



UDK 811.163.6'367.624

Polona Gantar

Inštitut za slovenski jezik Frana Ramovša ZRC SAZU

PROSTORSKA INFORMACIJA V STAVČNO ZASNOVANI SLOVARSKI DEFINICIJI

V prispevku je prikazana zasnova pomenskih shem – skladenjskih realizacij posameznih glagolskih pomenov – v Leksikalni bazi za slovenščino, ki predstavljajo izhodišče za oblikovanje stavčnih razlag v slovarju in omogočajo uporabo v jezikovnotehnološke namene. Glagolski pomeni, ki vsebujejo prostorsko komponento v pomenski shemi, so analizirani z vidika stavčnih vzorcev, v katerih se tipično realizirajo, in z vidika leksikalnih zapolnitev posameznih vezljivostnih mest, kot se kažejo v kolokacijah.

Ključne besede: leksikalna baza za slovenščino, pomenske sheme, stavčne razlage, semantični tipi, stavčni vzorci, kolokacije, prostorska pomenska komponenta

The paper describes the design of semantic frames, i.e., syntactic realizations of individual verbal meanings in the Slovene Lexical Database, which represent the basis for formulation of sentence-level lexical definitions and allow the use with language technologies. Verbal meanings containing a spatial component in their semantic frames are analyzed from the point of view of the syntactic patterns in which they are typically manifested and from the point of view of lexical realization of individual arguments as evident from collocations.

Keywords: Slovene lexical database, semantic frames, sentence definitions, syntactic patterns, collocations

1 Uvod

V prispevku je opisana zasnova *pomenskih shem*, kot smo jih oblikovali v Leksikalni bazi za slovenščino v okviru projekta Sporazumevanje v slovenskem jeziku, in sicer za posredovanje leksikalnogramatičnih podatkov v slovarju v obliki *stavčnih razlag* in za strojno procesiranje slovenščine. Zasnova tovrstnih opisov izhaja iz dveh teoretičnih modelov: evociranja pomenov besed na podlagi t. i. *shemske semantike* v projektu Framenet (Fillmore idr. 2003) in na podlagi opisov *stavčnih vzorcev* po teoriji jezikovnih konvencij in možnosti njihove izrabe v projektu Corpus Pattern Analysis (Hanks 1994, 2004; Hanks in Pustojevsky 2005). Obe izhodišči smo pri oblikovanju pomenskih shem nadgradili z upoštevanjem posebnosti slovenščine in glede na namebnost slovarske baze. Pri analizi prostorske informacije, kot je v stavčno oblikovanih pomenskih shemah razvidna iz *semantičnih tipov*, ki jih razumemo kot abstraktne zastopnike tipičnih leksikalnih zapolnitev na posameznih vezljivostnih mestih določenega glagolskega pomena, npr. *če PREDMET zaradi sile težnosti pade na nek PROSTOR ali POVRŠINO, prileti tja iz zraka ali z višine*, smo se osredotočili na tiste glagole oz. njihove pomene, pri katerih je mogoče na podlagi korpusne analize in z analizo besednih skic v orodju Sketch Engine (Kilgarrieff idr. 2004) ugotoviti prostorsko determiniranost pomena.

Obenem smo analizirali tudi načine skladenjskega in pomenskega izražanja prostorske komponente. V ta namen smo analizirali stavčne vzorce, značilne za določen pomen besede, npr. *pasti* – 'premikati se od zgoraj proti navzdol': *kaj pade kam, kaj pade na kaj, kaj pade v kaj* ipd. Ob tem smo beležili tipične leksikalne zapolnitve na posameznih mestih evidentiranih stavčnih vzorcev v obliki kolokacij, npr. [bomba, granata, drevo] pade, pasti na [dno, zemljo, tla], pasti v [prepad, globino, morje] ipd.

2 Leksikalna baza za slovenščino v okviru projekta SSJ

V okviru projekta Sporazumevanje v slovenskem jeziku (SSJ)¹ je bila med drugim izdelana Leksikalna baza za slovenščino (LBS), ki naj bi, kot izhaja iz projektne dokumentacije, služila (a) določitvi standardov za izdelavo slovarskih gesel v sodobnih slovarskih priročnikih za slovenščino (Gantar idr. 2009; Gantar idr. 2009a) in (b) računalniški obdelavi jezika oz. uporabi v jezikovnotehnoloških aplikacijah za slovenščino.

Leksikalnogramatični podatki so v LBS organizirani v obliki med seboj povezanih nivojev: pomenskem, skladenjskem, kolokacijskem in ponazarjalnem, ki jih povezuje lema oz. iztočnica. Najpomembnejši del – v smislu, da so mu podrejeni vsi drugi nivoji – je pomenski nivo, ki prinaša informacije o pomenski razčlenjenosti besede v obliki *pomenskih indikatorjev*, ki so primarno namenjeni oblikovanju pomenskega menija in služijo lažji navigaciji po zgradbi večpomenskega gesla v spletni postavitvi. V pričujočem prispevku nas bo zanimal predvsem pomenski opis, ki je zapisan v obliki pomenskih shem, ki se konceptualno približujejo semantičnim okvirjem v projektu Framenet in stavčnim vzorcem v projektu CPA, ki jih nekoliko podrobneje opredelimo v nadaljevanju. Pomenske sheme predstavljajo osnovo za oblikovanje stavčnih definicij Cobuildovega tipa (Hanks 1987, Barnbrook 2002, Dzemianko in Lew 2013), kar je za slovenščino novost, hkrati pa so z navedbo udeleženske zgradbe in s prikazom zapolnitve posameznih mest s semantičnimi tipi osnova za strojno procesiranje skladenjsko-pomenskih povezav na ravni stavka.

3 Teoretična in metodološka izhodišča

Pri oblikovanju pomenskih shem smo se teoretično naslonili na projekta Framenet, ki ga je zasnoval Charles Fillmore, in Corpus patterns Analysis (CPA) Patricka Hanksa, ki zaradi združitve skladenjskih in pomenskih informacij najbolj ustrežata na korpusu temelječemu opisu sodobne slovenske leksike.

3.1 Projekt Framenet

Logika pomenskega opisa, kot jo predstavlja Framenet (Fillmore in Atkins 1998; Fillmore idr. 2003),² temelji na ugotavljanju potrebnih (jedrnih in nejedrnih) shemskih

¹ Projekt je financiralo Ministrstvo za izobraževanje, znanost in šolstvo ter Evropski socialni skladi. Rezultati projekta so dostopni na: <http://www.slovenscina.eu/>.

² Teorija in praksa projekta Framenet je podrobno predstavljena v posebni številki *International Journal of Lexicography*, Zv. 16, Št. 3, september 2003.

elementov za evociranje posameznih pomenov besede, ki jih je nato na podlagi skupnih oz. razlikovalnih elementov mogoče združevati v semantične okvirje oz. sheme. Tako npr. shema »Staranje« (ang. *Aging*) med drugimi vključuje glagola *postarati se*, *starati se* ter pridevnik *ostarel* (ang. *age*, *aging* in *mature*). Obvezni ali jedrni element te skupine je »entiteta«, ki je podvržena staranju, nejedrni, torej opcijski elementi pa so lahko še: »okoliščine«, »način«, »lokacija«, »rezultat« in »čas«. Če skušamo pravilo zelo poenostavljeno prenesti na slovenski zgled, bi npr. v povedi: *Pretirano je kadila, njen obraz se je hipoma postaral*, zvezo *njen obraz* označili za »entiteto«, *hipoma* za »način«, *pretirano je kadila* pa za »okoliščine«, v katerih je entiteta podvržena procesu staranja. Dokumentacija shemskih elementov po vzoru Frameneta predvideva tudi opis njihove skladenjske realizacije.³ Kot pravita ATKINS in RUNDELL (2008: 145), je framenetovski pristop najbolj primeren za raziskovanje obnašanja besed na podlagi korpusnih podatkov, ker omogoča dosledno analizo besedilnega okolja, hkrati pa ne spregleda nobenega ključnega pomenskega dejstva.

3.2 Projekt CPA

Drugi projekt, ki izhaja z korpusne analize stavčnih vzorcev v povezavi s konkretnimi pomeni obravnavane besede, je projekt analize korpusnih vzorcev (Corpus Patterns Analysis; Hanks 2004, 2008, 2013; Hanks in Pustejovsky 2005). Projekt 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 Pustojevsky 2005). V postopku analize so posamezne besede (poleg glagolov so na ta način obravnavani tudi samostalniki) na podlagi primerov realne rabe povezane v semantično motivirane sintagmatske vzorce. Tem vzorcem je nato pripisan pomen, ki je izražen s t. i. implikaturami, na njihovi podlagi pa je mogoče postopoma oblikovati pomensko ontologijo. Identifikacija semantičnega vzorca nikakor ni preprost in avtomatiziran postopek, pač pa zahteva veliko leksikografske spretnosti in prakse: med najtežjimi leksikografskimi odločitvami je prav gotovo določitev ustrezne stopnje posplošitve, znotraj katere je mogoče razločevati posamezne (pod)pomene.

V slovarju *Pattern Dictionary of English Verbs*,⁴ ki izhaja iz omenjenega projekta, je za vsak glagol določeno število osnovnih vzorcev, ki ustrezajo registriranim pomenom, ter vzorci, ki predstavljajo ustvarjalno rabo ali t. i. eksploatacije. V Hanksovi analizi korpusnih vzorcev tako npr. glagol *adore* (oboževati; častiti) predvideva dva osnovna vzorca, ki izkazujeta tudi možnost eksploatacije:

- 66 % [[Human 1]] adore [[Human 2 | Deity | {Animal = Pet}]] conc.⁵
[[Human 1]] has extremely strong positive emotions with regard to [[Human 2 | Deity | {Animal = Pet}]] exploit.
- 34 % [[Human]] adore [[Entity | Eventuality]] conc.
[[Human]] enjoys [[Entity | Eventuality]] very much exploit.

³ Več o tem v prispevku S. Kreka (2008).

⁴ Celotni opis korpusnih vzorcev, vključno s povezavo na korpusne konkordance, ki ustrezajo posameznemu vzorcu, je za trenutno 877 angleških glagolov dostopen na: <http://deb.fi.muni.cz/pdev/>.

⁵ Tako pri osnovnih vzorcih kot pri eksploatacijah so predvidene povezave na konkordance, ki ustrezajo vzorcu.

Tak način označevanja pomenskih lastnosti glagolov služi predvsem uporabi v jezikovnotehnološke namene ter za jezikoslovne analize, ki temeljijo na empiričnih podatkih. Obema projektoma, kot tudi oblikovanju pomenskih shem v LBS, je skupno to, da izhajajo iz kontekstualnega principa določanja besednega pomena oz. iz spoznanja, da pomeni besed sami na sebi ne obstajajo (Kilgarriff 1997), obstajajo samo pomenske tendence, ki jih besede uresničujejo v sobesedilu (Hanks 2009).

4 Pomenske sheme v Leksikalni bazi za slovenščino

Pomenska shema je v zgradbi LBS tisti element geselskega članka, ki je namenjen skladijskemu opisu konkretnega pomena obravnavane besede. Kot taka vsebuje ključne informacije o realizaciji pomena v okviru skladijsko zaključene celote, kot je prosti stavek,⁶ in predstavlja vir podatkov za oblikovanje stavčnih definicij, ki vključujejo vse ključne skladijsko-pomenske elemente posameznega pomena, tj. udeležence in okoliščine, njihovo oblikoskladijsko organizacijo kot tudi konotativne in pragmatične jezikovne prvine, potrebne za njegovo realizacijo.

Stavčni vzorec, ki ga vsebuje pomenska shema, kot jo razumemo v LBS, praviloma vključuje vse pomensko relevantne udeležence v obliki t. i. semantičnih tipov (gl. v nadaljevanju), oz. prikazuje »največjo« udeležensko zgradbo, ki jo je mogoče ugotoviti iz večjega števila realnih besedilnih zgledov⁷ za konkretni glagolski pomen. Pri tem je pomembno opozoriti, da z vidika stavčnega vzorca, znotraj katerega prepoznavamo pomensko intencno polje glagola, obravnavamo kot pomensko relevantna tako (številčno omejena) vezljivostna določila kot skladijsko nevezjavne udeležence, ki so (običajno številčno neomejena) nevezljivostna dopolnila (prim. ŽELE 2003: 35). V pomenskih shemah in stavčnih vzorcih se o vrsti skladijskega razmerja (ujemanje, vezava, primik) ne odločamo, zanima nas predvsem pomenska (ne)nujnost oz. realna prisotnost/odsotnost udeležencev oz. konkretne leksikalne zapolnitve na njihovih mestih.

Udeležence in okoliščine, ki predstavljajo udeležensko zgradbo posameznega pomena glagola, zapisujemo v pomenski shemi z velikimi črkami v obliki t. i. semantičnih tipov zaradi možnosti avtomatskega pridobivanja prototipičnih vzorcev in zaradi povezovanja elementov pomenske sheme s kolokacijami in stavčnimi vzorci. Hkrati z določitvijo števila udeležencev določimo v pomenski shemi tudi njihove sklonske oblike in razporeditev v stavčni zgradbi, ki naj bi odražala tudi tipično skladijsko realizacijo pomena besede v stavku, ter t. i. pomenski scenarij, ki povezuje udeležence in okoliščine v konkretno sporočanje situacijo, ki je za dani pomen značilna:

⁶ Vidik pomenske povezovalnosti, ki izhaja iz koncepta besedne kompatibilnosti in kolokabilnosti (ČERMÁK 2010: 282), predvideva izraženost pomenske kolokabilnosti tudi zunaj stavčne povedi obravnavanega glagola, kar koncept vezljivosti, ki navadno pomensko kompatibilnost in kolokabilnost oblikoskladijsko kategorizira le znotraj stavka, širi na raven besedila. V zvezi s tem prim. ŽELE (2010, 2012: 251 in dalje).

⁷ Ročna analiza je zajela najmanj 150 in navadno ne več kot 300 konkordanc v korpusu Gigafida (Logar idr. 2012).



4.1 Semantični tipi

O udeleženski zgradbi, kot jo beležimo v pomenski shemi, ni mogoče govoriti, ne da bi pri tem opozorili na vlogo semantičnih tipov kot abstraktnih zastopnikov leksikalnih zapolnitev na predvidenih udeleženskih mestih. Leksikalno realizacijo semantičnih tipov predstavljajo na kolokacijskem nivoju kolokacije kot konkretne statistično izstopajoče leksikalne zapolnitve udeleženskih mest. Kolokacije za predvidene skladienske pozicije je na podlagi slovnice besednih skic,⁸ ki delujejo na oblikoslovno označenem in lematiziranem korpusu v orodju Sketch Engine (Kilgarriff idr. 2004), pridobiti s pomočjo besednih skic (ang. *Word Sketches*). Slika 1 prikazuje del besedne skice za glagol *pasti* v kombinaciji s predlogom *na*:

S na-d-ef	20963	24.4
<input type="checkbox"/> z pamet	2671	10.21
<input type="checkbox"/> z tla	4710	9.35
<input type="checkbox"/> z pleča	236	8.0
<input type="checkbox"/> z finta	162	7.76
<input type="checkbox"/> z rame	306	7.58
<input type="checkbox"/> z vodšče	119	6.58
<input type="checkbox"/> z glava	1014	6.45
<input type="checkbox"/> z zemlja	540	6.4
<input type="checkbox"/> z pokrov	129	6.29
<input type="checkbox"/> z hrbet	218	6.28
<input type="checkbox"/> z misel	508	6.22
<input type="checkbox"/> z rit	34	6.15
<input type="checkbox"/> z bok	122	5.95
<input type="checkbox"/> z dno	207	5.89

Slika 1: Del besedne skice za glagol *pasti* v korpusu Gigafida.

⁸ Prva različica slovenske slovnice besednih skic, ki je bila izdelana ob vključitvi korpusa FidaPLUS v orodje sKE (Krek in Kilgarriff 2006), je bila prvič nadgrajena po končani prvi fazi izdelave LBS, ko je bilo mogoče na podlagi pomensko-skladijskih analiz pri poskusno izdelanih geslih ugotoviti, katere skladienske strukture so za leksikalnogramatični opis slovenščine relevantne in jih je potrebno vključiti v novo različico. Druga nadgradnja je potekala v povezavi z izvedbo postopka avtomatizacije, ki je poleg prilagojenih in dodanih gramatičnih relacij upoštevala tudi možnosti, ki so bile na novo razvite v samem orodju, hkrati pa je bila ta različica aplicirana tudi na korpus Gigafida.

V prvi fazi izdelave LBS smo semantične tipe zapisovali brez enotne taksonomije, saj gre za postopek, ki ga je potrebno reševati induktivno, tj. od posameznega (konkretnega kolokatorja) k skupnemu oz. abstraktnemu pomenskemu zastopniku. Poleg tega je bilo mogoče pričakovati, da bodo na začetku abstraktni pomenski zastopniki zapisani z različnimi izrazi, saj je posamezne skupne pomenske imenovalce mogoče izraziti na več načinov. V drugi fazi izdelave LBS smo zato posamezne semantične tipe poenotili in izdelali štiristopenjsko taksonomijo. Kot hierarhično najvišje smo določili t. i. (1) pomensko polje, ki vsebuje posamezne pomenske razrede, in sicer (1.1) pomenske tipe, (1.2) pomenske podtipe, (1.3) kolokatorske tipe in (1.4) konkretne kolokatorje. **POMENSKA POLJA** zastopajo vrhnje pomenske kategorije, ki služijo združevanju osnovnih pomenskih tipov. Ob zaključku LBS so bila taka pomenska polja tri: (1) živo bitje, (2) konkretna entiteta in (3) abstraktna entiteta. **POMENSKI TIPI** predstavljajo samostojne pomenske kategorije znotraj pomenskih polj in smo jih v pomenski shemi uporabljali za označevanje udeležencev, kadar besedilnih realizacij ni bilo mogoče natančneje opredeliti. **POMENSKI PODTIPI** določajo relativno samostojne pomenske skupine znotraj pomenskega tipa, **KOLOKATORSKI TIPI** pa natančneje opredeljujejo pomenske tipe ali podtipe, in sicer s konkretnjšimi predstavniki, ki pa so še vedno zastopniki večje skupine sorodnih primerkov, medtem ko predstavljajo **KOLOKATORJI** v vlogi semantičnih tipov konkretne leksikalne zapolitve na udeleženskih mestih. Z njimi smo označevali udeležence pri posameznih pomenih besede, kjer je realizacija pomena vezana na konkretno leksikalno zapolnitev. Taksonomijo semantičnih tipov s prostorskim pomenom v pomenskih shemah LBS prikazuje Tabela 1.

Tabela 1: Pomenska taksonomija za označevanje semantičnih tipov s prostorskim pomenom v pomenskih shemah LBS.

Pomensko polje	Pomenski tip	Pomenski podtip	Kolokatorski tip	Kolokator
KONKRETNA ENTITETA				
	LOKACIJA			
		OBMOČJE		morje, jezero, reka ...
				pokrajina, dolina, obala, otok ...
			OKOLICA	
			OKOLJE	
			DRŽAVA	dežela
			OZEMLJE	
		KRAJ		vas, mesto
		PROSTOR		hodnik, balkon, etaža ...

				sadovnjak, gozd, vrt, rastlinjak ...
			STAVBA	hiša
			OBJEKT	
		MESTO		delovno mesto
			TOČKA	vrh, postaja
		POVRŠINA		gladina
			ZEMLJA	prst, gredica
			TLA	dno
	SMER			
		CILJ		
		POT		

4.2 Stavčni vzorci

Stavčni vzorci, kot jih razumemo v kontekstu izdelave LBS, so formalni prikaz vezljivosti posameznega glagolskega pomena znotraj stavčne zgradbe, ali natančneje, prikaz različnih realiziranih (ne torej zgolj sistemsko možnih) skladenjskih in obliko-skladenjskih načinov uresničevanja določenega glagolskega pomena v zgradbi prostega stavka. Na ta način stavčni vzorci vzpostavljajo povezavo s pomensko shemo, ki vsebuje »največji« stavčni vzorec (v smislu pomensko in skladdenjsko nujnih, čeprav v besedilu ne vedno izraženih, določil in dopolnil), in s kolokacijami kot leksikalnimi zapolnitvami posameznih vezljivostnih (vezavnih in primičnih) mest.

Vezljivostna mesta, kot jih predvideva posamezni glagolski pomen, so v stavčnih vzorcih izražena s poljubnostnim zaimkom *kdo/kaj*, ki poleg sklona izraža tudi slovnico-pomenski kategoriji neživosti in človeškosti (TOPORIŠIČ 2004: 311), redkeje s prislovnimi zaimki kraja, npr. *kaj pade kam*; *kdo se pobere od kod*, ki so v kolokacijah zapolnjeni s prislovi: *pasti [dol, notri, stran]*; *vrniti se [domov, nazaj]*. Stavčne vzorce smo nato združevali v samostojne skupine glede na to, ali izražajo okoliščine, kot so kraj, smer in čas, ter lastnosti, kot so način pa tudi namen in vzrok. V pričujoči razpravi nas bodo zanimali predvsem glagoli oz. njihovi pomeni, kjer smo stavčne vzorce beležili pod skupnim imenovalcema LOKACIJA in SMER. Posamezni deli stavčnega vzorca se, če so kolokabilni, potrjujejo s kolokacijami.

5 Glagoli s prostorsko pomensko komponento v LBS

V LBS so glagoli oz. posamezni glagolski pomeni s prostorsko komponento obravnavani na tri načine: (a) v svoji udeleženski zgradbi, izraženi v pomenski shemi, lahko vsebujejo semantični tip, ki označuje prostorski pomen (gl. Tabela 1), ali pa je (b) prostorska komponenta vključena v stavčne vzorce, ki predstavljajo različne skladdenjske realizacije osnovnega stavčnega vzorca v pomenski shemi, združene pod skupni »prostorski« imenovalc: SMER ali LOKACIJA, npr. *če ŽIVAL ali ČLOVEK*

bega po PROSTORU, se po njem premika v različnih smereh, navadno zato, ker je preplašen ali zmeden. Tretja (c) možnost je vključitev prostorske komponente v t. i. scenarijski del stavčno oblikovane pomenske sheme, ki prinaša tiste pomenske vsebine, ki pridejo do izraza v konkretnih sporočanjih oz. v odnosu do in med udeleženci sporočanja. Glede na to, da pomenska shema za posamezni glagolski pomen vključuje najširši stavčni vzorec, tj. tako skladenjsko in pomensko obvezna kot neobvezna določila in dopolnila, izražena vsaj v večini realnih besedilnih realizacij, je logično, da se bodo v pomenski shemi zabeleženi semantični tipi s pomenom prostora realizirali tudi na ravni stavčnih vzorcev in se potrjevali s kolokacijami, ki vključujejo besede s prostorskim leksikalnim pomenom, npr. *begati po čem*: begati po [hodnikih, sobi].

V spodnjih primerih, s katerimi potrjujemo vse tri načine izražanja prostorskega pomena v LBS, navajamo najprej glagolsko iztočnico, ki ji sledijo posamezni prostorsko determinirani pomeni, zapisani v obliki pomenskih indikatorjev (ležeče).⁹ Indikatorju sledi pomenska shema kot izhodišče za oblikovanje stavčne razlage (v okvirčku).¹⁰ Semantični tipi, s katerimi označujemo udeleženska mesta, so zapisani z velikimi črkami. Prostorsko determinirani deli stavčnih vzorcev, ki sledijo, se lahko potrjujejo s konkretnimi leksikalnimi zapolnitvami v obliki kolokacij.¹¹ Kolokacijam sledijo korpusni zgledi, ki potrjujejo tiste stavčne vzorce, katerih prostorsko determinirani deli niso kolokabilni, jih je bilo pa mogoče ugotoviti iz analize konkordanc.

obiskati

iti kam z določenim namenom

če ČLOVEK obišče določeno USTANOVO, KRAJ ali DOGODEK, gre tja zaradi osebne ali službenega interesa ali potrebe

obiskati koga/kaj

- obiskati [državo, občino, domovino, deželo]
- obiskati [prestonico, kraj, mesto]
- obiskati [muzej, knjižnico, razstavo, galerijo, prireditve]

obiskati koga v čem

- obiskati v [bolnišnici, zaporu]

obiskati koga na čem

- obiskati na [domu]

pobrati

priiti iskat

če ČLOVEK pobere PREDMET ali ČLOVEKA na neki LOKACIJI, ga pride tja iskat, da bi ga odpeljal ali vzel s seboj

⁹ Pri večpomenskih glagolih izpostavljamo samo tiste pomene, ki vključujejo prostorski semantični tip ali izkazujejo prostorsko determinirane stavčne vzorce.

¹⁰ Oblikovanje stavčnih razlag na podlagi pomenskih shem je opisano v Gantar in Krek (2009), več o stavčnih razlagah pri prenosu na slovenščino pa Gantar, Krek in Kosem (2014) ter Šorli (2014).

¹¹ Kolokacije v LBS razumemo kot statistično izstopajoče vzajemno družljive besede. Na ta način ločujemo med individualnimi realizacijami prostorskega pomena in kolokacijami, ki so vedno statistično izstopajoče realizacije posameznega dela stavčnega vzorca. Na ta način se posredno izkazuje tudi relativna pogostost posameznega glagolskega pomena.

pobrati koga/kaj v čem
pobrati koga/kaj po čem
pobrati koga/kaj na čem
• pobrati na [letališču, postaji]

iti stran; oditi

če se ČLOVEK pobere z nekega KRAJA, od tam odide, navadno zato ker ga kdo preganja ali ker tam ni zaželen

pobrati se od česa/kod¹²

Večkrat sem zagrozil, da se bom pobral od doma.

Iz Metala so nam grozili, da bodo metali šarže na nas in naj se poberemo od tod.

pobrati se iz česa

Augusto zagrozi Margariti, naj se pobere iz Santiagovega življenja.

križariti

pluti

če ČLOVEK s PLOVILOM ali če PLOVILO križari po MORJU, REKI ali JEZERU, po njem pluje

križariti po čem

- križariti po [morju, oceanu], križariti po [Sredozemskem, Severnem] morju
- križariti med čim
- križariti med [otoki]

križariti okoli česa

križariti ob čem

- križariti ob [obali]

vreti

množično prihajati

če LJUDJE vrejo iz PROSTORA, množično prihajajo iz njega

vreti iz česa

Iz Opere trumoma vrejo ljudje, večinoma nedeljsko oblečeni.

vreti v kaj

Veliko romarsko središče privablja vernike celo iz Indije; reke ljudi vrejo vanj.

migrirati

seliti se

če ČLOVEK migrira, se seli iz enega KRAJA ali DRŽAVE v drug KRAJ ali DRŽAVO, navadno iz ekonomskih razlogov

migrirati v kaj

- migrirati v [Avstralijo]

migrirati na kaj

- migrirati na [Zahod, Florido]; migrirati na [delo]

¹² Čeprav v LBS, kot rečeno, nismo določali prostorskih udeležencev glede na način izražanja skladijskega razmerja, je to posredno razvidno – vsaj pri prislovih v vlogi krajevnih prislovnih določil – v zapisu stavčnega vzorca.

smučati

voziti se s smučmi

če ČLOVEK smuča, se s smučmi večkrat spusti po zasneženem in za to urejenem
OBMOČJU

smučati

smučati po čem

- smučati po [strmini, prog]

smučati na čem

- smučati na [smučišču, prog]

smučati z/s česa

- smučati z [vrha]

smučati v čem

- smučati v [tujini]

Prostorskost, kot je razvidno iz zgornjih primerov, se po pričakovanju izraža s predložnimi zvezami, kjer izstopajo zlasti predlogi *na*, *v*, *po* in drugi. Pri izražanju prostorskega pomena z brezpredložnim tožilnikom obstaja možnost pretvorbe v glagol premikanja ali nahajanja in ustreznega predloga: *obiskati frizerja, zdravnika, bolnišnico, sejem, soseda* → *iti k frizerju, zdravniku; iti v bolnico, na sejem, do soseda; nahajati se/ biti v bolnici, pri frizerju, sosedu, na sejmu*. Preureditev delovalniških vlog, ki vključuje prostorski pomen,¹³ je sicer mogoče opazovati tudi pri drugih členih stavčnega vzorca, npr. *razred buči od smeha* – *v razredu buči od smeha*; *srce stisne koga* – *pri srcu stisne koga*; *križ boli koga* – *v križu boli koga*, pri čemer leksikalne zapolnitve na teh mestih niso arbitrarne in vedno vzajemno zamenljive: *ušesa/v ušesih boli, križ/v križu boli* : *boli v prsih, glava boli* (neizkazano: *prsi bolijo, v glavi boli*).

V drugi skupini je mogoče obravnavati glagolske pomene, pri katerih se noben od udeležencev v pomenski shemi ne izraža s prostorskim semantičnim tipom, vendar pa je prostorska komponenta razvidna iz stavčnih vzorcev in potrjena s kolokacijami. Govoriti je mogoče o glagolih, kjer je prostorska komponenta običajen vendar ne pomensko nujen element glagolskega pomena.

anketirati

pridobivati podatke

če ČLOVEK ali USTANOVA anketira LJUDI, jim zastavlja vnaprej pripravljena vprašanja o določeni TEMI ali PROBLEMATIKI, o kateri želi pridobiti mnenje ali zbrati podatke

anketirati na čem

- anketirati na [ulici, terenu]

anketirati v čem

- anketirati v [Ljubljani]

¹³ Prim. tudi utemeljitev pomenskopretvorbenege razmerja med svojino in nahajanjem na nekem mestu (VIDOVIČ MUHA 1998: 294).

cveteti

pognati cvetove

RASTLINA cveti, ko so RAZMERE UGODNE, da lahko požene in razpre cvetove

kaj cveti na čem

- cveteti na [vrtu, polju]

kaj cveti v čem

- cveteti v [zimskem] vrtu

klestiti

povzročati škodo; uničevati

če NARAVNI POJAV klesti PRIDELEK, ga uničuje

kaj klesti po čem

- klestiti po [vinogradih, dolini]
- klestiti po [občini, Sloveniji]

kaj klesti v čem

nakupovati

pridobivati kaj za plačilo

če POTROŠNIK nakupuje IZDELKE ali DOBRINE, jih pridobiva v zameno za plačilo

nakupovati v čem

- nakupovati v [trgovini, supermarketu, nakupovalnem centru, trgovskem središču]
- nakupovati v [tujini]

nakupovati na čem

- nakupovati na [tržnici]

V tretji skupini moramo izpostaviti glagole, ki prostorsko komponento izražajo zgolj na ravni pomenskega scenarija. Glede na koncept pomenskega scenarija v LBS, bi lahko rekli, da gre za glagolske pomene, kjer je prostorska komponenta del pragmatičnih pomenskih vsebin (spodaj poudarjeno), npr.

blatiti

grdo govoriti o kom

če ČLOVEK blati drugega ČLOVEKA ali njegovo DEJAVNOST, o njem v **javnosti** govori slabe stvari, navadno zato, da bi škodil njegovemu ugledu

blatiti v [medijih, javnosti, časopisih, tujini]

blatiti po [časopisih, televiziji]

zadremati

zaspati za kratek čas

če ČLOVEK zadrema, za kratek čas zaspi, navadno podnevi in na **neobičajnem mestu**

zadremati za [volanom, mizo]

zadremati pri [mizi]

zadremati na [soncu]

6 Zaključek

Za prikazane načine vključevanja prostorske informacije v slovarski opis glagolskih pomenov v LBS je najbolj pomembno to, da prostorsko komponento posredujejo na podlagi realnih besedilnih realizacij, kar slovarskemu uporabniku posreduje vsaj tri vrste medsebojno povezanih relevantnih skladenjskopomenskih informacij.

Uporabnik najprej prek semantičnih tipov, ki so v stavčni razlagi nadomeščeni s konkretnimi kolokatorji oz. so abstrahirani do te mere, da zastopajo realno skupino pomensko sorodnih udeležencev, uzavesti udeležence (število in oblikoskladenjsko organizacijo znotraj stavka) in njihove tipične leksikalne realizacije, ki se eksplicitno potrjujejo s kolokacijami. Nadalje, prek različnih stavčnih vzorcev, ki predstavljajo realizacijske možnosti osnovnega stavčnega vzorca v pomenski shemi oz. stavčno oblikovani slovarski razlagi, uzavešča možnost (ne)opustljivosti oz. pomenske (ne)nenujnosti določenih – v našem primeru prostorskih – vsebin. Nenazadnje je za aktivno uporabo konkretnega glagolskega pomena v besedilu, kot tudi za njegovo ustrezno dekodiranje pomembno, da se prek posameznih stavčnih vzorcev sooči in posledično uzavesti ne le tiste skladenjske realizacije, ki so za določeni pomen tipične, pač pa predvsem tiste, ki so znotraj številnih jezikovnosistemskih možnosti v besedilih tudi dejansko realizirane.

Opredeljevanje prostorskih pomenskih komponent z vidika vezavnosti nasproti družljivosti (oz. ločevanje med določili in dopolnili), se nam ob dejstvu, da je mogoče govoriti tudi o številnih vmesnih primerih, opredeljenih s pojmi kot *vezavna družljivost* ipd., zdi za slovarsko informacijo manj pomembno. Ob tem je potrebno poudariti, da nam način zapisovanja semantičnih tipov v pomenski shemi in združevanje predložnih stavčnih vzorcev pod skupne imenovalce na ravni slovarske baze omogoča avtomatsko pridobivanje vseh glagolov, ki v slovarski bazi vključujejo pomensko komponento prostora (tudi časa, načina, vzroka in namena), ter možnost njihovega nadaljnega računalniškega procesiranja in uporabo v jezikovnotehnoloških aplikacijah za slovenščino.

VIRI IN LITERATURA

- Sue B. T. ATKINS, Michael RUNDALL, 2008: *The Oxford Guide to Practical Lexicography*. Oxford: University Press.
- Geoff BARNBROOK, 2002: *Defining Language: A Local Grammar of Definition Sentences*. Amsterdam, Filadelfija: John Benjamins Publishing Company.
- František ČERMÁK, 2010: *Lexikon a sémantika*. Praga: Nakladatelství Lidové Noviny.
- Anna DZIEMIANKO, Robert LEW, 2013: When-definitions revisited. *International Journal of Lexicography* 26/2. 154–75.
- Charles FILLMORE, Sue B. T. ATKINS, 1998: FrameNet and Lexicographic Relevance. *Proceedings of the First International Conference On Language Resources And Evaluation*. Granada, Španija, 28–30 Maj 1998.
- Charles FILLMORE, Christopher J. JOHNSON, Miriam R. L. PETRUCK, 2003: Background to Framenet. *International Journal of Lexicography* 16/3. 235–50.

- Polona GANTAR, Katja GRABNAR, Polonca KOCJANČIČ, Simon KREK, Olga POBIRK, Rok REJC, Mojca ŠORLI, Simon ŠUSTER, Petra ZARANŠEK, 2009: *Specifikacije za izdelavo leksikalne baze za slovenščino: standard za izdelavo posamezne leksikalne enote v leksikalni bazi*. Projekt Sporazumevanje v slovenskem jeziku ESS in MŠŠ.
- , 2009a: *Specifikacije za izdelavo leksikalne baze za slovenščino: Opis analize referenčnega korpusa*. Projekt Sporazumevanje v slovenskem jeziku ESS in MŠŠ.
- Polona GANTAR, Simon KREK, 2009: Drugačen pogled na slovarske definicije: Opisati, pojasniti, razložiti? *Infrastruktura slovenščine in slovenistike*. Ur. M. Stabej. Ljubljana: FF (Obdobja 28). 151–59.
- Patrick HANKS, 1987: Definitions and Explanations. *Looking Up: An Account of the COBUILD Project in lexical computing*. Ur. J. Sinclair. London: Collins. 116–36. Na spletu.
- , 1994: Linguistic Norms and Pragmatic Exploitations, or Why Lexicographers Need Prototype Theory and Vice Versa. *Papers in Computational Lexicography*. Complex '94. Ur. F. Keifer, G. Kiss, J. Pajzs. Budimpešta: Hungarian Academy of Sciences. 89–113.
- , 2004: Corpus Pattern Analysis. *Proceedings of XI Euralex International Congress*. Ur. G. Williams, S. Vessier. Lorient: Université de Bretagne-Sud. 87–97.
- , 2009: Sestavljanje enojezičnega slovarja za domače govorce. *Jezik in slovstvo* 54/3–4. 7–24.
- , 2013: *Lexical Analysis: Norms and Exploitations*. Cambridge: MIT Press.
- Patrick HANKS, James PUSTEJOVSKY, 2005: A Pattern Dictionary for Natural Language Processing. *Revue Française de Linguistique Appliquée* 10/2. 63–82.
- Adam KILGARRIFF, 1997: 'I Don't Believe in Word Senses'. *Computers in the Humanities* 31/2. 91–113.
- Adam KILGARRIFF, Pavel RYLCHÝ, Pavel SMRZ, David TUGWELL, 2004: The Sketch Engine. *Proceedings of XI Euralex International Congress*. Ur. G. Williams, S. Vessier. Lorient: Université de Bretagne-Sud. 105–16.
- Simon KREK, 2008: FrameNet in slovenščina. *Jezik in slovstvo* 53/5. 37–54.
- Simon KREK, Adam KILGARRIFF, 2006: Slovene Word Sketches. *Jezikovne tehnologije* 5. Ur. T. Erjavec, J. Gros. Ljubljana: IJS. 62–65.
- Simon KREK, Polona GANTAR, Iztok KOSEM, 2014: Razlage v novem slovarju slovenskega jezika. *Novi slovar za 21. stoletje*. E-zbornik prispevkov s Posveta o novem slovarju slovenskega jezika na Ministrstvu za kulturo 12. 2. 2014. Ur. I. Grahek, S. Bergoč. Na spletu.
- Nataša LOGAR, Miha GRČAR, Marko BRAKUS, Tomaž ERJAVEC, Špela ARHAR HOLDT, Simon KREK, 2012: Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES: Gradnja, vsebina, uporaba. Ljubljana: Trojina, zavod za uporabno slovenistiko, FDV.
- Mojca ŠORLI, 2014: Sodobni sporazumevalni slovar slovenskega jezika: Izhodišča, viri, izvedba. *Novi slovar za 21. stoletje*. E-zbornik prispevkov s Posveta o novem

slovarju slovenskega jezika na Ministrstvu za kulturo 12. 2. 2014. Ur. I. Grahek, S. Bergoč. Na spletu.

Jože TOPORIŠIČ, ⁴2004: *Slovenska slovnica*. Maribor: Obzorja.

Ada VIDOVIČ MUHA, 1998: Pomenski preplet glagolov imeti in biti – njuna jezikovno-sistemska stilistika. *Slavistična revija* 46/4. 293–323.

Andreja ŽELE, 2003: *Glagolska vezljivost: iz teorije v slovar*. Ljubljana: Založba ZRC, ZRC SAZU.

--, 2010: Elipsa med glagolsko intenco in besedilno koherenco (Izpust med glagolsko usmerjenostjo in besedilno soveznostjo). *Slavistična revija* 58/1. 117–31.

--, 2012: *Pomensko-skladenjske lastnosti slovenskega glagola*. Ljubljana: Založba ZRC, ZRC SAZU.

SUMMARY

The paper describes the design of semantic frames as formulated for the Slovene Lexical Database within the project Communication in Slovene, namely, for providing lexico-grammatical data in a dictionary in the form of sentence-level explanations (Hanks 1987, Barnbrook 2002, Dzemianko and Lew 2013) and for use with language technologies. The design of this type of descriptions is based on two theoretical models: (1) lexical meanings that are evoked on the basis of the so-called frame semantics in the project FrameNet (Fillmore 2003); (2) descriptions of syntactic patterns following the theory of linguistic conventions and possibilities of their use in the project Corpus Pattern Analysis (Hanks 1994, 2013; Hanks in Pustojevsky 2004, 2005). In building semantic frames, both models have been upgraded taking into consideration the specific features of Slovene and the purpose of the lexical database. When analyzing spatial information as is evident in sentence-level semantic frames within the semantic types, the author focuses on the verbs in which it is possible to detect a spatial determination of meaning using corpus analysis and analysis of lexical frames with the tool Sketch Engine (Kilgarriff et al. 2004). The semantic types are understood as abstract representatives of typical lexical realizations of individual arguments in a particular verbal meaning, e.g., *if an OBJECT due to force of gravity falls on some SPACE or SURFACE, it flies there from the air or from a place high above*. In addition, the methods of syntactic or semantic expression of the spatial component were analyzed. For this purpose the author recorded syntactic patterns typical of a particular lexical meaning, e.g., *pasti* ‘to fall’—‘to move upside down’: *kaj pade kam*, ‘something falls somewhere’, *kaj pade na kaj* ‘something falls onto something’, *kaj pade v kaj* ‘something falls into something’, etc., and syntactic patterns like: *pasti pod kotom x stopinj* ‘to fall at a x-degree angle’, *pasti z višine x metrov* ‘to fall from a height of x-meters’. Typical lexical realizations of arguments in the recorded syntactic patterns emerged in the form of collocations, for example [*bomba, granata, drevo*] *pade* ‘[bomb, grenade, tree] falls’, *pasti na [dno, zemljo, tla]* ‘to fall to [bottom, earth, floor]’, *pasti v [prepad, globino, morje]*, ‘to fall [over a precipice /into deepness/the sea].