoslovnimi podatki*. Ljubljana: SAZU, ZRC SAZU, Inštitut za slovenski jezik, DZS.

Teubert, Wolfgang, 1999: Korpuslinguistik und Lexikographie. *Deutsche Sprache* 4/99. Amsterdam, Philadelphia: John Benjamins. 293–313.

Toporišič, Jože, 1982: *Nova slovenska skladnja*. Ljubljana: DZS.

Toporišič, Jože, 1992: *Enciklopedija slovenskega jezika*. Ljubljana: Cankarjeva založba.

Toporišič, Jože, 2000: *Slovenska slovnica*. Četrta, prenovljena in razširjena izdaja. Maribor: Obzorja.

Vidovič Muha, Ada, 1972: Kategorizacija in stilna opredelitev ozko knjižne leksike. Pogorelec, Breda (ur.): *7. seminar slovenskega jezika, literature in kulture*. Ljubljana: Filozofska fakulteta. 35–53.

Vidovič Muha, Ada, 1988: Nekateri jezikovnosistemske lastnosti strokovnih besednih zvez. Pogorelec, Breda, Sajovic, Tomaž, in Počaj-Rus, Darinka (ur.): *24. seminar slovenskega jezika, literature in kulture*. Ljubljana. 83–91.

Vidovič Muha, Ada, 2000: *Slovensko leksikalno pomenoslovje – Govorica slovarja*. Ljubljana: Znanstveni inštitut Filozofske fakultete.

Vrbinc, Marjeta, 1995: *Slovarska obravnava kolokacij v splošnih dvojezičnih slovarjih*. Magistrska naloga. Mentor: Dušan Gabrovšek. Ljubljana: Filozofska fakulteta.

Vrbinc, Marjeta, 2001: Kolokacije in rojeni govorcev – pomen vključevanja kolokacij v dvojezične slovarje. Kržišnik, Erika (ur.): *Skripta 5*. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovanske jezike in književnosti Filozofske fakultete. 51–63.

Spletne strani

<<http://framenet.icsi.berkeley.edu/>>. (Dostopno: 30. 10. 2008.)

<<http://nlp.fi.muni.cz/projekty/cpa/>>. (Dostopno: 30. 10. 2008.)

<<http://www.sketchengine.co.uk/>>. (Dostopno: 30. 10. 2008.)

