

SPECIFIKACIJE ZA IZDELAVO LEKSIKALNE BAZE ZA SLOVENŠČINO

STANDARDI ZA IZDELAVO POSAMEZNE LEKSIKALNE ENOTE V LEKSIKALNI BAZI (KAZALNIK 6)

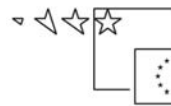
PROJEKT »SPORAZUMEVANJE V SLOVENSKEM JEZIKU«

Operacijo delno financira Evropska unija iz Evropskega socialnega sklada ter Ministrstvo za šolstvo in šport. Operacija se izvaja v okviru Operativnega programa razvoja človeških virov za obdobje 2007–2013, razvojne prioritete: razvoj človeških virov in vseživljenjskega učenja; prednostne usmeritve: izboljšanje kakovosti in učinkovitosti sistemov izobraževanja in usposabljanja 2007–2013.

Avtorji:

Polona Gantar, Katja Grabnar, Polonca Kocjančič, Simon Krek, Olga Pobirk, Rok Rejc, Mojca Šorli, Simon Šuster in Petra Zaranšek

Kamnik, 30. junij 2009



Kazalo vsebine

UVOD.....	9
1 TEORETIČNA IN METODOLOŠKA IZHODIŠČA.....	9
1.1 Kaj je leksikalna podatkovna baza.....	9
1.2 Pregled evropskih projektov, povezanih z leksikalno problematiko.....	12
1.3 Leksikalna podatkovna baza za slovenščino.....	14
1.3.1 Obseg.....	16
1.3.2 Zgradba.....	18
1.3.2.1 Organizacija leksikalnih podatkov po nivojih.....	18
1.3.2.2 Zgradba gesla v programskem vmesniku DPS (<i>Entry Editor</i>).....	23
1.3.2.2.1 Glava.....	24
1.3.2.2.2 Geslo.....	25
1.3.2.2.2.1 Pomeni in podpomeni.....	25
1.3.2.2.2.1.1 Skladienske skupine in skladienske zveze.....	26
1.3.2.2.2.1.2 Stalne besedne zveze.....	27
1.3.2.2.2.2 Frazeološke in idiomatske enote.....	28
1.3.3 Vsebina.....	29
1.3.3.1 Beleženje pomena: pomen ali pomenske tendence?.....	30
1.3.3.2 Opis pomena.....	32
1.3.3.2.1 New Oxford Dictionary of English.....	33
1.3.3.2.2 Macmillan English Dictionary.....	34
1.3.3.2.3 Longman Dictionary of Contemporary English.....	35
1.3.3.2.4 Pomenska shema in teorija FrameNet.....	36
1.3.3.3 Leksikalno-gramatični pristop: skladienske informacije v podatkovni bazi.....	40
1.3.3.4 Kolokacije.....	43
1.3.3.5 Večbesedne leksikalne enote.....	45
1.3.3.5.1 V razmerju do prostih besednih zvez.....	46
1.3.3.5.2 V razmerju do kolokacij.....	47



1.3.3.5.3 V razmerju do frazeoloških enot.....	47
1.3.4 Nadaljnje delo pri leksikalni bazi za slovenščino	48
2 STANDARDI ZA IZDELAVO LEKSIKALNE ENOTE V LEKSIKALNI BAZI ZA SLOVENŠČINO	49
2.1 Osnovni podatki o leksikalni enoti	49
2.2 Pomen leksikalne enote	50
2.2.1 Pomenski indikator.....	51
2.2.2 Pomensko razdvoumljanje.....	52
2.2.2.1 Kontekstualna informacija	52
2.2.2.2 Razmerja med jedrnim pomenom in podpomeni	54
2.2.2.2.1 Figurativni nasledek jedrnega pomena.....	55
2.2.2.2.2 Specializirani pomen jedrnega pomena.....	58
2.2.2.2.3 Druge razširitve in spremembe jedrnega pomena	59
2.2.2.3 Razmerja med podpomeni znotraj jedrnega pomena.....	60
2.2.2.4 Slovnične, funkcijske in pragmatične omejitve pomena.....	62
2.2.3 Razvrstitev pomenov.....	63
2.2.4 Pomensko in besednovrstno povezane LE	64
2.2.5 Homonimija in večpomenskost.....	65
2.2.6 Stalne besedne zveze	65
2.2.6.1 Pomensko podrejanje stalnih besednih zvez.....	68
2.2.6.2 Pomenski indikatorji pri stalnih besednih zvezah	69
2.2.6.3 Zapisovanje stalnih besednih zvez	70
2.2.7 Pomenska shema – splošno	72
2.2.7.1 Postopek oblikovanja pomenske sheme	73
2.2.7.2 Elementi pomenske sheme	74
2.2.7.3 Razmerje med pomenskimi indikatorji in pomenskimi shemami – primer glagola <i>pasti</i>	77
2.3 Skladenjski nivo	79
2.3.1 Skladenjske skupine.....	80
2.3.1.1 Glagol.....	81
2.3.1.1.1 Megastrukture	81
2.3.1.1.2 Skladenjske strukture pri glagolih.....	84



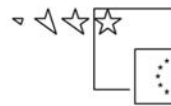
2.3.1.1.3 Skladenjski vzorci	85
2.3.1.1.3.1 Zapisovanje skladenjskih vzorcev	88
2.3.1.1.3.2 Razmerje med skladenjskimi vzorci in pomensko shemo.....	89
2.3.1.2. Samostalnik.....	91
2.3.1.2.1 Jedrne samostalniške strukture	91
2.3.1.2.2 Nejedrne samostalniške strukture	94
2.3.1.2.3 Zapisovanje samostalniških struktur	98
2.3.1.3 Pridevnik.....	99
2.3.1.3.1 Jedrne pridevniške strukture.....	99
2.3.1.3.2 Nejedrne pridevniške strukture	102
2.3.1.3.3 Zapisovanje pridevniških struktur	103
2.3.1.4 Prislov	103
2.3.1.4.1 Jedrne prislovne strukture	103
2.3.1.4.2 Nejedrne prislovne strukture	104
2.3.1.5 Kodiranje skladenjskih struktur in skladenjskih vzorcev	106
2.3.1.5.1 Glagol.....	106
2.3.1.5.2 Samostalnik.....	110
2.3.1.5.3 Pridevnik in prislov	111
2.3.1.6 Slovnične omejitve in preference	114
2.3.2 Skladenjske zveze.....	116
2.3.2.1 Skladenjske zveze pri glagolih	117
2.3.2.1.1 Kolokacijsko razširjeni segmenti jezika.....	117
2.3.2.1.2 Strukturno ustaljene skladenjske zveze	118
2.3.2.1.3 Zapis skladenjskih zvez.....	119
2.3.2.2 Skladenjske zveze pri samostalnikih.....	121
2.3.2.2.1 Vrste in zapis skladenjskih zvez pri samostalnikih.....	122
2.3.2.2.2 Primeri skladenjskih zvez (samostalnik <i>čas</i>)	123
2.3.2.3 Skladenjske zveze pri pridevnikih in prislovih	125
2.4 Kolokacijski nivo.....	126
2.4.1 Povezava med skladenjskim in kolokacijskim nivojem v LBS	127
2.4.1.1 Pri glagolu	129



2.4.1.2 Pri pridevnikih in prislovih	131
2.4.2 Zapis kolokacij v programskem vmesniku	132
2.5 Frazeologija in idiomatika	136
2.5.1 Pomenska členitev frazeoloških in idiomatskih enot.....	137
2.5.2 Zapisovanje frazeoloških enot.....	138
2.6 Primeri rabe	141
3 LITERATURA	143

Kazalo slik

Slika 1: Vrste leksikalnih podatkovnih baz.....	12
Slika 2: Element <i>glava</i> v programskem vmesniku.....	24
Slika 3: Elementa <i>pomen</i> in <i>podpomen</i> v programskem vmesniku.....	25
Slika 4: Nadelementa <i>skladenjske skupine</i> in <i>skladenjske zveze</i> v programskem vmesniku.....	26
Slika 5: Nadelement <i>stalne zveze</i> v programskem vmesniku	28
Slika 6: Nadelementa <i>frazeološke</i> in <i>idiomatske zveze</i> v programskem vmesniku.....	28
Slika 7: Meni pomenov večpomenske LE <i>time</i> (čas) v MED.....	34
Slika 8: Opis jedrnih elementov pomenskega okvirja <i>Manipulation</i> za glagol <i>squeeze</i> v FrameNetu	38
Slika 9: Zapis variantnih sestavin stalne besedne zveze v programskem vmesniku.....	71
Slika 10: Analiza konkordančnega niza z vidika udeleženske zgradbe za glagol <i>stisniti</i>	72
Slika 11: Kodiran zapis megastrukture za glagol <i>pogajati se</i>	83
Slika 12: Zapis skladenjskih vzorcev v 1. in 2. fazi izdelave LBS.....	108
Slika 13: Zapis glagolskih struktur v 1. in 2. fazi izdelave LBS.....	109
Slika 14: Označevanje restrikcij pri glagolskih vzorcih v programskem vmesniku	114
Slika 15: Zapis zaporednih skladenjskih zvez znotraj enega nadelementa.....	119
Slika 16: Zapis samostojnih skladenjskih zvez znotraj ločenih nadelementov.....	119
Slika 17: Zapis številčnih delov zvez v skladenjskih zvezah.....	120
Slika 18: Kolokabilni deli skladenjskega vzorca za glagol <i>pogajati se</i>	127



Slika 19: Kolokabilni deli skladijske strukture pri samostalniku <i>mir</i>	128
Slika 20: Zgledi rabe za podpomen samostalnika <i>rana</i> v programskem vmesniku	142

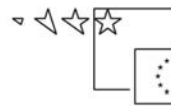
Kazalo tabel

Tabela 1: Ožji geslovnik – deleži besednih vrst v odstotkih in glede na število lem	17
Tabela 2: Razporeditev podatkov v LBS po vsebinskih nivojih.....	19
Tabela 3: Podrobna razčlenitev podatkov po nivojih: glagol <i>stisniti</i> in samostalnik <i>spomin</i>	20
Tabela 4 : Pomensko in besednovrstno povezane LE.....	65
Tabela 5: Skladijske strukture pri glagolih – (a).....	84
Tabela 6: Skladijske strukture pri glagolih – (b).....	84
Tabela 7: Jedrne strukture pri samostalnikih – (a).....	92
Tabela 8: Jedrne strukture pri samostalnikih – (b).....	93
Tabela 9: Nejedrne strukture pri samostalnikih – (a).....	94
Tabela 10: Nejedrne strukture pri samostalnikih – (b).....	95
Tabela 11: Nejedrne strukture pri samostalnikih – (c)	97
Tabela 12: Jedrne strukture pri pridevnikih – (a).....	99
Tabela 13: Jedrne strukture pri pridevnikih – (b).....	101
Tabela 14: Nejedrne strukture pri pridevnikih.....	102
Tabela 15: Jedrne strukture pri prislovih	104
Tabela 16: Nejedrne strukture pri prislovih – (a).....	104
Tabela 17: Nejedrne strukture pri prislovih – (b).....	105
Tabela 18: Elementi v skladijskih vzorcih, njihov opis, vrednost atributa in zapis	107
Tabela 19: Elementi v glagolskih strukturah, njihov opis, vrednost atributa in zapis	107
Tabela 20: Elementi v samostalniških strukturah, njihov opis, vrednost atributa in zapis	111
Tabela 21: Kodiranje pridevniških struktur.....	112
Tabela 22: Kodiranje prislovnih struktur	113



Kazalo zgledov

Zgled 1: Pomenska členitev in razvrstitev pomenov v NODE in LB za iztočnico <i>spomin</i>	33
Zgled 2: Primerjava definicij v MED in pomenskih indikatorjev v LBS za iztočnico <i>čas</i>	35
Zgled 3: Pomenski indikatorji jedrnih pomenov za iztočnico <i>spomin</i> v LDOCE in LBS	36
Zgled 4: Jedrne in nejedrne strukture ter pomeni pri samostalniških iztočnicah	42
Zgled 5: Primerjava definicij v slovarju <i>BRIDGE</i> in indikatorjev v LBS za iztočnico <i>spomin</i>	52
Zgled 6: Figurativni nasledki jedrnega pomena za glagol <i>pasti</i> in samostalnik <i>rana</i>	55
Zgled 7: Osamosvojitve figurativnega pomena glagola <i>pasti</i> in samostalnika <i>rana</i>	56
Zgled 8: Metonimično povezani podpomeni znotraj jedrnega pomena samostalnika <i>šola</i>	56
Zgled 9: Specializirani pomen jedrnega pomena.....	58
Zgled 10: Večpomenskost specializiranega pomena za samostalnik <i>spomin</i>	58
Zgled 11: Spremembe pomena, ki ohranjajo enega ali več elementov jedrnega pomena.....	59
Zgled 12: Jedrni pomen kot nadpomenka za pripadajoče podpomene.....	60
Zgled 13: Slovnične in funkcijske omejitve pomena.....	62
Zgled 14: Pragmatične omejitve pomena	63
Zgled 15: Pomensko podrejanje stalnih besednih zvez.....	68
Zgled 16: Elementi pomenske sheme za pomen 'ukrasti' glagola <i>vzeti</i>	75
Zgled 17: Razmerje med pomenskimi indikatorji in pomenskimi shemami pri glagolu <i>pasti</i>	77
Zgled 18: Združevanje struktur v megastrukturo pri glagolu.....	82
Zgled 19: Povezava med skladijskimi vzorci in kolokacijami pri glagolu <i>roditi</i>	86
Zgled 20: Povezava med skladijskimi vzorci in kolokacijami pri glagolu <i>pasti</i>	86
Zgled 21: Zapis skladijskih struktur z vzorci	87
Zgled 22: Zapis skladijskih vzorcev na podlagi konkordanc.....	87
Zgled 23: Razširitev pridevniških struktur z elementi besedilnega okolja	100
Zgled 24: Razširitev glagolskih kolokacij v skladijske zveze – (a)	117
Zgled 25: Razširitev glagolskih kolokacij v skladijske zveze – (b)	117
Zgled 26: Razširitev glagolskih kolokacij s količinskimi izrazi.....	118



Zgled 27: Strukturno ustaljene glagolske zveze	118
Zgled 28: Skladenjske zveze pri samostalniku <i>čas</i>	123
Zgled 29: Skladenjske zveze pri pridevnikih in prislovih	125
Zgled 30: Zapis kolokacij pri glagolskih iztočnicah.....	129
Zgled 31: Potrjevanje glagolskih struktur s kolokacijami	130
Zgled 32: Kolokacije pri pridevnikih in prislovih.....	131
Zgled 33: Potrditev pomena z zgledom (ob odsotnosti tipičnih struktur in kolokacij).....	131
Zgled 34: Potrditev pomena z vzorcem (ob odsotnosti tipičnih kolokacij).....	132
Zgled 35: Pomensko združevanje kolokatorjev znotraj kolokacijskega niza.....	133
Zgled 36: Združevanje kolokatorjev glede na slovnične kategorije	134
Zgled 37: Večbesedni kolokatorji	135
Zgled 38: Večpomenski kolokatorji.....	135
Zgled 39: Leksikalno-gramatične informacije v LBS: od abstraktne pomenske členitve do konkretnih primerov rabe	141



UVOD

Namen specifikacij za izdelavo leksikalne podatkovne baze za slovenščino (dalje LBS) je dvojen:

- opisujejo teoretična in metodološka izhodišča pri oblikovanju LBS s pregledom sorodnih jezikovnotehnoloških projektov po evropskih jezikih;
- predstavljajo navodila za izdelavo posamezne leksikalne enote (LE) v leksikalni bazi.

Poleg omenjenega je namen specifikacij na splošno opredeliti vlogo, namen in možnosti dolgoročne in večkratne izrabe LBS ter izpostaviti možnosti in potrebe za njeno nadgradnjo in vzdrževanje.

1 TEORETIČNA IN METODOLOŠKA IZHODIŠČA

V nadaljevanju opredelimo pojem *leksikalna podatkovna baza*, predstavimo različne vrste leksikalnih baz glede na namen njihove izrabe, na kratko predstavimo sorodne jezikovnotehnološke projekte po posameznih evropskih jezikih ter LBS v okviru projekta »Sporazumevanje v slovenskem jeziku« (dalje SSJ). Znotraj tega opredelimo njeno namembnost, obseg, strukturiranost ter vrsto jezikovnih podatkov, ki jih vključuje. Posamezne podatke opredelimo s teoretičnega in metodološkega vidika. Poglavje zaključimo z opredelitvijo in potekom nadaljnjega dela v okviru projekta.

1.1 Kaj je leksikalna podatkovna baza



Leksikalna baza¹ je organiziran opis besedišča določenega jezika v elektronski obliki ali na kratko: zbirka podatkov o leksikalnih enotah.² S pojmom *besedišče* pa nimamo v mislih samo zbira vseh leksikalnih enot konkretnega jezika, pač pa vse znanje, ki ga obsega besedišče maternega govorca, da se ta lahko sporazumeva v svojem jeziku.

Nekoliko bolj praktičen pogled, kot ga zastopa Patrick Hanks,³ opredeljuje leksikalno podatkovno bazo kot skup (korpusnih) dokazov, ki predstavljajo vsako LE v jeziku, pri čemer je poudarek na sintagmatiki oz. kolokabilnosti, pomenu in morfologiji. Taka leksikalna podatkovna baza velja kot primarni vir za številne aplikacije, zlasti za izdelavo slovarjev, jezikovnih učbenikov, popravljanje napak in učenje jezika, računalniško obdelavo naravnega jezika (ang. NLP: Natural Language Processing) in kodifikacijo besednih pomenov.

Če obstaja do neke mere soglasje o tem, da je leksikalna podatkovna baza skladišče leksikalnih informacij določenega jezika, pa obstajajo v jezikoslovni teoriji in praksi različna razumevanja pojma besedišče jezika in nanj vezanih informacij. V jezikoslovni literaturi sta se načeloma izoblikovali dve skrajni teoriji pojmovanja tega, čemur pravimo besedišče jezika, in sicer v razmerju do slovnice oz. položaja leksike v jeziku kot celoti. Chomskyjeva generativna slovnica temelji na ločevanju leksikona kot zbira leksikalnih enot, predvsem tistih, ki jih ni mogoče tvoriti oz. generirati iz obstoječih slovničnih pravil jezika. Na drugi strani – do neke mere izhajajoč iz same generativne slovnice (prim. Ooi 1998: 6/7) – so se uveljavila izhodišča, po katerih ne obstaja (bistvena) razlika med semantičnim oz. jezikovnim in enciklopedičnim znanjem oz. med besediščem in preostalim delom slovnice (Jackendoff 1995: 138). V sodobnih jezikoslovnih študijah se pojmovanje leksikona vse bolj usmerja v pripisovanje skladenjskih informacij, za katere je še do nedavnega veljalo, da so izpeljive iz slovničnih pravil, leksikalnim enotam oz. t. i. leksikalnim

¹ Glossary Of Linguistic Terms: <http://www.sil.org/linguistics/GlossaryOfLinguisticTerms/Index.htm>.

² <http://www.thefreedictionary.com/lexical+database>.

³ Gl. opombo 5.



pravilom. Koncept takega razumevanja širšega ali vsevključujočega (ang. *richer*; *all-inclusive*) leksikona zahteva torej kombinacijo in integracijo fonoloških, morfoloških, kolokacijskih, skladijskih, pomenskih in pragmatičnih informacij na različne načine. Skupni imenovalc vsake leksikalne baze je torej: inventar vseh⁴ v jeziku znanih LE; prikaz pomenov in oblik LE; kategorizacija LE, zlasti z vidika slovničnih in pomenskih kategorij, ki so LE imanentne, razmerja med LE ter njihova ustrezna raba.⁵ Pri tem je treba poudariti, da je izhodišče LBS v prvi vrsti opis leksikalnega potenciala slovenščine, medtem ko je interpretacija in predpis stvar slovarskih opisov in drugih aplikativnih del, ki temeljijo na leksikalni bazi. Bistvena informacija, na katero se navadno osredotočajo leksikalne podatkovne baze, je pomen LE. V tem smislu je vsaka LE v bazi pomensko razčlenjena; opremljena z zgledom, ki ponazarja posamezni pomen, opredeljena s kulturološkimi in pragmatičnimi oznakami, ki so pomembne za posamezni pomen, in opredeljena glede na semantične povezave z drugimi besedami. Poleg omenjenega, temelji bistvo jezikovnega opisa v sodobnih leksikalnogramatičnih leksikalnih bazah, ki so oblikovane na podlagi korpusnih podatkov, na spoznanju, da pomenski opis LE ni strogo ločen od njenega skladijskega opisa oz. povedano drugače, LE se poleg pomenskih pripisujejo tudi vse tiste slovnične in skladijske informacije, ki so potrebne za razumevanje pomena in rabe posamezne LE, in sicer ne glede na to, da jih je mogoče izpeljati iz slovničnih pravil, ki delujejo znotraj posameznega jezika.

Vsaka leksikalna baza je specifična glede na namen, za katerega je ustvarjena, je pa vsaka leksikalna baza vedno tudi: sekundarni jezikovni vir (podatki v LB so vedno na nek način hierarhizirani, kar zahteva določen jezikoslovni premislek, katerega posledica je vsaj do neke mere tudi jezikoslovna interpretacija jezikovnih podatkov); vsebuje več informacij, kot jih dejansko

⁴ Čeprav leksikalne baze poudarjajo težnjo po beleženju vseh v jeziku rabljenih LE, je taka težnja omejena najprej že iz povsem tehničnih vidikov, saj je njen obseg odvisen od časa izdelave, finančnih sredstev, števila sodelavcev itd., drugič pa tudi zato, ker se besedišče jezika in pomenska vrednost LE nenehno spreminjata, iz česar sledi, da je vsaka leksikalna baza zgolj presek danih možnosti v konkretnem obdobju in konkretnih namenov, za katere se oblikuje.

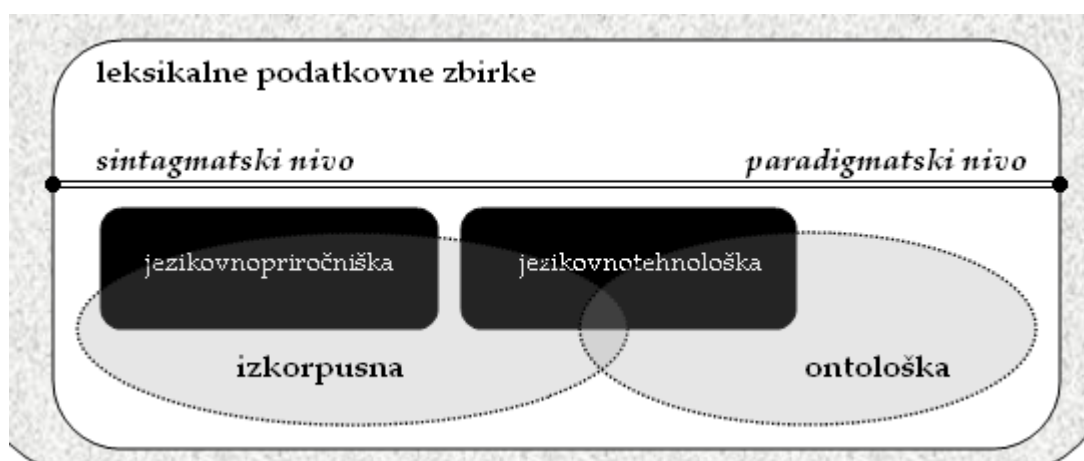
⁵ Zlasti leksikalne baze, ki so namenjene izdelavi enojezičnih slovarjev, naj bi upoštevale dejstvo, da materni govornici uporabljajo slovarje za preverjanje/konzultiranje glede ustrezne rabe in izbire besed (Hanks 2009).

vključujejo iz nje izhajajoči produkti oz. priročniki. Leksikalna baza naj bi bila v prvi vrsti baza podatkov in ne elektronska verzija (natisnjene) slovarja. Sodobna leksikalna baza je vedno zbir podatkov v elektronski obliki, kar pomeni, da je iz nje izhajajoče produkte mogoče izpeljati hitro in čim bolj avtomatično.

1.2 Pregled evropskih projektov, povezanih z leksikalno problematiko

Vprašanje, katere informacije naj bi leksikalna baza vsebovala, je v prvi vrsti vezano na to, za katere namene je oblikovana (Ooi, 1998: 11). Na splošno pa je mogoče ločevati med t. i. jezikovnotehnološkimi leksikalnimi bazami, kot so denimo STO, WordNet, EuroWordNet, Mikrokosmos, SENSUS ipd., in jezikovnopriročniškimi leksikalnimi bazami, ki so navadno v lasti založniških hiš, ki slovarje pripravljajo, npr. za slovenščino: leksikalna baza za izdelavo večjezičnih slovenskih-x slovarjev (Gorjanc, Krek in Gantar 2005), obrnjena angleško-slovenska slovarska podatkovna baza (Krek, Šorli in Kocjančič 2008), ki je nastala na podlagi Oxford-DZS-jevega Velikega angleško-slovenskega slovarja (Grabnar in Šorli 2003; Krek 2005–6; Šorli et al. 2006).

Slika 1: Vrste leksikalnih podatkovnih baz





Pojem leksikalne baze je v evropskem jezikoslovnem kontekstu znan vsaj že 20 let, dejavnosti v zvezi s tem pa so se začele z evropskimi jezikovnotehnološkimi projekti v 90-ih letih. Takrat še Evropska gospodarska skupnost in posamezne močnejše evropske države so jezikovne tehnologije in vire za »svoje« jezike izdatno raziskovalno podprle in v tem času – od leta 1993 v okviru Evropske unije – so bili za leksikalne vire jezikov EU izdelani predlogi standardov, npr. *EAGLES: Expert Advisory Group on Language Engineering Standards, 1993–1996*, ter digitalni leksikalni viri v okviru projektov, med katerimi je treba omeniti zlasti ACQUILEX (*Acquisition of Lexical Knowledge, 198–1995*), MULTILEX (*A Multi-Functional Standardised Lexicon for European Community Languages, 1990–1993*), Genelex I⁶ in Genelex II (*GENERIC LEXICON, 1990–1994*), CEGLEX (*Central European Genelex Model, 1995–1996*), DELIS (*Descriptive Lexical Specifications and Tools for Corpus-based Lexicon building, 1993–1995*), MECOLB (*Multilingual Environment for Corpus-Based Lexicon Building, 1994–1995*), PAROLE (*Preparatory Action for Linguistic Resources Organization for Language Engineering, 1996–1998*) in njegovo nadaljevanje SIMPLE (*Semantic Information for Multifunctional Plurilingual Lexicons, 1998–2000*), MULTEXT (*Multilingual Text Tools and Corpora, 1994–1996*), Multext-East (*Multilingual Text Tools and Corpora for Central and Eastern European Languages, 1995–1997*), ELAN (*European Language Activity Network, 1998–1999*), CONCEDE (*Consortium for Central European Dictionary Encoding, 1998–2000*) in TELRI (*Trans European Language Resources Infrastructure, I: 1995–1998 in I 1999–2001*). Hkrati so se oblikovala telesa in organizacije za hrambo in distribucijo leksikalnih virov: RELATOR (*A European Network of Repositories for Linguistic Resources, 1993–1995*) in ELRA (*European Language Resources Association, 1995–*).

⁶ Akronim za generični leksikon, tj. teorijo, ki jo je konec 80. in v začetku 90. let razvil James Pustejovsky s sodelavci (Pustejovsky 1995), in katere namen je ponuditi potrebno infrastrukturo za v celoti generativno, podrobno razčlenjeno in leksikalno orientirano semantično teorijo jezika. Prav ta teorija je tudi v izhodšču projekta CPA (Corpus Pattern Analysis; Hanks 2004; Hanks in Pustejovsky 2005).



Znotraj omenjenih projektov je bil denimo temeljni cilj projektov GENELEX I in II sistematično združiti izkušnje z razvojem virov in orodij za potrebe računalniške obdelave naravnih jezikov ter oblikovati standarde/modele za izdelavo elektronskih slovarjev evropskih jezikov. Od tod je bil način strukturiranja informacij po posameznih ravninah, morfološki, skladenjski in semantični, prenesen v druge evropske projekte:

- PAROLE (1993–1998) projekt, katerega rezultat so usklajene (tj. izdelane z isto metodologijo in opisane v skupnem DTD-ju) morfološko in sintaktično označene elektronske leksikalne zbirke za 13 jezikov takratne EU (zbirke so na začetku vsebovale 20. 000 besed);
- SIMPLE (1998–2002) kot nadgradnja projekta PAROLE s semantičnimi informacijami (vse podatkovne zbirke imajo skupno jedro osnovnih konceptov t. i. base concepts), povzetih po EuroWordNetu;
- DELIS (Descriptive Lexical Specifications and Tools for Corpus-based Lexicon Building) – gradnja korpusa za pridobivanje leksikalnih podatkov in razvoj semantičnega okvira za opis podatkov. Za semantično označevanje priredijo in dopolnijo model, ki temelji na teoriji generativnega leksikona J. Pustejovskyja. Reprezentacija podatkov temelji na razširjenem modelu Genelex.

Leksikalne podatkovne zbirke in njim sorodni projekti po posameznih evropskih jezikih so v večji ali manjši meri nasledki/nadgradnje omenjenih projektov, med njimi omenimo zlasti: [elexico](#) za nemščino, [CLIPS](#) za italijanščino, CORNETTO za nizozemščino, DAFLES in [ALFALEX](#) za francoščino, [STO](#) za danščino, [ADESSE](#) in [GRIAL](#) za španščino, [CEGLEX](#) za poljščino, [SPRÅKBANKEN](#) za švedščino, [PRALED](#) za češčino itd. Leksikalne baze, ki so nastale v okviru EU projektov za posamezne jezike, so v prvi vrsti namenjene računalniški obdelavi naravnih jezikov, v nekaterih segmentih pa tudi izdelavi spletnih slovarjev in on-line pripomočkov za učenje jezika.

1.3 Leksikalna podatkovna baza za slovenščino



Oblikovanje leksikalne podatkovne baze za slovenščino v okviru projekta SSJ sledi dvema temeljnima ciljema:

- zadostiti želi leksikografskim potrebam, konkretnije izdelavi za slovenščino najpomembnejših eno- in dvojezičnih slovarjev ter specializiranih jezikovnih priročnikov ter
- potrebam računalniške obdelave naravnega jezika oz. uporabi v jezikovnotehnoloških aplikacijah za slovenščino.

Da bi pri oblikovanju LBS lahko sledili zastavljenim ciljem, smo pri njeni zasnovi upoštevali naslednje:

- s podatki, vključenimi v LBS, mora biti mogoče v čim krajšem času priti do konkretnih, za jezikovno skupnost najbolj strateško pomembnih jezikovnih priročnikov, zlasti slovarja sodobnega slovenskega jezika, dvojezičnih slovarjev, posameznih poglavij slovnice, specializiranih slovarjev (šolski, frazeološki, frazni itd.);
- pri oblikovanju LBS se osredotočamo predvsem na to, kar je v jeziku osrednje/tipično: zanima nas torej jedrno besedišče in jedrni pomeni, skladijske strukture in vzorci, kolokacije;
- osredotočamo se na podrobno in natančno pomensko členitev jedrnega besedišča;
- poleg pomenskih nas zanimajo tudi skladijski in slovnični podatki, ki so pomembni za slovarske in slovnične opise in za izdelavo slogovnega priročnika za slovenščino.

Jezikovni podatki bodo v leksikalni bazi shranjeni v elektronski obliki, ki omogoča v prvi vrsti izdelavo elektronskih (po potrebi pa tudi tiskanih) slovarjev. E-slovarji ponujajo v slovarski prihodnosti vrsto zanimivih možnosti predstavitve novih informacij na povsem nove načine. Ključne lastnosti e-slovarjev prihodnosti naj bi bile predvsem prilagodljivost uporabniku in osebna nota (Atkins in Rundell, 2008: 239). Tako pojmovan »slovar« je dejansko zbirka leksikalnih virov (po možnosti večjezičnih), med katerimi lahko uporabnik izbira in jih upravlja glede na svoje potrebe.



1.3.1 Obseg

Leksikalna baza za slovenščino bo vsebovala podatke o frekvenčnih, oblikoslovnih, pomenskih, skladijskih, kolokacijskih, frazeoloških in drugih lastnostih jedrnega besedišča sodobne slovenščine.

V končni obliki bo LBS v okviru projekta SSJ vsebovala približno 2500 iztočnic predmetnopomenskih besed, in sicer samostalnikov, glagolov, pridevnikov in prislovov. LBS ne bo vsebovala slovničnih (veznik in predlog) ter funkcijskih (členek in medmet) besednih vrst. So pa zlasti predlogi in vezniki vključeni v LBS znotraj skladijskih in kolokacijskih segmentov pri predmetnopomenskih besednih vrstah, kar pomeni, da je mogoče naknadno pridobiti tudi informacije o pomenskih in skladijskih lastnostih predlogov in veznikov ter njihovo vlogo v ustaljenih tipičnih segmentih jezika. LBS tudi ne bo vsebovala pomenskih razlag/definicij, ker je odločitev za tip razlage/definicije, odvisna od namena, obsega in ciljnega uporabnika vsakega konkretnega slovarja, ki bo nastal na podlagi LBS. Pomenska členitev LE, ki je središčni del LBS in hierarhično nadrejena informacija, bo prikazana s pomenskimi indikatorji in s pomenskimi shemami, ki jih navajamo pri glagolskih (poskusno tudi pri samostalniških in pridevniških) iztočnicah. Zlasti pomenske sheme predstavljajo izhodišče in pomoč leksikografom za oblikovanje razlag opisnega (komunikacijskega) tipa (Sinclair in Coulthard 1975).

V LBS tudi ne beležimo lastnih imen, kadar nastopajo kot iztočnica, stalna besedna zveza, npr. *Partnerstvo za mir, Klub za iskanje zaposlitve, Glasbena matica*, ali frazeološka enota. Vključujemo jih, kadar so del širše občne zveze ali kolokacije: *roditi se v [Ljubljani, Mariboru ...]*. Ločnice med lastnimi imeni in prehodom v občno ime ni vedno lahko potegniti, zato se poleg pogostnosti zanašamo tudi na potrebo po pomenski opredelitvi take besede ali zveze, kar je zadosten razlog za uvrstitev v LBS.



V izhodišču smo za LBS oblikovali dve listi besed, tj. širši in ožji geslovník. Širši geslovník je oblikovan na podlagi 620-milijonskega referenčnega korpusa FidaPLUS v obsegu 5000 najpogostejših polno oz. predmetnopomenskih lem. Znotraj tega smo izločili 2500 samostalniških, glagolskih, pridevniških in prislovnih lem v naslednjih razmerjih: samostalnik 40 %, glagol 30 %, pridevnik 25 % in prislov 5 %, kar ustreza dejanski distribuciji predmetnopomenskih besednih vrst v jezikovni rabi in predstavlja ožji geslovník (Priloga 1).

Tabela 1: Ožji geslovník – deleži besednih vrst v odstotkih in glede na število lem

	ODSTOTEK	ŠTEVILO (/2500)
SAMOSTALNIK	40	1000
GLAGOL	30	750
PRIDEVNIK	25	625
PRISLOV	5	125
Skupaj	100	2500

Zaradi večje usklajenosti pri obdelavi pomensko povezanih besednovrstno različnih lem, npr. glagol: *mešati* – deležnik: *mešan* – glagolnik: *mešanje*, ali glagol: *pasti* – pridevnik: *padel* – posamostaljeni pridevnik: *padli* – samostalnik: *pad*, in besednodružinskih povezav med njimi, npr. *pomešati* – *premešati* – *zamešati*; *pasti* – *padati*; *praktičen* – *praktično* – *praktičnost* itd., smo se odločili, da končnega seznama lem znotraj odstotkovno določenih deležev med besednimi vrstami ne določimo vnaprej. Na ta način omogočimo tudi vzporedno obdelavo pomensko oz. tematsko povezanih LE⁷ (ang. lexical sets), pri čemer se zdi smiselno ločevati med odprtimi npr. sadje (*jabolko, hruška, sliva*); barve (*rdeča, črna, zelena* ...) in zaprtimi (dnevi v tednu itd.) leksikalnimi nizi. Z namenom, da bi v LBS prikazali jedrni del slovenskega besedišča in hkrati zajeli čim bolj raznolike tematske tipe LE, nam fleksibilnost geslovníka omogoča tudi vključevanje posameznih

⁷ Po nekaterih teorijah (Atkins in Rundell 2008: 127) bi bilo mogoče na tak način obdelati okrog 25 % besedišča.



LE kot »predstavnic« svoje skupine v pomensko zaključenih nizih, kot so npr. imena mesecev (npr. *januar, februar ...*), dnevi v tednu (npr. *ponedeljek, torek ...*), merske enote (npr. *kilogram, meter, ura ...*), jeziki (*slovenščina, angleščina*) ipd. in ne celotnega pomenskega niza. S tem, ko seznama lem v geslovníku ne določimo trdno že vnaprej, imamo tudi možnost vključevanja oz. izločanja sinonimno povezanih LE, ki zahtevajo razmeroma enako obdelavo v leksikalni bazi, nimajo pa nujno tudi enako pomembne pogostosti, npr. *časnik – časopis; predor – tunel*.

1.3.2 Zgradba

Leksikalna baza za slovenščino je zasnovana tako, da predstavlja na eni strani skladišče leksikalnih enot, ki tvorijo jedrno besedišče slovenskega jezika, hkrati pa je tudi skladišče jezikovnih informacij, vezanih na posamezne iztočnice. Dvojna narava LBS, ki je povezana z namenom njene izrabe v leksikografske in jezikovnotehnološke namene, določa zgradbo celotne leksikalne baze po vsebinskih nivojih. Posamezni nivoji, ki jih bomo predstavili v nadaljevanju predvsem z vsebinsko teoretičnega vidika in z vidika hierarhične zgradbe, so zajeti v XML zapisu geselske strukture, t. i. DTD-ju (ang. Document Type Definition) in vključeni v programsko opremo DPS (ang. Dictionary Publishing System) podjetja IDM. Zapis geselske strukture v XML formatu in programska oprema za izdelavo podatkovne baze je podrobno predstavljena v samostojnem dokumentu: Specifikacije za izdelavo leksikalne baze za slovenščino: Opis analize referenčnega korpusa (Kazalnik 5).

1.3.2.1 Organizacija leksikalnih podatkov po nivojih

Vsako iztočnico (lemo) analiziramo v leksikalni bazi s treh temeljnih vidikov: semantičnega, skladišjskega in kolokacijskega oz. sintagmatskega. Informacije, ki jih prinašajo izpostavljeni vidiki obravnavane LE, so v LBS organizirani po nivojih. Znotraj posameznega nivoja je za vsako LE oz. njen pomen predvidenih 5 različnih tipov jezikovnih podatkov. Na semantičnem nivoju pripišemo LE pomen(e) oz. podpomen(e), ki jih opredelimo s t. i. pomenskimi indikatorji. Drugi del pomenske informacije predstavlja t. i. pomenska shema, ki se teoretično približuje pomenskimi shemam, kot jih predvideva projekt FrameNet. Skladišjski nivo je namenjen formaliziranju skladišjskih struktur, v katerih se realizira posamezni pomen obravnavane LE. Ta del LBS je v



prvi vrsti namenjen procesiranju naravnega jezika – tj. računalniškemu branju jezikovnih podatkov. Leksikografsko verzijo skladijskih struktur zapisujemo s skladijskimi vzorci, s katerimi beležimo tudi konkretne slovnične omejitve LE oz. t. i. koligacije. Tretji nivo je kolokacijski. Tu navajamo tipične kolokacije, v katere vstopajo obravnavane LE glede na predhodno registrirane pomene in skladijske strukture. Posebej evidentiramo tudi razširjene kolokacijske vzorce in omejene zveze, ki predstavljajo tipične delčke jezika, ki pa niso leksikalizirane do te mere, da bi potrebovale samostojno razlago. Frazeološke enote, v katerih nastopa obravnavana LE, navajamo v samostojnem razdelku LB. Vsaki FE pripišemo pomen v obliki pomenskega indikatorja, hkrati pa evidentiramo tudi njene realizacijske oblike in korpusne zglede.

Tabela 2: Razporeditev podatkov v LBS po vsebinskih nivojih

Nivo	Podatki	Primer
I. Leksikalna enota	iztočnica	stisniti
	besedna vrsta	glagol
II. Pomenski nivo	pomenski indikator	(1) <i>oprijeti z dlanjo</i>
	pomenska shema	ČLOVEK stisne PREDMET z dlanjo
III. Skladijski nivo	skladijska (mega)struktura	sbz1 GBZ sbz4 (za sbz4)
	skladijski vzorec (1)	kdo stisne kaj
	skladijski vzorec (2)	kdo stisne koga za kaj
IV. Kolokacijski nivo	(1)	stisniti [roko, vrat]
	(2)	stisniti za [roko, ramena, vrat]
V. Zgled	zgled (1)	<i>Ko je uro pozneje Michaels z landroverjem vozil nazaj h koči, je ob menjalniku otipal ženino roko in jo močno stisnil.</i>
	zgled (2)	<i>Stisnil jo je za vrat in jo začel daviti ter jo skušal posiliti, vendar mu zaradi motenj s potenco to ni uspelo.</i>



VI. Frazeologija	FE	srce (se) stisne komu
	pomenski indikator	razžalostiti, postati žalosten
	zgled	<i>S kratkim pozdravom se poslovim, srce pa mi stisne boleča otožnost.</i>

Tabela 3: Podrobna razčlenitev podatkov po nivojih: glagol *stisniti* in samostalnik *spomin*

NIVO	Jezikovni podatki		Primeri	
I. Leksikalna enota	iztočnica		stisniti	spomin
	zapis		stisniti	spomin
	besedna vrsta		glagol	samostalnik
II. Pomenski nivo	pomenski indikator		(1) <i>oprijeti z dlanjo</i>	(1) pomnjenje česa iz preteklosti
			(2) <i>iztisniti tekočino</i>	(1.2) pomnjenje in dejstva o pomembnih ljudeh ali dogodkih iz preteklosti
	pomenska shema		(1) ČLOVEK stisne PREDMET z DLANJO	
			(2) ČLOVEK s PRIPOMOČKOM stisne SADJE ali stisne TEKOČINO iz SADJA ali stisne, SADJE v TEKOČINO	
III. Skladenjski nivo				
	skladenjske skupine			
	skladenjska struktura (1/a)		sbz1 GBZ sbz4	Prid-SAM1
	vzorec (1/a)		kdo stisne kaj	



IV. Kolokacijski nivo	kolokacija		stisniti [dlan, vrat, roko]	[lep, nostalgичen, prijeten, nepozaben, mladosten] spomin
V. zgledi	zglede		Ljubeče je stisnil njeno drobno dlan	Na Ljubljano me vežejo prelepi mladostni spomini.
	skladenjska struktura (1/b)		*sbz1 GBZ sbz4 z/s sbz6	SBZ1-Vez-gbz-Prid
	vzorec (1/b)		kdo stisne kaj s čim	
IV. Kolokacijski nivo	kolokacija		stisniti s/z [prsti, roko, dlanjo]	spomin je [boleč, živ]
V. zgledi		zgledi	Kamen se je tesno prilegal dlani , tako da ga je lahko močno stisnil s prsti.	Tih in boleč je spomin na 28. junij , ko sta minili dve leti , odkar nas je zapustil naš dragi mož , oče in dedek.
	skladenjska struktura (1/c)		sbz1 GBZ sbz4 za sbz4	Glag-SBZ4
	vzorec (1/c)		kdo stisne koga za kaj	
IV. Kolokacijski nivo	kolokacija		stisniti za [ramo/ramena, roko, prste, vrat, boke]	[obujati, obuditi, zbudjati, priklicati, prebuditi] spomine
V. zgledi		zgledi	Mirno bi ga lahko stisnil za goltanec ali ga zadušil z blazino	Slika rojstne hiše priklіče nepozabne otroške spomine.
	Skladenjske zveze			
	zveza (1)		**stisniti ustnice v [tanko, ozko] črto	*v spomin
	vzorec			narediti kaj v spomin komu



	vzorec			narediti kaj v spomin koga
	vzorec			narediti kaj v spomin na koga
IV. Kolokacijski nivo	kolokacija			[turnir, pohod] v spomin [slovesnost, svečanost, praznik] v spomin
V. zgledi		zgledi	Tommy Duggan je stisnil ustnice v tenko, jezno črto.	Knjiga je napisana v spomin Nejcju Zaplotniku, vrhunskemu alpinistu.
	zveza (2)			ostati v spominu
IV. Kolokacijski nivo	kolokacija			ostati v [lepem, trajnem] spominu
V. zgledi		zgledi		Hvala vsem, ki ste ga ohranili v lepem spominu.
	zveza (3)			ostati zapisan v spominu
IV. Kolokacijski nivo	kolokacija			Za vedno boš ostal zapisan v našem spominu
	SBZ (1)			kratkoročni spomin
	indikator			medicina
	struktura			(L)ema: kratkoročni (L)ema: spomin
IV. Kolokacijski nivo	kolokacija			
V. zgledi		zgledi		Z leti upade sposobnost učenja in predstave, kratkoročen spomin opeša



VI. Frazeologija				
	enota		stisniti zobe	Če me spomin ne vara
	indikator		<i>potrpeti</i>	<i>izraža omilitev, dvom v trditev</i>
V. zgledi		zgledi	Kljub hudim bolečinam sem stisnil zobe in si privozil pomembne točke v boju za naslov svetovnega prvaka	
	enota		stisniti rep med noge	
	indikator		ustrašiti se in se umakniti	
V. zgledi		zgledi	Čim se pojavi nasprotnik z argumenti, stisnejo rep med noge ali pa pred poslušalci in gledalci preprosto pogorijo.	

* Ker prvi pomen samostalnika, ki ga navajamo za zgled, ne izkazuje skladijskih zvez, navajamo skladijske zveze, ki sodijo pod 2. pomen samostalnika, tj. 'kar se ljudje spominjajo o pomembnem človeku ali dogodku'.

** Glagol v obravnavnem pomenu ne izkazuje skladijskih zvez, zato navajamo zvezo za pomen 'približati enake dele telesa'.

1.3.2.2 Zgradba gesla v programskem vmesniku DPS (*Entry Editor*)

Ker je zgradba gesla v programskem vmesniku (DPS vmesnik)⁸ – t. i. DTD struktura – neposredno povezana z vsebino LBS, si bomo najprej ogledali zgradbo dokumenta, kot je razvidna iz programskega vmesnika, v katerem avtorji izdelujejo posamezna gesla, in sicer glede na vsebino predvidenih leksikalnih informacij in glede na hierarhična razmerja med elementi DTD-ja. Ob

⁸ Programska oprema za izdelavo leksikalne baze je podrobno predstavljena v samostojnem dokumentu: Opis analize referenčnega korpusa (Kazalnik 5).



posameznih elementih bomo navedli tudi njihova poimenovanja, ki jih bomo uporabljali v specifikacijah, zlasti v istem delu, kjer pojasnjujemo standarde za izdelavo iztočnic v LBS.

Gesla so v programskem vmesniku razdeljena na več zaključenih enot, ki so bodisi med seboj v hierarhičnem razmerju bodisi na enaki ravni znotraj celotne hierarhije. Prvi del DTD-ja, ki je hierarhično najvišji in tako nadrejen vsem drugim elementom, vsebuje podatke o delih leksikalne baze na sploh, ki jih imenujemo *članki*. Vsak članek znotraj leksikalne baze ima kompleksno zgradbo, ki predvideva dva obvezna elementa: *glavo* in *geslo*.

1.3.2.2.1 Glava

Glava vsebuje podatke o obliki LE – nadelement *oblika*, znotraj katere ločimo *zapis* iztočnice predvsem iz tehničnih razlogov, in zapis iztočnice same – tj. element *iztočnica*. Element *zapis* je namenjen lažjemu iskanju posameznih iztočnic, npr. v primeru, da bi na tem mestu vključili tudi zapis naglasnega mesta in vrsto naglasa, saj iskalnika ni potrebno obremenjevati s poznavanjem diakritičnih znamenj. Sledi nadelement *zaglavje*, kjer je predviden zapis podatka o besedni vrsti obravnavane iztočnice – element *besvrs*, ter povezave z leksikonom⁹, kjer so leksikonske enote morfosintaktično označene (sezname oznak temeljijo na projektu Multext-East in JOS – jezikoslovno označevanje slovenščine).¹⁰ Možne so seveda tudi druge povezave in informacije v zvezi s pravopisom, slovnico itd. obravnavane besede v iztočnici.

Slika 2: Element *glava* v programskem vmesniku

⁹ Leksikon besednih oblik bo prinašal okvirno 100.000 leksikonskih enot, vsebujočih oblikoslovne paradigme posameznih besed z naborom informacij, prekrivnih s sistemom oblikoskladenjskega označevanja JOS. Predvidena je vključitev informacij o (besedotvorni) povezanosti leksikonskih enot, v primeru v jezikovni rabi izpričane oblikovne variantnosti pa bodo vključeni tudi podatki o pogostnosti oblik ter njihovi trenutni opredeljenosti v normativnih virih. Vključitev večbesednih enot je predvidena na ravni večbesednih lastnih imen ter oblik, ki se variantno pišejo skupaj oz. narazen.

¹⁰ Projektna dokumentacija ter rezultati projekta so na voljo na strani: <http://nl.ijs.si/jos/>.



glava
 oblika
 zapis *besedilo* /zapis
 iztočnica *besedilo* /iztočnica
 /oblika
 zaglavje
 besedna vrsta *besedilo* /besedna vrsta
 /zaglavje
 /glava

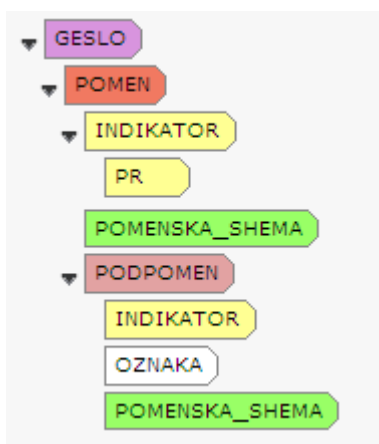
1.3.2.2.2 Geslo

Jedrni del LB, kjer so shranjeni vsi predvideni podatki o posamezni LE, predstavljajo *gesla*. Opis LE znotraj DTD strukture se dejansko začne na ravni pomena, ki je nadrejen vsem drugim informacijam o iztočnici.

1.3.2.2.2.1 Pomeni in podpomeni

Vsi podatki, ki jih v zvezi z obravnavano besedo v iztočnici navajamo znotraj gesla, se podrejajo pomenski členitvi besede v iztočnici. Vsako geslo tako predvideva obvezni nadelement *pomen*, ki mu sledita obvezna elementa *pomenski indikator* in *pomenska shema* ter poljubno število *podpomenov*, ki niso obvezni element geselske strukture. V primeru, da ima pomen enega ali več podpomenov, mu obvezno sledita elementa *indikator* in *pomenska shema*. Znotraj elementa *pomenski indikator* je za navajanje pragmatičnih elementov predviden element *pr* (pragmatika) ter element *oznaka* za navajanje slovničnih in pomenskih restrikcij, ki se nanašajo na celotni pomen besede v iztočnici.

Slika 3: Elementa *pomen* in *podpomen* v programskem vmesniku



geslo
 pomen
 indikator *besedilo* /indikator
 pr *besedilo* /pr



pomenska shema *besedilo* /pomenska shema

podpomen

indikator *besedilo* /indikator

oznaka *besedilo* /oznaka

pomenska shema *besedilo* /pomenska shema

/podpomen

/pomen

/geslo

1.3.2.2.1.1 Skladenjske skupine in skladske zveze

Na pomensko členitev besede v iztočnici so vezane skladske informacije, kar pomeni, da opišemo skladsko obnašanje besede za vsak posamezni pomen in podpomen. V ta namen sta v zapisu DTD-ja predvidena dva nadelementa: *skladske skupine* in *skladske zveze*.

Nadelement *skladske skupine* je predviden za opis skladskih struktur pri vseh obravnavanih besednih vrstah, pri čemer se glagolska gesla v tem delu nekoliko razlikujejo od drugih besednih vrst: skladsko obnašanje glagolov opišemo tako z združeno skladsko strukturo – element *skladska struktura*,¹¹ kot tudi s skladskimi *vzorci*. Vsako skladsko strukturo (in vzorce pri glagolih) potrdimo s kolokabilnimi deli struktur ali vzorcev v elementu *kolokacija* in z zgledi v elementu *zglede* (nadelement *zglede*). Na istem nivoju kot skladske skupine so pri samostalniških, glagolskih, pridevniških in prislovnih geslih predvidene tudi skladske zveze (nadelement *skladske zveze*), ki predstavljajo bodisi kolokacijsko razširjene bodisi ustaljene pogoste dele jezika. Posamezne skladske zveze navajamo v elementu *zveza* zaporedoma, ki jim po potrebi sledijo *vzorci*. Skladske zveze potrjujemo s kolokacijami in zgledi.

Slika 4: Nadelementa *skladske skupine* in *skladske zveze* v programskem vmesniku

¹¹ Za razliko od skladskih struktur pri drugih besednih vrstah, govorimo pri glagolih o *megastrukturah*.



pomen

indikator *besedilo* /indikator

pomenska shema *besedilo* /pomenska shema

skladenjske skupine

skladenjska struktura

struktura *besedilo* /struktura

vzorec *besedilo* /vzorec

vzorec *besedilo* /vzorec

kolokacija *besedilo*/ kolokacija

kolokacija *besedilo*/ kolokacija

zgledi

zgleđ *besedilo* /zgleđ

zgleđ *besedilo* /zgleđ

/zgleđi

/skladenjska struktura

/skladenjske skupine

skladenjske zveze

skladenjska zveza

zveza *besedilo* /zveza

zveza *besedilo* /zveza

vzorec *besedilo* /vzorec

vzorec *besedilo* /vzorec

kolokacija *besedilo*/ kolokacija

kolokacija *besedilo*/ kolokacija

zgleđi

zgleđ *besedilo* /zgleđ

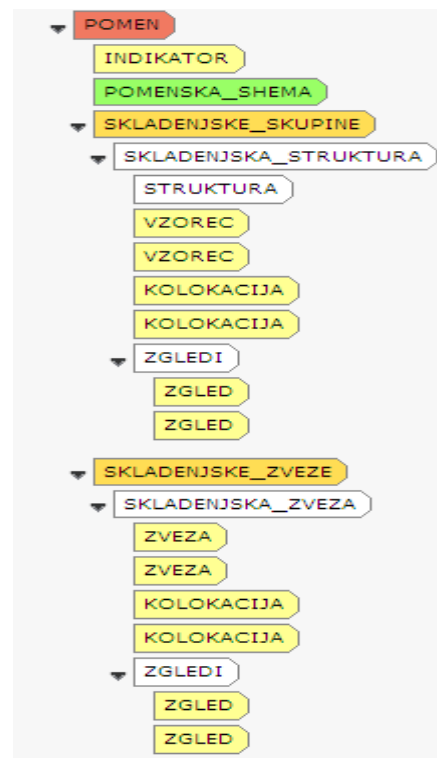
zgleđ *besedilo* /zgleđ

/zgleđi

/skladenjska zveza

/skladenjske zveze

/pomen

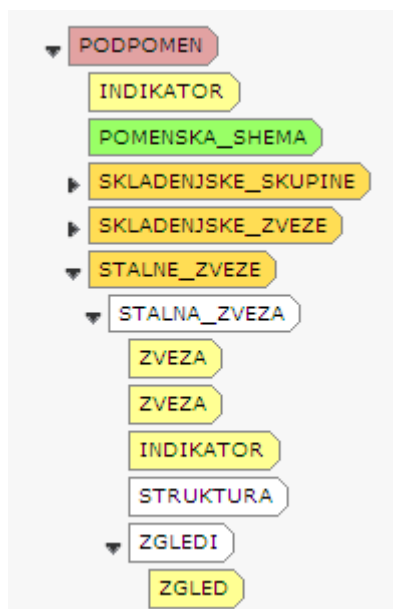


1.3.2.2.1.2 Stalne besedne zveze

Znotraj pomenskega nadelementa v nadelementu *stalne zveze* beležimo pri posameznih iztočnicah stalne besedne zveze kot večbesedne leksikalne enote, ki potrebujejo razlago, hkrati pa vzpostavljajo razmerje s katerim od registriranih pomenov ali podpomenov. V strukturi geselskega članka jih zato vključujemo pod posamezne pomene (na koncu vseh pripadajočih

podpomenov) ali pa znotraj posameznega podpomena. Pomensko povezane stalne besedne zveze (npr. variantne oblike) si znotraj nadelementa lahko sledijo zaporedoma v elementu *zveza*. Ker gre za samostojne leksikalne enote, struktura DTD predvideva še elemente *indikator*, *skladenjska struktura*, *kolokacija* in *zgledi*.

Slika 5: Nadelement *stalne zveze* v programskem vmesniku

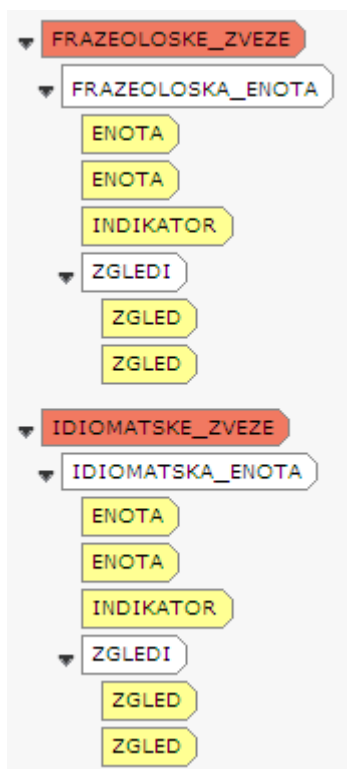


podpomen
 indikator *besedilo* /indikator
 pomenska shema *besedilo* /pomenska shema
 skladsenjske skupine /skladsenjske skupine
 skladsenjske zveze /skladsenjske zveze
stalne zveze
 stalna zveza
 zveza *besedilo* /zveza
 zveza *besedilo* /zveza
 indikator *besedilo* /indikator
 struktura *besedilo* /struktura
 zgledi
 zglede *besedilo* /zglede
 /stalna zveza
/stalne zveze
 /podpomen

1.3.2.2.2 Frazeološke in idiomatske enote

Kot popolnoma samostojne LE beležimo v samostojnih razdelkih LBS (tj. ne znotraj pomenskega razdelka) frazeološke in idiomatske enote, ki si, kadar so pomensko povezane (tj. variante in pretvorbe), lahko sledijo zaporedoma v elementu *enota* znotraj nadelementov *frazeološke zveze* ali *idiomatske zveze*. Kot za druge LE predvidevamo tudi zanje pomenski indikator, beleženje kolokabilnega okolja in možnost navajanja različnih tipičnih realizacijskih možnosti.

Slika 6: Nadelementa *frazeološke* in *idiomatske zveze* v programskem vmesniku



frazeološke zveze

frazeološka enota

enota *besedilo* /enota

enota *besedilo* /enota

indikator *besedilo* /indikator

zgledi

zgled *besedilo* /zgled

zgled *besedilo* /zgled

/frazeološka enota

/frazeološke zveze

idiomatske zveze

idiomatska enota

enota *besedilo* /enota

enota *besedilo* /enota

indikator *besedilo* /indikator

zgledi

zgled *besedilo* /zgled

zgled *besedilo* /zgled

/idiomatska enota

/idiomatske zveze

1.3.3 Vsebina

LBS predstavlja skladišče podatkov o LE, vsebina teh podatkov pa izhaja iz prepričanja o neodtujljivi povezanosti pomenskih in skladiščenjskih lastnosti besed, hkrati pa sledi zadovoljevanju potreb jezikovne skupnosti in njenega življenjsko relevantnega interesa na področju jezika. Zadnje določa vsebino LBS v tem, da mora vsebovati spekter podatkov, ki bodo v najširšem smislu večkrat uporabljivi v jezikovnopriročniških in računalniških aplikacijah različnih vrst. Zato se zdi pomembno na eni strani upoštevati sodobno stanje (tako v teoriji kot praksi) na področju leksikologije, leksikografije, jezikovnih tehnologij in potreb uporabnika v širšem jezikovnem okviru (zlasti v evropskem kontekstu), na drugi strani pa izhajati iz obstoječega stanja in potreb glede slovarskopriročniške infrastrukture za slovenščino. V nadaljevanju zato pojasnimo nekaj temeljnih teoretičnih premislekov, ki smo jih upoštevali pri snovanju LBS, in predstavimo modele, ki so nam služili kot inspiracija pri oblikovanju vsebine leksikalnih in slovničnih podatkov, vključenih v LBS.



Pri tem je potrebno že v izhodišču (ponovno) poudariti, da se v LBS odrekamo kakršnimkoli interpretacijam, ki so neposredno povezane z organizacijo podatkov, ki zadeva rešitve bodisi v konkretnem slovarju (eno- ali večjezičnem, specializiranem itd.) bodisi v katerem drugem jezikovnem priročniku (slovnici, slogovnem priročniku, jezikovnem priročniku za tujce itd.). S tem razlogom LBS ne vsebuje definicij oz. pomenskih razlag, meril za razločevanje homonimije in večpomenskosti, opredelitev do različne stopnje frazeološke pomenske (ne)prozornosti, kategorizacije LE glede vrstnosti, kakovostnosti, terminološkosti ipd. Iz istih razlogov se tudi ne odločamo za hierarhično podrejanje besednodružinsko povezanih iztočnic. Na ta način ne izgubljam o podatkov o pomenu pri podrejenih iztočnicah (npr. pridevnik – prislov; samostalnik – glagolnik itd.) in se izognemo opredeljevanju sistemskih hierarhičnih razmerij, ki pogosto sledijo sistemskim možnostim jezika,¹² nimajo pa potrditev v realni rabi, so obvestilno prazna oz. zahtevajo jezikoslovno nadpovprečnega uporabnika. Na drugi strani težimo k temu, da bo LBS vključevala vse temeljne podatke, ki so za zgoraj omenjene aplikacije bistvenega pomena: sem sodi zlasti evidentiranje in podrobna pomenska členitev jedrne sodobne slovenske leksike, beleženje leksikalnogramatičnih lastnosti LE in njihova formalizacija v računalniško berljivem jeziku, beleženje kolokabilnega potenciala in variantnosti LE, registracija jedrne večbesedne leksike in frazeologije.

1.3.3.1 Beleženje pomena: pomen ali pomenske tendence?

Večkrat smo že poudarili, da je primarni namen izrabe LBS izdelava slovarjev in večpredstavnih jezikovnih priročnikov. To, s čimer se slovarji tipično ukvarjajo, pa je pojasnjevanje besed v njihovih različnih pomenih. Ob tem ne moremo mimo – na današnji stopnji leksikografske teorije in prakse že uveljavljenega spoznanja (prim. Kilgarriff 2006a: 29, Atkins in Rundell 2008: 264) –, da namreč ni jasno, kaj besedni pomeni dejansko so in celo, če kaj takega, kot je besedni pomen, v

¹² Gre za eno bistvenih ugotovitev korpusnega jezikoslovja, da se namreč sistemske možnosti jezika ne uresničujejo sistemsko.



resnici sploh obstaja.¹³ Bistvo te na prvi pogled kontroverzne ugotovitve je, da večina pomenske informacije konkretne LE izhaja iz njenega bodisi neposrednega bodisi širšega besedilnega okolja, iz česar sledi, da vsega, kar se nahaja vsakokrat v sobesedilu vsake konkretne LE, ni mogoče zajeti v njen pomenski opis, kot si ga zamišlja oz. omogoča slovar, in, da govorci v medsebojnem sporazumevanju in dekodiranju pomenske vrednosti LE počnemo nekaj povsem drugega, kot počnejo leksikografi, ko opisujejo pomene besed v slovarjih in jim pripisujejo različne slovnične in pomenske kategorije. Obstaja torej bistvena razlika med slovarskimi pomeni (tj. oštevilčenimi pomeni, na katere so navadno razdeljene iztočnice v slovarjih) in pomeni, kot jih dojemajo slovarski uporabniki (Atkins in Rundell 2008: 263). Dejstva, da sta postopek pomenskega dekodiranja v vsakdanji jezikovni komunikaciji in postopek pomenskega razdvoumljanja pomenov v slovarju (ang. WSD – Word Sense Disambiguation) med seboj neprimerljiva, bi se, kot pravita Atkinsonova in Rundell morali leksikografi pri svojem delu jasno zavedati. Logična posledica tega je, da v LBS ne težimo k beleženju izoliranih besednih pomenov, pač pa beležimo najočitnejše pomenske tendence. Vendar pa, tako kot je na eni strani jasno, da ne obstaja splošno sprejeti in najboljši način pomenskega členjenja večpomenskih LE, obstaja na drugi strani tudi težnja, da vsi člani leksikografske skupine razumejo in uporabljajo iste principe pomenskega členjenja, s čimer se ustvarja potrebna konsistenca znotraj podatkovne baze. Kot pravita Atkins in Rundell (2008: 275), je največ, kar lahko leksikograf doseže pri pomenskem členjenju LE to, da uporabnik pri pregledovanju večpomenske slovarske enote nima občutka, da gre za nesmiselen in nerazumen način predstavljanja dejstev. Prav to je tudi eno od osnovnih vodil pri pomenskem členjenju leksike v LBS.

Če govorimo o leksikalni bazi, na podlagi katere bo (med drugim) mogoče izdelati slovarje, potem je eden od temeljnih ciljev priti do celotnega inventarja pomenov za vsako LE in jih predstaviti tudi z drugih vidikov, zlasti skladenjskega in kolokabilnega. Podatkovna baza torej ni že sam

¹³ Če citiramo van der Meera 2006: 604: »Kljub stoletjem praktičnega ukvarjanja z leksikografijo, še vedno ne obstaja oprijemljiv konsenz glede delitve semantičnega prostora leksikalne enote«. (Prevod P. G.)



slovar, zaradi česar se tudi odrekamo oblikovanju slovarskih definicij, ampak strukturiran zbir podatkov, ki se oblikujejo v analitičnem procesu iz surovega jezikovnega materiala. Če se vrnemo k pomenskemu členjenju LE, lahko ugotovimo, da posamezni slovarji, celo taki, ki so namenjeni istemu ciljnemu uporabniku in imajo primerljiv obseg, različno členijo posamezne besede na pomene. To ni presenetljivo. Pomeni večpomenskih besed so namreč lahko predstavljeni do različnih stopenj razdrobljenosti. Kako podrobno bo konkretni slovar pomensko členil LE, je odvisno od njegovega namena, obsega, ciljnega uporabnika, specializiranosti itd. Ena izmed pomembnih lastnosti, ki jo izpostavlja sodobna leksikografska literatura v zvezi s tem pa je, da je v fazi oblikovanja podatkovne baze smiselno pomene razdeliti čim bolj podrobno (Atkins in Rundell (2008: 268) – in tega se držimo tudi pri oblikovanju LBS. S podrobno razdeljenimi pomeni v podatkovni bazi je namreč leksikografom na voljo celoten spekter možnosti, iz katerih lahko izpeljejo več različnih slovarjev (večje, manjše, eno- in dvojezične, specializirane itd.). Hkrati sledimo tudi „nasprotnemu“ vodilu: razbijanje pomenov mora slediti še obvladljivim in dovolj homogenim enotam, čemur denimo Čermák (2009) pravi »zadostna analogija«.

1.3.3.2 Opis pomena

Pri pomenskem opisovanju LE želimo v prvi vrsti, čeprav na račun sistematičnosti, ohranjati preprostost in težiti k razumljivosti. Uporabljamo torej pomenske opise v obliki pomenskih indikatorjev ki niso obremenjeni z rigidnostjo jezikoslovnega ali leksikografskega metajezika. Da pa bi kljub temu dosegli ustrezno stopnjo konsistentnosti, smo se pri pomenskem členjenju LE in pri oblikovanju pomenskih indikatorjev naslonili na 3 slovarske prakse: New Oxford Dictionary of English (NODE 1998²), Longman Dictionary of Contemporary English (LDOCE 2003⁴) in Macmillan English Dictionary for Advanced Learners (MED 2007²). Vsi trije so enojezični angleški slovarji, ki se osredotočajo na standardno angleščino ter na to, kar je v jeziku osrednje, pri čemer je NODE namenjen angleškim govorcem (po obsegu večji), ostala dva slovarja pa sta namenjena tujim govorcem za učenje angleščine (ang. EFL/ESL dictionary) in po obsegu relativno primerljiva (100.000 besed). Pri vsakem od slovarjev smo za izhodišče pomenskega členjenja upoštevali lastnosti, ki so najbolj ustrezale naši predstavi o podrobni pomenski členitvi besed ter lastnosti, ki



bi po našem mnenju uporabniku LBS ponudile čim bolj zanesljivo in glede na sodobni jezik realno informacijo na čim bolj uporabniku prijazen način.

1.3.3.2.1 New Oxford Dictionary of English

NODE, ki je namenjen angleškim govorcem, prikazuje angleščino, kot jo izkazuje sodobno gradivo in se osredotoča predvsem na to, kar je osrednje in tipično. Rezultat tega je manjše število pomenov, ki so jasno in jedrnato razloženi. Pomenska členitev besed temelji na naslednjih izhodiščih:

- za vsako besedno vrsto (sledijo si po pogostnosti) je prva definicija namenjena jedrnemu pomenu;
- vsaka iztočnica ima navadno vsaj en osnovni/jedrni pomen (ang. core sense), na katerega je vezanih več podpomenov, ki so z njim logično povezani;
- jedrni pomen¹⁴ predstavlja tipično, osrednjo rabo besede (kot jo izkazuje BNC)¹⁵, kar pomeni, da ga materni govorniki pepoznavajo kot najbolj uveljavljenega in osrednjega;
- hkrati je to najbolj „dobesedni“ pomen (ne pa nujno tudi najstarejši oz. izhodiščni);
- jedrni pomen ni nujno najfrekventnejši, najfrekventnejši so lahko tudi figurativni nasledki jedrnega pomena;
- vsak podpomen je jasno izpostavljen;
- logično razmerje med podpomenom in jedrnim pomenom uporabniku omogoča lažji pregled nad iztočnico ter vpogled v strukturiranost in medsebojno povezanost pomenov.

Zgled 1: Pomenska členitev in razvrstitev pomenov v NODE in LB za iztočnico *spomin*

LBS

1. pomnjenje česa iz preteklosti

NODE

1. the faculty by which the mind stores and

¹⁴ Po Eleanor Rosch (citirano v Atkins in Rundell 2008: 277 in 279) imenovan tudi *prototipski pomen*, ki pojasnjuje vse tipične in pogoste rabe neke besede: ne opredeljuje vseh možnih (so)besedilnih rab, a je hkrati dovolj širok, da ustrezno upošteva variacije na nivoju individualnih jezikovnih situacij.

¹⁵ <http://www.natcorp.ox.ac.uk/>.



- 1.1 človekova doživetja iz preteklosti
- 1.2 pomnjenje in dejstva o pomembnih ljudeh ali dogodkih iz preteklosti
- 1.3 nekaj, kar spominja na preteklost
- 2. sposobnost pomnjenja
- 2.1 del mišljenja, ki je namenjen pomnjenju
- 3. avtobiografsko literarno delo
- 4. o računalniku
- 4.1 del računalnika
- 4.2 zmogljivost računalnika




remembers information

- 1.1 the mind regarded as a store of things remembered
- 2. something remembered from the past
- 2.1 the remembering or commemoration of a dead person
- 2.2 the length of time over which a person or event continues to be remembered
- 3. the part of a computer in which data or program instructions can be stored for retrieval
- 3.1 capacity for storing information in this way

1.3.3.2.2 Macmillan English Dictionary

Pri MED smo se zgedovali po principu, po katerem je pri večpomenskih besedah (5 ali več pomenov) na začetku izpostavljen t. i. „meni pomenov“, ki predstavlja povzetek pomenov in podpomenov, ki so podrobneje predstavljeni v nadaljevanju gesla, in sicer tako, da so v meni pritegnjeni najpomembnejši deli razlage, navadno neposredne nadpomenke.

Slika 7: Meni pomenov večpomenske LE *time* (čas) v MED

time /  taim  / noun  ★★ ★

Menu

- 1 quantity: using clock
- 2 period
- 3 occasion/moment
- 4 time available/needed
- 5 how fast music is played
- + phrases



Tak meni pomenov predstavlja izhodišče za oblikovanje indikatorjev, kot si jih zamišljamo v LBS. V nadaljevanju gesla so posamezni pomeni znotraj MED natančneje pojasnjeni, kot prikazuje zgled.¹⁶

Zgled 2: Primerjava definicij v MED in pomenskih indikatorjev v LBS za iztočnico čas

MED – definicije

1. the quantity that you measure using a clock
- 1a a particular moment during a day, measured on a clock
- 1b the time in a particular part of the world
2. a particular amount of time
- 2a period in history
- 2b period in someone's life
- 3 an occasion
- 3a an experience
- 3b a moment
- 3c a particular point when something happens
- 3d a particular moment that is suitable or not suitable for something
- 3e used for saying how often something happens
- 4 the time that is available for something
- 4a the amount of time that you need for particular activity
- 4b the amount of time that someone takes to finish a race
- 5 the speed at which the piece of music is played, measured as the number of beats in each

LB – indikatorji

1. dimenzija obstoja
- 1.1 (v športu) merjen z uro
- 1.2 časovni pas
- 1.3 pri fotografiji
2. omejeno trajanje, obdobje
- 2.1 z opredeljenim namenom
- 2.2 o delu, zaposlitvi
- 2.3 (zgodovinsko) obdobje s svojimi značilnostmi
- 3 trenutek, priložnost
- 3.1 ura

1.3.3.2.3 Longman Dictionary of Contemporary English

LDOCE je namenjen zahtevnejšim študentom/učencem angleščine kot tujega jezika. Organizacija pomenov znotraj večpomenskih besed je za nas zanimiva zaradi t. i. pomenskih kazalcev (ang. signpost), ki so znotraj iztočnice posebej izpostavljeni (velike črke, barvno ozadje). Njihova naloga

¹⁶ Primerjava pomenske delitve, kot je zamišljena v LBS, in med slovarji, namenjenimi učenju tujega jezika, je sicer precej kompleksna: v EFL slovarjih je načeloma pomenov več, več je tudi osamosvojenih pomenov, čeprav obstajajo zgolj kot frazeološke enote.



je nakazati pomensko območje posameznega pomena, za razliko od pomenskega menija v MED v LDOCE kazalci ali indikatorji navadno niso del razširjene definicije, ki sledi v nadaljevanju gesla.

Zgled 3: Pomenski indikatorji jedrnih pomenov za iztočnico *spomin* v LDOCE in LBS

LDOCE: *memory*

1 ability to remember

2 something you remember

3 computer

LBS: *spomin*

1 pomnjenje česa iz preteklosti

2 sposobnost pomnjenja

3 nav. množina avtobiografsko literarno delo

4 o računalniku

1.3.3.2.4 Pomenska shema in teorija FrameNet

Opis pomenske sheme ali udeleženske strukture (ang. argument structure) za vsak posamezni pomen in/ali podpomen glagola predstavlja združitev skladijskih in pomenskih informacij pri opisu glagolskih LE. Pri oblikovanju pomenske sheme smo se v nekaterih segmentih približali teoriji pomenskih okvirjev v projektu FrameNet,¹⁷ katerega bistvo je, da pomaga leksikografom identificirati pomembna in uporabna dejstva na podlagi korpusnih besedil. Pomenska shema, kot jo vidimo znotraj LBS, ohranja glede na projekt FrameNet tudi nekatere pomembne razlike.

Projekt FrameNet, ki je za profesionalno leksikografijo izredno pomemben, nastaja pod vodstvom Charlesa J. Fillmoreja na Mednarodnem inštitutu za računalništvo na Kalifornijski univerzi v Berkeleyju.¹⁸ Namen projekta je za vsak pomen besede ali besedne zveze analizirati in beležiti celotni spekter pomenskih in skladijskih razmerij. Da bi to dosegli, so raziskovalci izoblikovali

¹⁷ Teorija in praksa projekta FrameNet je podrobno predstavljena v posebni številki International Journal of Lexicography, Vol. 16, No. 3, september 2003, Special Issue: Framenet and Frame Semantics. Projekt FrameNet je bil z vidika slovenščine že predstavljen (Krek 2008; tudi kot [predavanje](#) v okviru Solomonovih seminarjev na IJS: [ppt-predstavitev](#)). Birte Lönneker Rodman se ukvarja prav tako s FrameNetom za slovenščino, več o tem pa je mogoče najti na njeni [spletni strani](#).

¹⁸ ICSI: International Computer Science Institute; University of California, Berkeley.



niz oznak, ki poimenujejo pomenske vloge (shemske elemente) in omogočajo podrobno dokumentacijo korpusnega sobesedila, v katerem se nahaja obravnavana beseda oz. njen pomen. Delo pri projektu, znotraj katerega poteka označevanje vzorčnih stavkov polavtomatično, je v celoti računalniško podprto. Nastajajoča podatkovna baza omogoča vrsto različnih avtomatičnih prikazov in iskanj podatkov, ki so namenjeni podrobnim človeškim analizam. Podatkovna baza je prosto dostopna, njen obseg pa se nenehno širi: trenutno jo uporablja na stotine raziskovalcev, učiteljev in študentov jezika po vsem svetu. Podobne analize, ki so tesno povezane z izhodiščnim projektom v Berkeleyju, nastajajo tudi za nemščino, španščino in japonščino. Kot pravita Atkins in Rundell (2008: 145), je framenetovski pristop, tako kot je zasnovan, najbolj primeren in uporaben za raziskovanje obnašanja besed na podlagi korpusnih podatkov, ker omogoča dosledno analizo besedilnega okolja, hkrati pa ne spregleda nobenega ključnega dejstva.

Vsaka beseda, ki je vključena v FrameNet, pripada določenemu pomenskemu okvirju ali shemi. Če je beseda večpomenska, pripada več različnim pomenskim okvirjem. Pomensko povezane besede, npr. glagoli premikanja, govorjenja, zaznavanja itd., se tako pojavljajo znotraj istih pomenskih okvirjev. Vsaka pomenski okvirje vnaprej določen, ima svoje ime in definicijo, v kateri so izraženi jedrni shemski elementi. Poleg nejedrnih lahko okvir vključuje tudi nejedrne elemente. Znotraj pomenskega okvirja ima vsak element svoje ime, opis in je opredeljen glede na semantični tip. Jedrni elementi so za evociranje konkretnega pomena glagola (ali samostalnika) obvezni, vendar v nekaterih natančno določenih primerih lahko tudi neizraženi oz. izpuščeni. Npr. glagol *squeeze* (*stisniti*) pripada dvema pomenskima okvirjema: *Cause_to_move_in_place* (povzročiti premik v prostor(u)) in *Manipulation* (upravljanje). Poleg glagola *squeeze* se znotraj prvega okvirja med drugim nahajajo tudi glagoli kot: *rotate*, *seesaw*, *shake*, *spin*, *swing*, *swirl*, *turn* ipd., v drugem okvirju pa glagoli kot: *caress*, *clasp*, *claw*, *clutch*, *diddle* ipd. Spodnja slika prikazuje opis glagola *squeeze* v pomenskem okvirju *Manipulation*, kamor sodijo vse besede (oz. njihovi pomeni), ki opisujejo *upravljanje z neko entiteto*, ki jo izvede *aktant*, navadno z *delom svojega telesa*. Iz tega izhaja, da so *entiteta*, *aktant* in *del njegovega telesa* jedrni elementi pomenskega okvirja *Manipulation*.

Slika 8: Opis jedrnih elementov pomenskega okvirja *Manipulation* za glagol *squeeze* v FrameNetu

Manipulation

Definition:

The words in this frame describe the manipulation of an **Entity** by an **Agent**.

Ian **SQUEEZED** Mara's hand.

A number of lexemes listed in this frame also occur in the frame Cause-to-move (for example, *push*, *yank*). Only the non-motion uses of these words belong in this frame, e.g., combinations such as *push at*, *push on*, *yank at*.

FES:

Core:

Agent [Agt]

Semantic Type

The Agent manipulates an object. Agent typically refers to animate beings, but is also extended to mechanical and other inanimate objects. Agent is most frequently expressed as the external argument of verbs: **A robot arm** **GRASPS** the box on the conveyor belt and seals it.

Bodypart of agent [AgtBP]

Semantic Type

Body_part

The part of the Agent's body being used to manipulate the **Entity** may also be expressed. While the default is for an **Agent** to use his/her hands, other body parts may be specified. This frame element usually occurs as a PP-complement.

Kim **GRASPED** the needle **between finger and thumb**.

Entity [Enty]

The Entity being manipulated often occurs as the direct object of verbs: **SQUEEZE** **the ball** as many times as you can. Pat **SQUEEZED** **Jo's hand**.

Tako kot pri projektu FrameNet temelji tudi v LBS opis pomenskega scenarija v pomenski shemi konkretne LE na analizi velikega števila primerov oz. stavkov iz korpusa, ki vsebujejo obravnavano LE. Za razliko od FrameNeta pa naš namen ni opisati pomena določene besede na podlagi obstoječega ali vnaprej izdelanega seznama pomenskih okvirjev in pomenskih elementov znotraj njih. V LBS je opis pomenskega scenarija v pomenski shemi prost; edini pogoj je, da redaktor natančno opiše situacijo pomena določene LE. Pri opisu pomenskega scenarija v pomenski shemi tako kot pri FrameNetu jasno izpostavimo za vsak pomen obvezne udeležence in okoliščine ter vlogo, ki jo pri tem opravljajo. Za razliko od FrameNeta, ki za vsak shemski element predvideva ime, opis in semantično kategorijo, poimenovanje udeležencev in okoliščin v pomenskih shemah v LBS ni fiksirano in vnaprej določeno, pač pa je poimenovanje udeležencev, npr. ČLOVEK, PREDMET, ŽIVAL, SNOV, DEL TELESA itd. podrejeno opisu konkretnega



pomenskega scenarija besede oz. njenega pomena. Kljub temu težimo pri opisu pomenskega scenarija tudi v LBS k natančnosti in navajanju tistih podrobnosti, ki omogočajo razločevanje med posameznimi pomeni oz. podpomeni istega glagola, medtem ko bo abstrakcijo in posplošitev v smislu framenetovskih shemskih elementov mogoče izpeljati naknadno, ko bo izdelanih več pomenskih shem. Za razliko od FrameNeta, kjer formalizirani pomenski okvirji in elementi znotraj njih omogočajo medsebojno povezovanje in vzpostavljanje pomenskih razmerij med LE, namen LBS ni združevati LE na podlagi pomenskih shem, čeprav naknadna analiza omogoča tudi to. Menimo namreč, da je postopek generaliziranja v času oblikovanja LBS manj primeren, saj bi s posploševanjem lahko izgubili marsikatero pomembno pomenskorazločevalno informacijo. Bistvo FrameNeta, ki se kaže v povezovanju pomensko sorodnih besed, narekuje obravnavo LE glede na njihovo pomensko povezanost. Konkretno to pomeni, da v primeru, če framenetov seznam ne vsebuje pomenskega okvirja, ki bi ustrezal opisu določene nove besede, te besede ni mogoče vključiti v obdelavo, dokler se ne pokaže zadostna potreba oz. zadostno število besed, ki bi ustrezale novemu pomenskemu okvirju. V nasprotju s takim sistemom je LE v LBS mogoče obravnavati neodvisno od njihove pomenske (ne)povezanosti, kar omogoča, da v LBS vključimo jedrni del besedišča neodvisno od možnosti pomenskega povezovanja med LE. Iz istega razloga lahko v LBS več pozornosti posvetimo pomenskemu opisu večpomenskih besed, tj. opisu tudi tistih pomenov, ki nimajo v celotni leksikalni bazi izkazane sorodne pomenske povezave.

Glede na opisane skupne točke in razlike med projektom FrameNet in rešitvami v LBS je za nas zanimiva tudi Hanksova (2008)¹⁹ nekoliko kritična primerjava med CPA (ang. Coprpus Patterns Analysis, Hanks 2004, 2008; Hanks in Pustejovsky 2005) in FrameNetom. Kot pravi Hanks, skuša CPA za razliko od FrameNeta določiti sintagmatska merila za razlikovanje pomenov večpomenskih besed na t. i. semantičnopovršinski način (ang. »semantically shallow« way), medtem ko FrameNet izraža globinsko semantiko (ang. deep semantics) situacij, izraženih v pomenskih okvirjih. FrameNet leksikalne enote analizira glede na obstoječe okvire in ne glede na

¹⁹ Ppt-predstavitev predavanja je mogoče najti na spletni strani: <http://nlp.fi.muni.cz/projekty/cpa/>.



(abecedni) seznam besed; situacije v pomenskih okvirjih so v FrameNetu analizirane z vidika shemskih elementov; FrameNet preučuje pomenske razlike in povezave med različnimi besedami v okvirju, ne ukvarja pa se eksplicitno s pomenskimi razlikami večpomenskih besed. Po Hanksovem mnenju FrameNet ne analizira korpusnih podatkov sistematično, pač pa v korpusu lovi samo primere, ki potrjujejo hipoteze (izražene v okvirjih). Poleg tega ima FrameNet problem z vključevanjem novih besed v okvirje, pri čemer se marsikatera beseda izgubi.

1.3.3.3 Leksikalno-gramatični pristop: skladienske informacije v podatkovni bazi

Bistvo leksikalnogramatičnega pristopa pri opisu leksikalnega potenciala jezika je v povezovanju pomenskih in skladijskih lastnosti LE. Ta princip je neposredno povezan že s pomensko členitvijo večpomenskih LE, kjer upoštevamo vpliv skladijskega okolja na realizacijo konkretnega pomena (gl. tudi poglavje o oblikovanju pomenskih shem), v LBS pa ga upoštevamo tudi pri beleženju skladijskih struktur in skladijskih vzorcev, ki so specifični za obravnavano LE v vsakem od njenih pomenov in podpomenov. Če predstavlja pomenska shema opis pomenskega scenarija, v katerem so za glagolske iztočnice jasno izraženi udeleženci, njihova razporeditev v skladijskem vzorcu in okoliščine, predstavljajo skladijske strukture in skladijski vzorci beleženje možnih različnih tipičnih skladijskih realizacij konkretnih pomenov preučevane LE. S skladijskimi vzorci izpostavljam predvsem funkcijski nivo preučevanega glagola, na strukturni ravni pa nas zanima zgradba besednih zvez (Toporišič 1982: 39 – stavčnih fraz) na posameznih stavčnih delih in zgradba tipičnih struktur ne glede na jedrnost iztočnice ali stavčnočlensko mejo. **S skladijskimi strukturami in skladijskimi vzorci torej beležimo in formaliziramo različne skladijske realizacije pomensko obveznih (navadno pa tudi tipičnih opcijskih) elementov, izraženih v pomenski shemi.** V LBS se problema lotevamo na način, ki bo računalniško berljiv in zato uporaben pri računalniškem procesiranju jezika.

Izhajajoč iz leksikografske teorije in prakse, nas na skladijskem nivoju LBS zanimajo podrobnosti skladijskega okolja preučevane LE, ki jih je mogoče ugotoviti z analizo konkordančnega niza in s pomočjo orodja za avtomatično profiliranje besed, kot je Sketch Engine, zlasti aplikacija Word



Sketches ali t. i. *besedne skice* (Krek in Kilgarriff 2006). Z ugotavljanjem skladijskih struktur (ang. constructions) in skladijskih vzorcev (ang. syntactic patterns), v katerih se LE tipično pojavljajo, je mogoče beležiti tako leksikalno in s tem leksikografsko pomembne elemente stavka, v katerem se pojavlja LE, in ki jih mora poznati vsak govorec jezika, če želi ustrezno uporabljati besede in besedne zveze v njem, kot tudi informacije, ki so zanimive za slovnične opise jezika. Zlasti prvi tip informacij je z razmahom korpusnega jezikoslovja dobil pomembno mesto tako v eno- kot v dvojezičnih slovarjih, neizogibna pa je ta informacija v slovarjih, ki so namenjeni učenju tujega jezika (Atkins in Rundell 2008: 331).

V LBS smo z opisom skladijskih struktur želeli slediti dvema ciljema: s pomočjo kod formalizirati strukture, v katerih se tipično pojavlja preučevana LE in to informacijo narediti uporabno tudi za leksikografske opise slovenščine. Pri glagolskih iztočnicah, kjer glagoli s svojo intenco odpirajo vezljivostna mesta, njihovo število, razporeditev, slovnične in pomenske omejitve pa so pomensko relevantni, beležimo poleg struktur tudi skladijske vzorce kot potencialno leksikografsko in slovnično zanimiv podatek. S spodnjimi primeri želimo pokazati, kako se posamezni pomeni izražajo z različnimi skladijskimi strukturami oz. vzorci in obratno, kako posamezne skladijske strukture izražajo različne pomene LE.

ZGLED: povezava med skladijskimi strukturami/vzorci in pomeni glagola *leteti*

struktura/vzorec	pomen glagola <i>leteti</i>	kolokacijski zgled
sbz1 GBZ ▪ kdo/kaj leti	→ <i>'premikati se po zraku'</i>	[ptič, ptica, čebela] leti
	→ <i>'potovati'</i>	[potniki, letalo] letijo
	→ <i>'padati'</i>	[žoge, stvari, kovanci, žlice] letijo
sbz1 GBZ na sbz4 ▪ kdo/kaj leti na kaj	→ <i>'premikati se po zraku'</i>	leteti na [jug]
	→ <i>'padati'</i>	leti na [tla]



- kaj leti na kaj

sbz1 GBZ z/s sbz2

→ **'potovati'**

leteti z [letališča]

- kdo/kaj leti z/s česa

→ **'padati'**

leti z [mize, police]

- kaj leti z/s česa

Pri vseh besednih vrstah smo skladijske strukture razdelili glede na to, ali je jedro skladijske strukture iztočnica ali ne. V prvem primeru torej formaliziramo samostalniške, pridevniške in prislovne besedne zveze ali stavčne fraze (Toporišič 1982: 39 in dalje), v drugem primeru pa tipične strukture, v katerih se iztočnica pojavlja ne glede na svojo jedro vlogo.

Zgled 4: Jedrne in nejedrne strukture ter pomeni pri samostalniških iztočnicah

struktura/vzorec

pomen samostalnika *kos*

kolokacijski zgled

(a) JEDRNE

pbz SBZ1

→ **'del celote'**

[velik, zajeten] **kos**

→ **'ptica'**

[črni, povodni] **kos**

SBZ1 sbz2

→ **'enota'**

kos [pohištva, prtljage]

→ **'del celote'**

kos [torte, potice]

(b) NEJEDRNE

gbz SBZ4

→ **'del celote'**

odrezati, odlomiti [kos]



1.3.3.4 Kolokacije

Konec 50. let 20. stoletja so se nasledniki J. R. Firtha, McIntosh, Halliday, Sinclair in drugi, v nasprotju s strukturalističnimi pristopi, ki so temeljili na (ne)razstavljivosti, komponentni analizi in pomenskih razmerjih med leksikalnimi elementi, posvetili preučevanju besedne povezovalnosti, ki ni bila vezana izključno na vlogo besednega pomena. Pozornost je bila usmerjena v preučevanje distribucije besed v besedilu in predvidljivosti pogostejšega sopojavljanja nekaterih besed z enimi kot z drugimi. S tem so bila dana izhodišča za obravnavanje besednega povezovanja kot statistično predvidljivega in hkrati semantičnega pojava. Bistvo omenjenih sintagmatskih pogledov na jezik je spoznanje, da je razmerje med pomenom in obliko zelo subtilno, saj različni pomeni leksikalne enote pogosto ustrezajo samo izbranim besednim kombinacijam. Prav tako je postalo očitno, da lahko dokazi o besedni povezovalnosti spodbijajo introspekcijo rojenih govorcev in da se večina naravnega jezika pojavlja v vsaj deloma vnaprej oblikovanih frazah, ki jih določajo omejene izbire (ang. single choices), čeprav se sicer na prvi pogled zdi, da jih je mogoče analizirati/razstaviti na posamezne elemente. Pojem kolokacije, ki se je oblikoval na tej podlagi, je tudi v nadaljnjih obravnavah pri posameznih avtorjih izpostavljaj bodisi statistično prepoznavanje karakterističnih vzorcev sopojavitve besed (t. i. formalnoverjetnostni, statistični pristop) oziroma se je skušalo določiti karakteristične vzorce tudi z vidika pomenske odvisnosti znotraj zveze, torej kot pojav, ki je poleg svoje povezovalne trdnosti tudi pomensko določen.

- **Kolokacija kot pogost, značilen vzorec**

Pri preučevanju simultano izbranih značilnih vzorcev besednih sopojavitev znotraj korpusnega okolja, ki ga omogoča statistično izračunavanje besedne povezovalnosti (Kilgarriff 2006), so v analizah pozornost zbudile zveze, ki se v jeziku pojavljajo najpogosteje, obenem pa navadno niso del osrednje slovarske problematike. V celoti je tak statistični pristop izpeljal obsežen trizvezkovni slovar Dictionary of English Collocations (Kjellmer 1994), ki temelji na korpusu Brown²⁰ in pojmuje kolokacije kot

²⁰Uporaba korpusa: <http://www ldc.upenn.edu/cgi-bin/ldc/textcorpus?doc=yes&corpus=BROWN> in dokumentacija: <http://icame.uib.no/brown/bcm.html>.



ponavljajoče sekvence elementov (najpogosteje besed), ki ustrezajo slovničnim pravilom jezika, pri čemer se mora zaporedje elementov v korpusu pojaviti več kot enkrat in mora ustrezati predhodno ugotovljenemu skladijskemu vzorcu. Na podlagi tega so v slovar sprejete številne zveze, ki jih v angleščini določata npr. člena *the* in *a* (samo zveze s *the* zavzemajo 156 strani v slovarju in predstavljajo skupaj 15.000 kolokacij). Slovar je doživel precej nasprotujočih si ocen (prim. Johansson 1998), zlasti v smislu svoje klasične slovarske uporabnosti, medtem ko se je na drugi strani jasno pokazala uporabnost pristopa pri učenju angleščine kot tujega jezika (Johansson 1998; Hoey 2000: 224). Pojmovanje kolokacije kot oblikovno verjetnostnega pojava pa je hkrati usmerilo pozornost tudi na zveze, ki delujejo kot celota tako, da vzpostavljajo razmerje s pomensko obvestilnimi deli neposrednega ali širšega besedilnega okolja (Siepmann 2005: 427–429).

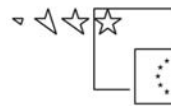
• Kolokacija kot pomenski pojav

Na drugi strani izpostavlja pojmovanje kolokacije kot pomenskega pojava vsaj dve vrsti tipičnih besednih sopojavitvev: take, ki nakazujejo posamezne pomene večpomenske besede kot v primerih (a), in take, ki se tipično sopojavljajo samo pri izbranem/določenem pomenu besede kot v primerih (b). V zvezi z zadnjimi se znotraj frazeološke problematike govori o *omejenih kolokacijah* (Cowie 1981; Aisenstadt 1979; 1981), posebej pa se poudarja tudi vloga pomensko odvisnega oz. neodvisnega dela zveze (Hausmann 1999, 2003; Mel'čuk 1998 idr.).

(a) Kolokatorji nakazujejo posamezne pomene večpomenske besede v jedru²¹ konkordančnega niza

jedro	kolokator	pomen besede v jedru
črn	obleka, avto	črne barve

²¹ Izraz *jedro* je tu vezan izključno na pojmovanje besede ali zveze, ki jo preučujemo kot središče konkordančnega niza (ang. Key word in Context). Čeprav je organizacija znotraj stavčnega vzorca neposredno povezana s kolokabilnim potencialom besede v izhodišču, puščamo za zdaj vlogo skladijskega razmerja in pomenske (NE)jedrnosti, ki jo denimo izpostavlja Hausmann (1999, 2003), ob strani.



črn	gradnja, borza	nezakonit
črn	slutnja, misel	neprijeten, pesimističen
črn	lista, seznam	nezaželen; nedovoljen

(b) Kolokatorji se pojavljajo samo pri izbranem pomenu besede²²

jedro	kolokator	pomen zveze
živo	srebro	kemijski element
živi	zid	ovira, ki jo sestavljajo ljudje
živa	glasba	izvajanje glasbe pred občinstvom
živa	meja	pregrada iz okrasnega grmičevja

Brisanje ostrih razlik med pomensko predvidljivimi in pomensko nepredvidljivimi kolokacijami omogoča razumevanje pojava besedne povezovalnosti v smislu stopnjevalnosti in ne različnih vrst (Bollinger 1976). Temu sledi tudi Sinclairjeva (1991: 71) opredelitev kolokacije kot vmesnega člana znotraj skrajno določenih točk pomena, kjer je na eni strani mogoče ločiti neodvisni pomen besede in odvisni pomen zveze na drugi. Ob tem, da je pri kolokaciji tendenca sopojavljanja besed očitna, gre hkrati za zvezo, kjer lahko posamezna beseda znotraj zveze glede na drugo ohranja tudi svoj neodvisni položaj, npr. beseda *izraz* v zvezi *priti do izraza* in v besednem nizu *to je izraz medsebojne ljubezni*.

1.3.3.5 Večbesedne leksikalne enote

Pojem večbesedna leksikalna enota se v teoretičnem in terminološkem smislu dotika dveh skrajnih točk pojmovanja in ukvarjanja z besednimi kombinacijami: (a) *frazeoloških enot* ali *stalnih besednih zvez* (ang. fixed expressions) kot krovnega izraza za vse tipe frazeoloških enot in (b) *stavčnih fraz* (ang. lexical phrases), poimenovanih tudi (proste) besedne zveze. Besedne zveze (obeh skrajnosti

²² V teh primerih, kot kažejo zgledi, gre frazeološki pristop k pojmovanju kolokacij (prim. Cowie 1981).



in vmesnih tipov) so bile predmet preučevanja že pred obstojem korpusov, ukvarjali pa so se predvsem z njihovo pomensko (ne)transparentnostjo in posledično z ugotavljanjem stopnje leksikaliziranosti. Obstoj in razvoj korpusov in orodij za njihovo analizo je omogočil prepoznavanje pogosto ponavljajočih se jezikovnih sekvenc, ki pa se v izhodišču ni oziralo na stopnjo leksikaliziranosti, pač pa predvsem na tipičnost in pogostost. Preučevanje takih sekvenc na korpusni podlagi je postalo zanimivo z leksikografskega in pedagoškega vidika, zadnje predvsem pri poučevanju in učenju tujega jezika. Korpusno jezikoslovje je namreč jasno pokazalo, da je celoten jezik na nek način vzorčen in da absolutno proste besedne zveze v resnici ne obstajajo (prim. Čermák 1985: 167, 173), zato se je zdelo smiselno preučevanje besednih zvez vključiti tudi v splošnejše in ne le v strogo skladijske ali strogo leksikalne opise jezika (Huston in Francis 1999: 14).

1.3.3.5.1 V razmerju do prostih besednih zvez

Obstoj različnih tipov večbesednih leksikalnih enot, ki segajo od kolokacij, fraznih glagolov, frazeoloških enot do idiomov, in različnih teoretičnih pristopov, ki skušajo fenomen klasificirati in opredeliti z nedvoumnimi in oprijemljivimi merili, jasno kaže, da meje niso strogo začrtane, zlasti ne znotraj obeh skrajnosti, ki ju predstavljajo »proste« besedne zveze (kamor v nekaterih pristopih sodijo tudi kolokacije) na eni strani in idiomatične oz. frazeološke enote na drugi. Če hočemo poiskati objektivno merilo za prepoznavanje večbesednih leksikalnih enot, ki jih želimo vključiti v LBS, jih je smiselno zamejiti glede na vse druge zveze besed. Na podlagi pojmovanja besednih zvez v slovenističnem jezikoslovju (Toporišč 1992: 7) kot »odvisnostno ali logično povezanih besed«, ki jih »v govoru delamo po skladijskih vzorcih, ne da bi jih jemali gotove iz spomina«, npr. *oblačno nebo* po vzorcu pridevnik + samostalnik (Toporišč 1982: 230), so (proste) besedne zveze ali stavčne fraze pojmovane kot gradbene enote višjih in zapletenejših skladijskih zvez. V LBS jih zato registriramo v skladijskem delu, kjer posebej ločujemo strukture, v katerih iztočnica nastopa kot jedro (tj. glagolske, samostalniške, pridevniške in prislovne fraze), in strukture, kjer se iztočnica pojavlja ne glede na jedrnost: z njimi beležimo skladijsko okolje, v katerih se tipično



pojavlja iztočnica kot odvisni del stavčne fraze, vendar je za izražanje konkretnega pomena prav tako zanimiva.

1.3.3.5.2 V razmerju do kolokacij

V jezikoslovni literaturi se v razmerju do (prostih) besednih zvez, katerih primarni kriterij je oblikovanost po slovničnih pravilih jezika, vedno znova postavlja tudi potreba po razmejitvi glede na tiste zveze, ki v jeziku obstajajo kot del njegovega leksikalnega fonda. Na tej podlagi je potrebno najprej razčistiti razmerje med besedami in tipičnimi besednimi sopojavitvami (kolokacijami), ki pa navadno niso naključne, nasprotno, pogosto so tudi pomensko relevantne, kot smo prikazali že v poglavju o kolokacijah. Če pustimo ob strani različne možnosti, ki zadevajo pisanje skupaj in narazen, je problem kompleksen tudi (ali predvsem) s pomenskega vidika: posamezne besede, ki tvorijo v kombinaciji z drugimi besedami razmeroma prepoznavne leksikalne enote, npr. *stara mama*, tj. 'očetova ali mamina mama', ohranjajo v različni meri tudi tiste svoje 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 v pomenih (po SSKJ) 'ki ima razmeroma veliko let' in 'ki je v zadnjem obdobju življenja', kot pomenskih lastnosti, ki jih vsebujejo drugi pomeni istega pridevnika, prav tako izkazani v SSKJ, npr. 'ki nima več potrebne, ustrezne kakovosti za uživanje' (*star kruh*) ali 'ki je dalj časa v uporabi' (*star avtomobil*). To kaže na dejstvo, da je sicer pogosto omenjani pomenski kriterij: »pomen zveze je več kot vsota pomenov njenih delov« (Atkins in Rundell 2008: 167), na podlagi katerega zlasti leksikografi ločujejo kolokacije od večbesednih leksikalnih enot, neizogiben a hkrati izmuzljiv, saj tovrstnih zvez tako materni kot tuji govorniki ne prepoznavamo kot očitnih nosilcev pomena. Kriterij pomenske (ne)transparentnosti pa se uporablja tudi za ločevanje večbesednih leksikalnih enot od frazeoloških in idiomatskih enot.

1.3.3.5.3 V razmerju do frazeoloških enot

Področje frazeologije je znotraj jezikoslovnih raziskav, ki so vezane na leksikalne enote jezika, usmerjeno v prvi vrsti na tiste zveze, ki glede na svoje sestavne elemente izkazujejo samostojni, od sestavnih elementov neodvisni – t. i. *frazeološki pomen*. O frazeologiji se torej govori predvsem



takrat, ko se ob večbesednosti in ustaljenosti upošteva tudi idiomatičnost kot tista lastnost, ki določa, da nekatere enote glede na slovnična in pomenska razmerja sestavnih elementov niso jasno razstavljive. Ob tem se, vsaj v smislu pojmovanja frazeologije v ožjem smislu, za FE zahteva tudi prenesenost pomena, ki jo izkazuje bodisi posamezna sestavina ali zveza kot celota, kar ji daje glede na nefrazeološka enobesedna poimenovanja v besedilu ekspresivno oziroma zaznamovano noto. Prav zadnje iz okvira FE izloča tudi vsa večbesedna terminološka poimenovanja, ne glede na sicer izkazano idiomatičnost, npr. *črna skrinjica* 'naprava ...'; *sončni veter* 'energija ...'.

1.3.4 Nadaljnje delo pri leksikalni bazi za slovenščino

Z izdelavo poskusnih gesel v fazi oblikovanja standardov za izdelavo posameznih iztočnic v LBS se je v splošnem potrdila osnovna strukturiranost leksikalnih in skladijskih podatkov po posameznih nivojih LBS, hkrati pa so se pokazale možnosti vključevanja dodatnih informacij in nadgradnje zlasti na pomenskem in skladijskem nivoju.

Na pomenskem nivoju pričakujemo, da bo analiza obsežnejšega števila gesel z vidika pomenskih indikatorjev in elementov pomenskih shem omogočila poenotenje in tipologizacijo, ki bo neposredno uporabna pri pomenski členitvi in oblikovanju slovarskih definicij. Pri nekaterih pomenskih shemah tako poskusno uvajamo »diskurzni« ali Cobuildovski način zapisovanja, ki se ponekod dobro obnese tudi pri pojasnjevanju FE, za katere se je, zlasti v primeru večpomenskosti, pokazalo, da jih je težko pomensko razdvoumljati zgolj s kratkimi pomenskimi indikatorji.

Na skladijskem nivoju je ena od takih možnosti vključevanje informacij o skladijskih vlogah elementov pomenske sheme, in sicer v razmerju do skladijskih vzorcev in struktur. Zaenkrat predvidevamo, da bo analizo posameznih delov skladijskega vzorca v smislu stavčnočlenske vloge, zlasti ko gre za izražanje okoliščin, mogoče narediti naknadno, saj potrebujemo več gradiva, na katerem bo mogoče opazovati razmerje med predmetno (vezljivostno) in prislovnodoločilno (primično) funkcijo delov registriranih skladijskih struktur ter njihovo vlogo pri pomenskem razdvoumljanju LE. Posebno zanimivo se to poglavje kaže za slovnične opise slovenščine, zato



bomo v nadaljevanju izdelave LBS razmišljali o vključevanju – in kodifikaciji – podatkov, ki bi bili neposredno uporabni za slovnične analize in analize slogovnih značilnosti sodobne slovenščine.

Kompleksna struktura LBS in podrobnost pomenskega, skladišnega in kolokacijskega opisa LE v njej, omogoča v časovnih in finančnih okvirih projekta SSJ, analizo ozkega dela jedrnega besedišča slovenščine, hkrati pa se kaže kot soliden temelj za analizo obsežnejšega leksikalnega potenciala in predvsem za nadgradnje glede potrebnih pedagoških, slovarskih in računalniških aplikacij.

2 STANDARDI ZA IZDELAVO LEKSIKALNE ENOTE V LEKSIKALNI BAZI ZA SLOVENŠČINO

V nadaljevanju predstavimo postopke in podrobnosti pri oblikovanju iztočnic v LBS po posameznih vsebinskih nivojih, znotraj tega pa posebej opozarjamo na različne rešitve pri posameznih besednih vrstah, kadar do njih prihaja.

2.1 Osnovni podatki o leksikalni enoti

Osnovne informacije v zvezi z obravnavano LE so v LBS vezane na zapis LE v iztočnici v osnovni obliki in na podatek o besedni vrsti. S povezavo na leksikon besednih oblik²³ so na tem nivoju za posamezno LE dostopni tudi podatki o frekvenci in oblikoskladenjski podatki iz nabora oznak

²³ Izdelava leksikona besednih oblik je predvidena v okviru projekta Sporazumevanje v slovenskem jeziku. Cilj aktivnosti je baza podatkov o oblikoslovnih lastnostih besedišča, ki je izdelana po sistemu in v skladu z naborom oznak, razvitih v okviru projektov MULTEXT-East in JOS (gl. naslednjo opombo).



Multext-East²⁴ in projekta JOS.²⁵ Nekateri slovnični in pomenski podatki, kot je števnost, neštevnost, omejenost LE v določeni obliki paradigme ipd. niso predvideni na tem nivoju, ampak so prestavljeni bodisi na pomenski nivo (npr. ločevanje med vrstnimi in kakovostnimi pridevniki v samostojnih pomenih) v element *oznaka* neposredno za pomenskim indikatorjem bodisi na skladenjski nivo v element *restrikcija*.

Pri zapisovanju pridevnikov kot kanonično določimo nedoločno obliko, npr. *zdravstven, glasben*. Pri teoretično možnih vendar realno neobstoječih oblikah iztočnice, npr. *severen, zmoči* (1. os. ednine *zmorem*), se odločamo glede na realno stanje v jeziku oz. glede na prevladujočo rabo. Odločitev podpiramo s tem, da se v LBS ne ukvarjamo s slovarsko določenim oz. uporabniško prijaznim iskanjem iztočnic. V LBS je namreč vedno možna povezava z leksikonom besednih oblik. Vse druge informacije, ki so povezane s pomenom obravnavane iztočnice, tudi prevladujočo obliko za posamezni pomen ipd., kot rečeno, beležimo v elementu *oznaka*.

V tej fazi izdelave LBS na nivoju, kjer beležimo osnovne podatke o LE, ne predvidevamo glasoslovnih in pravopisnih podatkov, jih je pa mogoče vključiti v bazo naknadno.

2.2 Pomen leksikalne enote

Ker je primarni namen LBS natančno in podrobno razčleniti pomen obravnavanih LE, predvidevamo na tej ravni dva tipa pomenskih informacij:

²⁴ Celotno bibliografijo v zvezi z slovenskim leksikonom je mogoče najti na spletni strani Tomaža Erjavca: <http://nl.ijs.si/et/et-dom.html> in na MULTEXT-East spletni strani: <http://nl.ijs.si/ME/>.

²⁵ Jezikoslovno Označevanje Slovenščine: <http://nl.ijs.si/jos/>.



- **pomenske indikatorje** pri vseh obravnavanih besednih vrstah, stalnih besednih zvezah ter frazeoloških in idiomatskih enotah in
- **pomenske sheme** pri glagolskih, poskusno pa tudi pri nekaterih samostalniških in pridevniških iztočnicah.

Leksikalna baza za slovenščino, kot rečeno, ne vsebuje pomenskih razlag, da pa bi bilo to delo leksikografom olajšano, vključuje ob pomenski razčlenitvi vsake LE, ki je znotraj zgradbe LBS nadrejena informacija, pomensko opredelitev na omenjenih dveh segmentih.

2.2.1 Pomenski indikator

Pomenska členitev LE v LBS je prikazana s pomenskimi indikatorji. Pomenski indikator na kratko določi pomenski obseg ali področje obravnavanega pomena, pri večpomenskih besedah pa vzpostavlja razliko glede na druge pomene.²⁶ Pri oblikovanju pomenskih indikatorjev se ne spuščamo v podrobno pojasnjevanje pomenskih komponent, prav tako se, če je mogoče, izogibamo tako sinonimom kot prevelikemu posploševanju. Bistvene lastnosti in naloge pomenskih indikatorjev so torej:

- kratkost in jedrnatost
- razumljivost
- pomenska diferenciacija večpomenskih besed.

S kratkostjo in jedrnatostjo pomenskih indikatorjev se želimo izogniti oblikovanju slovarskih razlag, hkrati pa doseči hitro prepoznavnost pomenskega dosega posamezne LE oz. njenega pomena. Uporaba sinonima je ustrezna takrat, kadar z njim ne zajamemo področja preširoko in kadar lahko z njim vzpostavimo ustrezno pomensko razmerje do drugih pomenov in med

²⁶ Lahko bi rekli, da je osnovna funkcija pomenskega indikatorja uporabniku pomagati identificirati določen koncept tako, da ga priključuje iz lastnega konceptualnega spomina (Geeraerts 1990: 196), ali kot pravi Hanks (2000: 8): »Slovarji niso namenjeni definiranju, pač pa je njihova naloga ustvariti vrsto namigov in asociacij, ki pomagajo povezati neznano z znanim«.



podpomeni. Z razumljivostjo želimo doseči jasno in preprosto predstavo o pomenskem dosegu LE, zato se izogibamo besed, ki same potrebujejo razlago ali so redko uporabljane v vsakdanjem govoru. Deloma se pri oblikovanju indikatorjev, zlasti pa pri oblikovanju pomenskih shem, naslanjamo na pravila šolskega diskurza (Sinclair in Coulthard 1975), ki je uporabljen pri oblikovanju opisnih razlag v slovarjih, namenjenih učenju tujega jezika, konkretno za slovenščino v angleško-slovenskem slovarju BRIDGE (2000).

Zgled 5: Primerjava definicij v slovarju *BRIDGE* in indikatorjev v LBS za iztočnico *spomin*.

spomin

Bridge

One's memory je zmožnost, da si zapomnimo stvari.

A memory je nekaj, česar se spomnimo iz preteklosti.

A computer's memory je zmožnost računalnika, da shranjuje podatke.

Kadar nekdo piše *memoirs*, piše knjigo o svojem življenju.

One's reminiscences so stvari iz preteklosti, ki se jih človek spominja in o katerih govori ali piše

LBS

zmožnost pomnjenja

pomnjenje česa iz preteklosti
o računalniku

avtobiografsko literarno delo

predmet, stvari, ki spominjajo
na preteklost

2.2.2 Pomensko razdvoumljanje

Kot rečeno, je eden temeljnih namenov LBS podrobna pomenska členitev LE. Pri tem se naslanjamo na realno jezikovno stanje, kot ga ugotovimo iz pregleda niza konkordanc. Zelo običajna izkušnja pri pomenski analizi besed na podlagi realnih besedil je, da obstajajo redki t. i. prototipični primeri rabe za konkretni pomen in na drugi strani, da obstaja cela vrsta primerov, ki jih je težko z gotovostjo pripisati konkretnemu slovarskemu pomenu. V nadaljevanju zato predstavimo nekaj strategij razdvoumljanja besednih pomenov.

2.2.2.1 Kontekstualna informacija

Eden izmed kazalcev pomenske diferenciacije, ki ga upoštevamo pri pomenski členitvi, je sprememba v »obliki« LE. Ker sledimo metodologiji leksikalno-gramatičnega pristopa, spremembe



v obliki ne zadevajo samo LE, pač pa navadno tudi njeno besedilno okolje. Tako izhodišče, ki se naslanja na Sinclairjevo ugotovitev, da je vsaka sprememba v pomenu neposredno povezana s spremembo v obliki (Sinclair, citirano v Moon 1987: 89), lahko pride do izraza šele z identifikacijo tipičnih oz. ponavljajočih se vzorcev in znotraj njih uresničenih pomenov. Pomemben kazalec pomenske spremembe pri glagolskih iztočnicah je namreč sprememba v skladijskem vzorcu glagola, zlasti sprememba števila, vloge ali vrste udeležencev, ki zasedajo vezljivostna mesta, pa tudi okoliščin, kadar so za priklic konkretnega pomena relevantne.

Če primerjamo zglede (a) in (b) za glagol *pasti*:

- (a)
- *Malo prej je v lokal vdrlo kakih osem fantov, ki so padli po 50-letnem lastniku lokala.*
 - *Na Plečnikovem trgu so obstopili nekaj fantov in deklet, zagrozili in zahtevali denar, potem pa padli po 18-letnem fantu.*
- (b)
- *Po nesrečnejžu je padlo nekaj močnih udarcev, tako da ga je oblila kri.*
 - *Nepripravljena sta najprej prosila za cigareto, potem pa je padlo po nesrečnem mladeniču.*

lahko rečemo, da gre v obeh primerih za izvajanje nasilja nad človekom, ki se izraža z zvezo glagola in predloga *pasti po (kom)*, vendar pa je v prvi skupini zgledov, ta, ki izvaja nasilje, človek (*kdo (=človek) pade po kom = 'napasti'*), v drugem primeru pa so posebej izpostavljeni udarci (*kaj (=udarci) pade po kom = 'o udarcih'*), čeprav pogosto tudi brezosebno, kot izkazuje drugi zglede v skupini (b) (*Ø pade po kom*). Razločevanje na ravni vrste udeleženca v skladijskem vzorcu glagola nas torej navaja na dva ločena podpomena: 'napasti koga' in 'o udarcih'. Posebej se ta razlika, kot bomo še opozorili, odraža pri oblikovanju pomenskih shem.

Kazalci pomenske diferenciacije so lahko vezani tudi na spremembo števila udeležencev, ki jih predvideva glagol v svojem pomenu. Spodaj ustreza zglede (a) jedrnemu pomenu glagola *pasti*:

'premikati se v smeri od zgoraj proti navzdol'. V tem pomenu zahteva glagol zgolj enega udeleženca, navadno je to predmet, in lokacijo, ki pa ni nujno izražena:

(a) *Propeler je razsekalo in helikopter je kot kamen padel v globino.*

Do pomenske spremembe, ki jo želimo zabeležiti v LBS, pride, ko se število nujnih udeležencev spremeni, kot kaže zgled (b). Za podpomen glagola *pasti* 'izgubiti, izpustiti' sta namreč potrebna vsaj dva udeleženca: predmet, ki je izgubljen ali izpuščen, človek, ki ta predmet izgubi ali izpusti, in lokacija, od koder (lahko tudi kam, vendar neobvezno) predmet pade:

(b) *Sneguljčici zastrupljeni košček jabolka pade iz ust in dekle oživi.*

Pri pomenski členitvi LE v LBS smo pozorni na dve vrsti razmerij med pomeni: na razmerje med jedrnim pomenom in podpomeni in na razmerje med podpomeni znotraj jedrnega pomena.

2.2.2.2 Razmerja med jedrnim pomenom in podpomeni

Kot smo videli pri NODE, obstajajo med jedrnim pomenom in podpomeni običajno trije tipi osnovnih pomenskih razmerij, in sicer, (a) figurativni nasledek jedrnega pomena, (b) specializirani pomen jedrnega pomena in (c) druge razširitve ali spremembe pomena, ki ohranjajo enega ali več elementov jedrnega pomena.

Zgoraj smo pri samostalniku *spomin* lahko videli, da obsega celoten spekter različnih pomenov. Ti pomeni so bodisi osamosvojeni ali pa se na različne načine oddaljujejo od jedrnega pomena. V tem procesu prihajajo v ospredje različni vidiki jedrnega pomena, ki se navadno aktivirajo glede na specifično besedilno okolje. Tako lahko *spomin* razumemo kot nekaj, česar se človek spominja iz preteklosti, npr. *lepi, nostalgični, prijetni, nepozabni, mladostni spomini*, ali pa nekaj, kar človeka spominja na preteklost, npr. *dragocen spomin, dobiti kaj za spomin*. Številne tovrstne rabe so dovolj stabilne in pogoste, da jih je mogoče obravnavati kot različne slovarske podpomene, medtem ko so



nekatero druge rabe, čeprav še vedno povezane z jedrnim pomenom, preveč obrobne in idiosinkratične, da bi jih bilo mogoče z gotovostjo obravnavati kot (pod)pomene. Na teh skrajnih robovih, kot trdita Atkins in Rundell (2008: 287), meje nikoli niso čiste, zato je leksikografska obdelava v veliki meri odvisna od namena in podrobnosti pomenske členitve kot tudi od stanja, ki ga izkazujejo korpusni zgledi.

2.2.2.2.1 Figuratívni nasledek jedrnega pomena

Zelo tipičen primer odstopanja od jedrnega pomena je figurativna raba jezika, kjer se očitno kaže podobnost med dobesednim ali primarnim pomenom, npr. rana kot 'poškodba na telesu' in nedobesednim, prenesenim, pogostokrat abstraktnim pomenom, npr. rana kot 'psihična ali čustvena bolečina'. Še posebno očiten je ta proces pri imenih za dele telesa in telesne organe, ki vzpostavljajo preneseno razmerje prek oblike, lokacije ali funkcije izhodiščnega jedrnega pomena (Atkins in Rundell 2008: 287).

Zgled 6: Figuratívni nasledki jedrnega pomena za glagol *pasti* in samostalnik *rana*

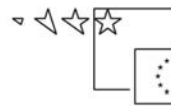
pasti

- ▶ *premikati se v smeri od zgoraj proti navzdol*
[bomba] pade
 - ▶▶ *o udarcih*
pade po [kom]
- ▶ *izgubiti ravnotežje*
pasti pod [traktor, kolo/kolesa]
 - ▶▶ *ne biti več v določenem stanju, razpoloženju*
pasti iz [ritma]

rana

- ▶ *poškodba na telesu*
rana [skeli, boli, zija]
 - ▶▶ *o pokrajini*
rana v [naravi, okolju]
 - ▶▶ *primanjkljaj, praznina*
[zevajoča] rana

Kadar se figurativni pomen v jeziku osamosvoji do te mere, da oblikuje svoje lastne odklone in specializacije, ga v LBS lahko povzdignemo na raven samostojnega pomena kot spodaj v primeru samostalnika *rana*.



Zgled 7: Osamosvojitvev figurativnega pomena glagola *pasti* in samostalnika *rana*

pasti

- ▶ *umreti*
- pasti pod [streli, krogli, udarci]
 - ▶▶ *prenehati obstajati*[berlinski zid, železna zavesa] pade

rana

- ▶ *psihična ali čustvena bolečina*
- rana na [duši]
 - ▶▶ šport *pri porazu*[zaceliti] rane

Zelo pogost primer figurativne rabe jezika je *metonimija*, ki se nanaša:

(a) na razmerja med posameznim delom in celoto, npr. organ – človek/žival (*nahraniti veliko ust = 'veliko otrok, ljudi'*); drevo – les drevesa; stavba – ljudje v stavbi (*dvorana je navdušeno vzklikala*); država kot ozemlje – država kot celota kulturnih, jezikovnih idr. značilnosti – vodstvo te države; mesto kot upravna enota – mesto kot prostor (*hoditi po mestu*) ali prebivalci mesta (*novica se je širila po mestu*);

(b) asociativno povezana poimenovanja, npr. rastlina, živilo, organ – jed iz te rastline/živila/organa, npr. *sušiti čaj : skodelica čaja; izvaliti jajce : trdo kuhano jajce* itd.; predmet – dejavnost povezana s tem predmetom, npr. *igrati kitaro – ure kitare; francoščina mi ne gre – iti na francoščino* ipd. Omenjene in podobne primere metonimije obravnavamo v LBS znotraj podpomenov, kot npr. v zgledu za samostalnik *šola*.

Zgled 8: Metonimično povezani podpomeni znotraj jedrnega pomena samostalnika *šola*

šola

1 o izobraževanju

- 1.1 ustanova** [osnovna, srednja, visoka, poklicna] šola
- 1.2 stavba** [igrišče, parkirišče] pred šolo
- 1.3 ljudje v stavbi** [cela, vsa] šola
- 1.4 proces** [hoditi, vpisati se] v šolo
- 1.5 storitev** šola [navtike, jahanja, jadriranja, smučanja, plavanja]

Jezikovna realnost, kot jo izkazujejo korpusi, zelo očitno kaže, da je vsak besedni pomen mogoče naknadno izkoristiti in da taka izkoriščanja, kot smo videli zgoraj, nastajajo zelo pogosto, lahko



tudi individualno za konkretno sporazumevalno potrebo. Kadar tak pomenski izkoristek postane splošno sprejet in uporabljan, ga, kot smo videli zgoraj, uvrstimo na raven podpomena. Veliko težje ulovljive so razpršene figurativne rabe, pomenske širitve in zožitve, ki jih glede na podrobnost pomenske členitve sicer ne želimo izgubiti, hkrati pa niso dovolj »čiste« in pogoste, da bi jih bilo mogoče registrirati kot jezikovni standard pri katerem od osnovnih pomenov. Taka odstopanja so npr. za glagol *stisniti* razvidna iz spodnjih zgledov:

- *Ves čas so jim govorili: **Stisnite**, bolje bo! In so **stisnili**, zdaj pa ni nič boljše. Slabše je!*
- *Nenadoma pa me je **stisnil** išias, zato nisem mogel hoditi.*
- *Enega bom **stisnila** na avionu.*
- *Vedel se je kot tipičen oče, ki sina spočetka krišpa kot blazen, a ko se ta nikakor noče poboljšati, pobere rekeljc in jo po tihem **stisne** v bife, skrb za vzgojo pa prepusti materi Cerkvi in njenemu večnemu nerganju.*

Ker frazeološke enote beležimo v samostojnem razdelku LBS in ker frazeologija (poleg strukturne ustaljenosti) navadno vključuje figurativne, zlasti metaforične in metonimične nasledke katerega od pomenov besede, ki jo obravnavamo v iztočnici, se moramo odločati tudi za bodisi samostojen figurativni pomen iztočnice bodisi za njeno vključitev v frazeološki razdelek (prim. zgoraj: *Ves čas so jim govorili: Stisnite, bolje bo! In so stisnili, zdaj pa ni nič boljše. Slabše je!*, kjer gre za skrajšano varianto FE: *stisniti zobe* – '*potrpeti*'). Odločitev redaktorja je povezana predvsem s tem, v kakšni meri je ustaljeno tudi besedilno okolje, v katerem se pojavlja iztočnica v prenesenem pomenu, in s katerimi slovničnimi možnostmi in skladenjskimi načini (s t. i. oblikami rabe) se konkretni pomen izraža v besedilu.

Če vzamemo za zgled glagol *stisniti*, ki se realizira v kolokacijah kot: [*srce*] *se stisne komu*; *koga stisne v* [*prsih, želodcu, grlu*] in *koga stisne pri* [*duši, srcu*], pri čemer so posamezni deli kolokacij razmeroma trdni (npr. *srce*) oz. zamenljivi znotraj ožjega pomenskega polja (določeni deli telesa: *srce, prsi; želodec, grlo*), kar je pogostokrat signal, da gre bodisi za samostojni pomen bodisi za FE z izkazanim variantnim nizom in različnimi oblikami rabe. Ker pa se glagol *stisniti* v istem pomenu tipično rabi



tudi v »izoliranem« (tj. nekolokabilnem) besedilnem okolju: *koga stisne (... odložil je kozarec in si pokril oči z rokami. Nič se mi ni smilil, ampak vseeno me je malo stisnilo)*, mu lahko pripišemo samostojni – oz. glede na katerega od osnovnih – preneseni podpomen: 'prestrašiti se'.

2.2.2.2.2 Specializirani pomen jedrnega pomena

To, kar pogostokrat vpliva na pomensko členjenje LE, je določeno področje, na katerem se LE tipično pojavlja, kot je razvidno iz spodnjega zgleda.

Zgled 9: Specializirani pomen jedrnega pomena

stisniti

► računalništvo

stisniti s [programom, algoritmom]

stisniti v [datoteko]

Tudi v teh primerih je informacija o specializiranem ali kakorkoli omejenem področju najpogosteje razvidna iz sobesedila:

- *Velike datoteke je zato priporočljivo stisniti s programi, kot sta WinZip ali PKZip.*
- *Program izbrane datoteke še dodatno stisne, tako da na praznem mediju zavzamejo kar najmanj prostora.*

V nekaterih primerih se tak specializirani pomen lahko tudi osamosvoji na raven jedrnega pomena, ker bodisi ne izkazuje oprijemljive odvisnosti od katerega jedrnih pomenov bodisi sam na sebi izkazuje večpomenskost.

Zgled 10: Večpomenskost specializiranega pomena za samostalnik *spomin*

spomin

► o računalniku

- del računalnika

►► *zmogljivost računalnika*

2.2.2.2.3 Druge razširitve in spremembe jedrnega pomena

Ob kontekstualnih informacijah, ki vplivajo na spremembe jedrnega pomena obravnavane LE, pogostokrat naletimo tudi na t. i. notranje indikatorje pomenskih sprememb, kot so že omenjeno skladijsko in leksikalno-gramatično obnašanje LE, tipični kolokatorji ali omejitve kolokabilnosti na določeno pomensko polje besed, ki se tipično sopojavljajo ob LE v konkretnem pomenu, kot prikazujeta spodnja zgleda.

Zgled 11: Spremembe pomena, ki ohranjajo enega ali več elementov jedrnega pomena

stisniti

- *uporabiti fizični pritisk*
stisniti [roko, dlan, vrat, ročaj]
 - *pritisniti in sprožiti*
stisniti [sklopko, gumb, tipko, ročico, zavoro]

spomin

- *pomnjenje česa iz preteklosti*
[prepustiti se, predajati se] spominom
 - *nekaj, kar spominja na preteklost*
[dragocen] spomin

Tu je potrebno opozoriti tudi na večpomenskost, ki nastaja pod vplivom časa in nove predmetnosti: nekatere LE namreč s časom pridobijo nove pomene ali pa novi pomeni v celoti nadomestijo stare. Za glagol *brisati* je na primer ob pomenih 'delati suho' 'delati čisto' in 'delati manj vidno' (SSKJ) glede na rabo v sodobni slovenščini postal očiten predvsem pomen 'odstranjovati' zlasti v smislu računalniških datotek, programov ali besedila na sploh.

V LBS dopuščamo – kot tehnično rešitev pri organizaciji jedrnih pomenov – možnost izpostavljanja jedrnega pomena, ki služi zgolj kot **krovna nadpomenka** za vse podpomene, ki jih navajamo pod njim. Gre za primere, kjer si podpomene sledijo v enakovrednem nizu, sodijo torej v okvir istega pomena, vendar nobenega izmed njih ni mogoče povzdigniti na raven jedrnega pomena. V takih primerih navedemo v programskem vmesniku pri jedrnem pomenu zgolj indikator, ostalih elementov (tj. pomensko shemo, vzorce, strukture in kolokacije) pa ne navajamo.



Zgled 12: Jedrni pomen kot nadpomenka za pripadajoče podpomene

pomešati

- ▶ združiti različno v novo celoto
 - ▶▶ o snoveh (konkretno)
pomešati [sestavine]
 - ▶▶ o predstavah (abstraktno)
pomešati [religije, sloge]

šola

- ▶ o izobraževanju
 - ▶▶ ustanova
 - ▶▶ stavba
 - ▶▶ process
 - ▶▶ storitev/tečaj

pasti

- ▶ znižati se na lestvici
 - ▶▶ o količini ali vrednosti
[tečaj, indeks, cena, vrednost] pade
 - ▶▶ o intenzivnosti
pasti pod [dopustno, kritično] mejo
 - ▶▶ o statusu ali kakovosti
pasti na [raven, ničlo, dno]

Pomembno: Pri podrobni pomenski členitvi LE v LBS smo sprejeli splošno pravilo, da se povsod tam, kjer je posamezne pomenske odtenke med seboj težko ločevati, odločamo za podrejanje jedrnemu pomenu, torej za podpomene in ne za pomensko osamosvajanje na raven jedrnih pomenov. Tipična situacija te vrste je, ko prihaja do pomenskega prepletanja tipa: šola kot *ustanova* – *stavba* – *učni proces* v večini korpusnih zgledov, ob njih pa v manjšem obsegu obstajajo tudi taki, kjer je razlika očitna. V splošnem torej velja, da določamo jedrne pomene na razmeroma široki ravni, kar pomeni, da težimo k manjšemu številu jedrnih pomenov, težo podrobne pomenske členitve pa nosijo podpomeni.

2.2.2.3 Razmerja med podpomeni znotraj jedrnega pomena

Med podpomeni znotraj jedrnega pomena obstajajo naslednje možnosti hierarhične razporeditve: jedrni pomen (ali eden od jedrnih pomenov) je osnovni ali osrednji pomen, od njega pa se pomensko odmikajo podpomeni. Ti so lahko med seboj neposredno povezani, kot v zgledu (a); drug od drugega neodvisni oziroma se v različnih smereh odmikajo od jedrnega pomena, kot v



zgledu (b), lahko pa se postopno odmikajo, tako da je zadnji podpomen v vrsti precej bolj odmaknjen od jedrnega pomena kot prvi podpomen, npr v zgledu (c).

(a) povezani podpomeni

pasti

- | | |
|--|---------------------------|
| ▶ premikati se od zgoraj proti navzdol | [bomba] pade |
| ▶▶ o svetlobi | [mrak, senca] pade |
| ▶▶ o padavinah | [padavine, kaplje] padejo |
| ▶▶ o drevesu | [drevesa] padejo |

(b): neodvisni podpomeni

pasti

- | | |
|--|------------------------------|
| ▶ premikati se od zgoraj proti navzdol | [bomba] pade |
| ▶▶ o oblačilih | [lepo, lepše, mehko] pasti |
| ▶▶ o denarju | pade v [blagajno, denarnico] |
| ▶▶ o delih telesa | [lase, lasje, roke] padejo |

(c) postopno odmikanje od jedrnega pomena

pasti

- | | |
|--------------------------|--------------------------------------|
| ▶ znižati se na lestvici | [tečaj, indeks, cena, vrednost] pade |
| ▶▶ o količini, vrednosti | pasti na [ničlo] |
| ▶▶ o intenzivnosti | pasti na [zadnje] mesto |
| ▶▶ o statusu, kakovosti | pasti na [družbeno dno] |
| ▶▶ o socialnem položaju | |

Pomembno: Kadar jedrnemu pomenu glagola sledi več podpomenov, navajanje indikatorjev pri posameznih podpomenih ni obvezno, če z indikatorjem ni mogoče vzpostaviti ustreznega pomenskega razločevanja med podpomeni. V takih primerih mora biti pomenska razlika med podpomenoma jasno razvidna iz pomenske sheme. Pri samostalniških, pridevniških in prislovnih iztočnicah, kjer pomenskih shem ne pišemo, je navedba indikatorja nujna.



2.2.2.4 Slovnične, funkcijske in pragmatične omejitve pomena

V programskem vmesniku navajamo neposredno za pomenskim indikatorjem v elementu *oznaka* področne, slovnične, pozicijske in funkcijske omejitve pomena, kot se odražajo v dejanski rabi.

Zgled 13: Slovnične in funkcijske omejitve pomena

(a) področne omejitve

- pri/v športu
- v računalništvu

rana 2. psihična ali čustvena bolečina
oznaka: v športu: *pri porazu*
... *Celjani še niso zacelili vseh ran iz Ljudskega vrta*

(b) slovnične omejitve/preference

- navadno v 3. osebi
- neštevno/števno

pasti 1. premikati se od zgoraj proti navzdol
oznaka: nav. v 3. osebi
... *gondoli sta padli v prepad*
senca 1. zaradi padanja svetlobe
oznaka: neštevno
... *v steklu vidim Petrino senco*

(c) funkcijske omejitve/preference

- v povedni rabi

vesel 2. tak, ki čuti veselje
oznaka: v povedni rabi
... *najbolj sta vesela obiskov otrok*

(d) pozicijske omejitve/preference

- pred samostalnikom

glasben 1. tak, ki se nanaša na glasbo
oznaka: pred samostalnikom
... *gost je bil glasbeni ustvarjalec*

Poleg omejitev in preferenc beležimo znotraj pomenskega indikatorja v elementu >pr< (pragmatika) tudi omejitve, ki se nanašajo na rabo besede v določenem pomenu, kadar je ta povezana bodisi z odnosom govorečega do vsebine besedila bodisi z udeleženci v govorni situaciji, z zunajjezikovnimi okoliščinami, kulturnimi značilnostmi diskurza ipd. Tu izhajamo iz spoznanja, da pragmatike ni mogoče povsem ločiti od semantike in da ima pravzaprav vsaka leksikalna enota v jeziku pomen, ki je slovarsko-klasifikatorne narave ter takega, ki odraža emotivne, družbeno



naravnane oz. medosebne elemente komunikacije, tj. pragmatični pomen. Pragmatike ni mogoče zaobiti pri obravnavi nobene od LE, ki eksplicitno izražajo odnos tvorca do vsebine izrečenega (npr. kletvice, pohvale, opravičila, prošnje, zahteve itd.) – na tej ravni govorimo zlasti o konotaciji –, kot tudi ne pri tistih LE, kjer je funkcija jezikovne enote bolj ključna za razumevanje pomena kot njen goli leksikalni pomen.

Zgled 14: Pragmatične omejitve pomena

pobirati

za izražanje kritike: o neupravičenem zaslužku

Mi delamo, a direktor pobira denar.

2.2.3 Razvrstitev pomenov

Pomenska razčlenitev je leksikografsko najpomembnejši in hkrati najzahtevnejši postopek. Ob tem, da je različne pomene težko prepoznati in še težje opisati, je pogosto tudi težko ločiti to, kar je pomensko jedrnega, središčnega od tega, kar je pomensko obrobnega oz. od jedrnega odvisno. Načeloma smo sprejeli odločitev, da pomene in podpomene razvrščamo glede na relativno pogostost, ki jo ugotavljamo na podlagi pregledanega konkordačnega niza in glede na obseg in število kolokatorjev pri posameznem pomenu in podpomenih. O relativni pogostosti govorimo zato, ker je tako s pomočjo korpusa kot s pomočjo besednih skic, kjer sicer izstopajo posamezni kolokatorji glede na število pojavitev in glede na statistične izračune besedne sopovezovanosti, frekvenco posameznega pomena nemogoče ugotoviti avtomatično. V izhodišču sprejeta težnja, da želimo pomene besed čim bolj drobiti in čim bolj podrobno zajeti vse pomenske odtenke, še dodatno zabrisuje pogostost jedrnih pomenov in relativizira pogostost posameznih podpomenov znotraj njih. Hkrati so isti kolokatorji, ki sicer izstopajo po frekvenci in statistikah, lahko različno razporejeni med posamezne pomene in podpomene. Ni redka tudi možnost, da je kateri od podpomenov frekventnejši od jedrnega pomena. Zaradi tega je razporeditev pomenov in

podpomenov v LBS glede na realno pogostnost v jeziku relativna, pogostokrat povezana tudi z besedilnim tipom ali zvrstjo, v katerem določena zveza ali pomen besede tipično nastopata.

2.2.4 Pomensko in besednovrstno povezane LE

Besednovrstno povezane LE, kot so: deležje – prislov, deležnik – pridevnik, glagolnik – samostalnik in števnik – pridevnik, obravnavamo znotraj ene iztočnice (lahko znotraj ene besednovrstne opredelitve ali z izbiro nadgesla). To pomeni, da lahko npr. deležnik *stisnjen* – *stisnjen med hribe*, obravnavamo (a) v okviru enega gesla pod iztočnico *stisnjen*, ki je besednovrstno opredeljena kot pridevnik (*stisnjena datoteka*), pomenske razlike in specifičnosti v razmerju deležnik – pridevnik pa so, če obstajajo, razvidne iz pomenske členitve znotraj gesla, ali (b) v okviru pridevniškega nadgesla s (pod)besednovrstno oznako deležnik. Prav tako imamo možnost opozoriti na preferenco določene oblike (npr. navadno deležnik, trpna oblika ipd.) v elementu <R> (restrikcija) znotraj skladijske strukture (gl. poglavje o slovničnih omejitvah in preferencah). Povratnosvojilne glagole, ne glede na njihovo dejansko povratnost, npr. *umivati se* (= sebe) ali *umivati koga*, v nasprotju s *poboljšati se*, *smejati se* ipd., kjer *se* nima pomenske vloge, navajamo pod skupno glagolsko obliko (brez *se*, razen če ni imanentni del glagola). Zgolj na določene pomene vezane povratnosvojilne glagole oz. morfem *se* ali *si* predstavljamo pod ustreznimi pomeni ali podpomeni, npr. *pobirati (koga/kaj)* – 'postavljati pokonci' : *pobirati se* – 'vstajati s tal' : *pobirati se* – 'postopoma si opomoči', tako tudi v primerih, kjer je *se* recipročen, npr. *objeti se*, *pozdraviti se* – tj. 'med seboj' itd. Problem trpnosti, tipa: *stvari se pozabljajo* ne beležimo, razen če je brezosebno izražanje v smislu splošnega vršilca dejanja za konkretni pomen LE tipično. V tem primeru tako preferenco LE označimo na ravni skladijske strukture ali vzorca v elementu *restrikcije*.

Samostalnike, ki opravljajo funkcijo pridevnika v vlogi levega prilastka, kot npr. *mini (krilo)*, *jeans (moda)* ipd., ter pridevnike v samostalniški vlogi (t. i. posamostaljene pridevnike), npr. *padli (v boju)*, *(število) brezposelnih* ipd., obravnavamo v okviru nadgesel. Tako povezanih LE namreč ne želimo obravnavati ločeno, zato posamostaljene pridevnike obravnavamo v okviru pridevniškega nadgesla, samostalnike v vlogi levega prilastka pa v okviru samostalniškega nadgesla.



Tabela 4 : Pomensko in besednovrstno povezane LE

besedna (pod)vrsta	obravnavava v LBS		
	primer	besednovrstna (pod)oznaka	nadgeslo
deležje	miže, grede	deležje	prislov
deležnik	<i>stisnjen, rojen</i>	deležnik	pridevnik
glagolnik	<i>pobiranje</i>	glagolnik	samostalnik
povratni glagol	smejati se	glagol	
posamostaljeni pridevnik	<i>padli (v boju)</i>	samostalnik	samostalnik
samostalnik v vlogi pridevnika	<i>jeans (jakna, moda)</i>	pridevnik	samostalnik
števnik	<i>provi</i>	števnik	pridevnik

2.2.5 Homonimija in večpomenskost

V LBS ne obravnavamo homonimnih LE tipa *kos* – 'del celote' in *kos* – 'ptica' kot samostojnih iztočnic, kadar gre za besednovrstno enake LE, pač pa jih obdelamo znotraj istega gesla kot večpomenske LE. V različnih iztočnicah obravnavamo le tiste homonimne LE, ki ne izkazujejo enake spregatvene oz. sklanjatvene paradigme, npr. *pasti* – *pasem* : *pasti* – *padem*, *zlagati* – *zlagam* : *zlagati se* – *zlažem se*, kar ustreza stanju v leksikonu besednih oblik.

2.2.6 Stalne besedne zveze

Pri samostalniških in pridevniških iztočnicah obravnavamo pod posameznimi pomeni (tj. za vsemi podpomeni in pripadajočimi skladenjskimi strukturami in zvezami) ali podpomeni, kadar je pomenska povezava med podpomenom iztočnice in pomenom zveze nedvoumna, večbesedne leksikalne enote ali stalne (besedne) zveze (ang. multi-word expressions, compounds).



Stalne besedne zveze vključujemo v LBS na podlagi večbesednosti, pomenske samostojnosti oz. vsaj delne neodvisnosti pomena zveze od njenih sestavin ter na podlagi ločevanja od frazeoloških enot, za katere velja metaforična ali kako drugače očitna ekspresivna ali pragmatična vrednost.²⁷

Poleg omenjenega si pri določanju stalnih besednih zvez pomagamo z naslednjimi merili:

- zlasti samostalniške stalne besedne zveze imajo konkretnega nosilca, npr. *bančna kartica, kisl juha*;
- poleg pomenske upoštevamo tudi strukturno ustaljenost zveze, npr. *krompir v oblicah, pire krompir*;
- variantni niz sestavin zveze je omejen na sinonime oz. na ozko pomensko polje, npr. *kratkoročni/ kratkotrajni spomin; glasbena skupina/zasedba*;
- stalne besedne zveze pogosto izkazujejo lastno kolokabilnost: *[ponuditi] krompir v oblicah; [pravilnik, odlok] o obratovalnem času, [skrajšan, nedeljski] obratovalni čas*, ali pa se razširjajo z "notranjimi" kolokatorji in se sestavljajo: *iskanje (skritega) zaklada – 'družabna igra', predstavnik predstavnic, za javnost/ stike z javnostjo, [ponuditi] [pečen] krompir v oblicah; [štajerska] kisl juha*.
- lahko so večpomenske, npr. *trojanski konj* – 1. 'po grškem mitu'; 2. 'nekaj, s čimer se prikriva čigave slabe namene'; 3. 'računalniški virus', oz. imajo poleg dobesečnega tudi celostni/nedobesečni pomen: *poletni čas – 'poleti'; poletni čas – geografija: 'časovni pas'*;
- navadno je njihova raba omejena na določeno (terminološko) področje: *trojanski konj – 'računalništvo'; kisl dež – 'ekologija'*.
- lahko podlegajo različnim pretvorbam in poenobesečenju oz. imajo sinonimno kratično ime, npr. *porodniški dopust – porodniška; upravna enota – UE* ipd.

V LBS kot večbesedne leksikalne enote ali stalne besedne zveze obravnavamo torej besedne zveze, ki so oblikovno in pomensko osamosvojene, kar pomeni, da imajo svoj lastni pomen, ki je pogosto vezan na določeno strokovno ali kako drugače omejeno področje rabe, pri čemer vsaj ena sestavina

²⁷ Načeloma velja, da besedne zveze s stalno besedilno funkcijo, kot so npr. diskurzni označevalci tipa: *konec koncev, če me spomin ne vara* ipd. uvrščamo v frazeološki razdelek.



kaže asociativno povezavo s katerim od pomenov besede v iztočnici. Tipičen pokazatelj leksikaliziranosti je tudi večpomenskost stalne besedne zveze oz. dejstvo, da obstaja tudi dobessedna, navadno neterminološka homonimna besedna zveza, npr. *črna luknja* v primeru: *črne luknje razbitih oken* in *črne luknje so ključ za razumevanje poslednje usode vesolja*. Variantnost sestavin stalne besedne zveze je omejena na sinonimijo ali določeno pomensko polje, hkrati pa imajo stalne besedne zveze pogostokrat tudi lastno tipično besedilno okolje, ki ga beležimo v obliki kolokacij.

Merila določanja stalnih besednih zvez so razmeroma ohlapna in pogosto se ni mogoče zanašati na eno samo, saj je razpon med relativno prostimi in že frazeološkimi zvezami velik, »vmesno področje«, ki zajema najproduktivnejši tip besednozveznih enot, pa je težko določljiv. Ker gre za področje leksike, kjer jasnih mej ni mogoče potegniti (te namreč ne obstajajo), se pogostokrat zanašamo tudi na lastni občutek o leksikaliziranosti zveze, o njeni terminološkosti (glede na referenčnost korpusa, iz katerega pridobivamo podatke, ni mogoče pričakovati obsežnega nabora terminoloških zvez, ki so značilne za komuniciranje v ozkem krogu strokovnjakov, ampak gre primarno za splošno znane izraze, ki so zgolj tipično vezani na določeno področje rabe), na poznavanje terminološkega področja in na jezikovno kontrastivo: številne stalne besedne zveze, ki maternemu govorniku ne predstavljajo pomenske neprozornosti, npr. *jedilna, temna čokolada* – ang. *plain, dark chocolate*; *črni humor* – ang. *black, dark humour*, imajo v tujem jeziku svoje lastne ustreznike. Kljub temu ostaja pri odločanju za uvrstitev med kolokacije ali zveze na eni in med stalne besedne zveze na drugi eno bistvenih meril prav produktivnost pomena posamezne sestavine: če namreč konkretna sestavina potencialne stalne besedne zveze izkazuje pomensko produktivnost v tem smislu, da pomen, ki ga izraža v konkretni zvezi, izraža tudi ob različnih drugih kolokatorjih, npr. *šola* kot ustanova: [osnovna, srednja, visoka] šola, gre po vsej verjetnosti za isti del zveze, ki ne pripomore k leksikaliziranosti zveze kot celote. Nasprotno je potreba po kombiniranju z omejenim številom kolokatorjev za izražanje konkretnega pomena znak, da imamo opraviti s pomensko trdnejšo zvezo, ki jo je smiselno beležiti kot stalno besedno zvezo. Če torej pridevniki kot: *osnovni, srednji, visok* ipd. izražajo ustrezni konkretni pomen samo v kombinaciji s *šola*, razmišljamo o kandidatih za stalne besedne zveze. Zlasti vrstni pridevniki pogosto tvorijo s



samostalnikom trdnejše besedne zveze, ki niso nujno frazeološke ali terminološke,²⁸ so transparentne, predvidljive in ne povzročajo nobenih težav pri razumevanju, a jih vendarle prepoznamo kot poimenovalne koščke jezika, ki odražajo bolj ali manj konkreten segment jezikovne realnosti. V okviru stalnih besednih zvez jih beležimo, ker so izrednega pomena pri jezikovni produkciji in v prevajalski praksi.

2.2.6.1 Pomensko podrejanje stalnih besednih zvez

Z vključevanjem stalnih besednih zvez pod pomene ali podpomene želimo v LBS doseči dvoje: slediti merilom, po katerih so stalne besedne zveze – oz. je pomen katere od njihovih sestavin – očitno pomensko povezane s katerim od registriranih jedrnih pomenov iztočnice (gl. zgled spodaj), obenem pa opozoriti na dejstvo, da gre za pomensko dovolj osamosvojene zveze kot celote (npr. z vidika tujega govorca), ki potrebujejo lastno razlago (v LBS samo v obliki indikatorja). Kot take so stalne besedne zveze v konkretnem slovarju lahko obravnavane kot samostojne iztočnice.

Zgled 15: Pomensko podrejanje stalnih besednih zvez

<i>konj</i>	<i>spomin</i>	<i>kisel</i>
žival	del mišljenja, namenjen pomnjenju	okus
<i>arabski konj</i>	<i>kratkoročni spomin</i>	zelenjava
<i>islandski konj</i>	<i>kratkotrajni spomin</i>	<i>kislo zelje</i>
<i>lipicanski konj</i>		<i>kisla repa</i>
<i>posavski konj</i>		mlečni izdelek
<i>hladnokroni konj</i>		<i>kislo mleko</i>
<i>toplokroni konj</i>		<i>kisla smetana</i>
<i>polnokroni konj</i>		
<i>nilski konj</i>		
telovadno orodje		
<i>konj z ročaji</i>		

²⁸ Pojem terminološkosti je problematičen sam na sebi, saj je pogojen vsaj z dvema dejavnikoma: (1) razširjenostjo zveze na področje splošne rabe, npr. razmeroma enotno se kot stalna besedna zveza dojema *črevesna kolika*, veliko več premisleka pa zahteva zveza *zdravstvena nega*; (2) lastna seznanjenost z določenim področjem.



Pomembno: V številnih primerih se ni mogoče nedvoumno opredeliti bodisi za besedo, kjer je njena tipična besedilna okolica v obliki kolokatorjev odraz relativno zamejenega kolokacijskega niza, npr. *kisel [dež, plin]*, kar v LBS beležimo na ravni kolokacij, bodisi za stalno besedno zvezo, kjer katera od sestavini izraža relativno omejen (sinonimno, pomensko) variantni niz. Glede na to, da dokončnih mej med pomensko določenimi kolokacijami in relativno pomensko prozornimi stalnimi besednimi zvezami ni mogoče potegniti, je tako vključitev med kolokacije kot med stalne besedne zveze ustrezna. Kljub temu pa smo v LBS v zvezi s tem sprejeli načelno odločitev, da težimo k razbijanju podobnih kolokativnih nizov na stalne besedne zveze z namenom, da se približamo možnim različnim ustreznikom stalnih besednih zvez v tujih jezikih, s čimer sledimo tudi potrebam LBS za dvojezično leksikografijo.

2.2.6.2 Pomenski indikatorji pri stalnih besednih zvezah

Pomenski indikator pri stalnih besednih zvezah na kratko opredeli njihov pomenski obseg. Ker gre pogosto za terminološka ali polterminološka poimenovanja, ki zahtevajo strokovne razlage, z indikatorjem označimo zgolj področje rabe ali poiščemo ustrezno nadpomenko. S pregledom in analizo večjega števila indikatorjev pri stalnih besednih zvezah bo mogoče oblikovati nabor področij in poenotiti način pomenskih opredelitev v smislu: žival: domača žival, škodljivec, rastlina: poljščina, okrasna rastlina Na tej stopnji LBS ločimo pomenske indikatorje za označevanje:

(a) pripadnikov vrste, sorte ipd. Navadno znotraj določenega pomenskega polja (rastlina, žival ...)

arabski konj – *pasma*

zlata delišes – *sorta*

(b) poimenovanja predmetov, stvari, stanj, prepričanj, ustanov, kjer je indikator bodisi

(b1) širša nadpomenka:

iskanje skritega zaklada – *družabna igra*



konj z ročaji – *telovadno orodje*

nilski konj – *žival*

krompir v oblicah – *jed*

glasbeni stolp – *naprava*

ali (b2) ožja nadpomenka:

črni kos – *ptica*

zelena uš – *škodljivec*

(c) poimenovanja stvari, predmetov, pojavov ipd. na določenem strokovnem ali omejenem področju rabe. Indikatorji na kratko označijo področje. Ponekod je očitna potreba po ožji specializaciji, očitni pa je tudi prehod v splošno rabo, zlasti na področju javne uprave, gospodarstva, financ, mobilne telefonije, računalništva ipd.

otrok iz epruvete – *medicina*

otrok s posebnimi potrebami – *medicina*

zimski čas – *geografija*

glasbena skupina – *glasba*

kisli dež – *ekologija*

mesto med vratnicama – *šport: nogomet*

(d) parafraza ali opis:

knežje mesto – *Celje*

mesto pod Pohorjem – *Maribor*

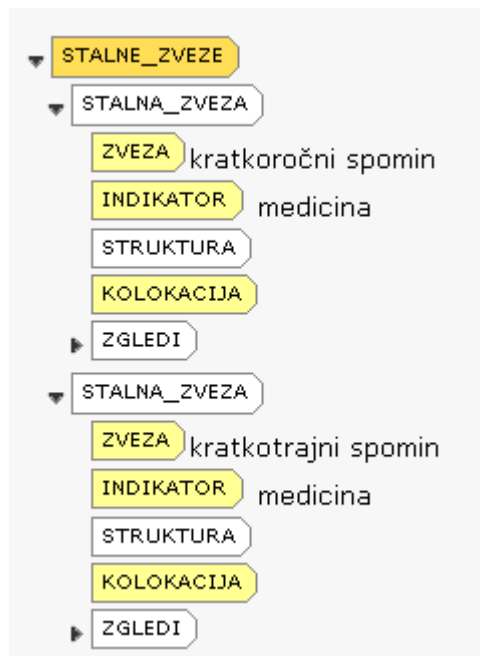
2.2.6.3 Zapisovanje stalnih besednih zvez

Pridevniške sestavine stalnih besednih zvez so pogostokrat vrstni pridevniki, njihov zapis na ni/ski ipd. pa je v rabi tipično nestabilen, kar pomeni, da ne sledi teoretičnim zahtevam razlikovalnega zapisovanja vrstnih in kakovostnih pridevnikov v njih, npr. *beli/bel kruh* : *bel pulover*. V LBS se ne želimo opredeljevati do ustreznosti zapisa pri vrstnih in kakovostnih pridevnikih, še

posebno zato ne, ker nas zanima dejanska raba in ne njena (pogostokrat tudi sporna) lingvistično-teoretična interpretacija. Menimo, da je bistvena osnova za vsako nadaljnjo analizo dejanska jezikovna raba, zato se v LBS osredotočamo na ugotavljanje pomena pridevnikov (vrstnost : kakovostnost) tudi v povezavi s pomensko samostojnostjo pridevnika ali pomenske nesamostojnosti pridevnikov kot sestavin stalne besedne zveze. V primerih, kjer katera od oblik prevladuje, npr. *glasbeni stolp*, se odločamo za prevladujočo, sicer zapisujemo obe: *glasbeni/glasben inštrument*.

Ker dojemamo stalne besedne zveze kot potencialne iztočnice (leme), je pomembno, da **variantne sestavine** zapisujemo ločeno, tj. vsako varianto kot enoto v elementu *zveza*, kot kaže spodnja slika.

Slika 9: Zapis variantnih sestavin stalne besedne zveze v programskem vmesniku



Iz istih razlogov, tj. zaradi obravnavanja stalnih besednih zvez kot možnih večbesednih LE, bomo njihove strukturne dele v drugi fazi izdelave LBS zapisovali kot leme.

Kot leksikalne enote imajo stalne besedne zveze pogosto tudi svoje tipično besedilno okolje, ki ga navajamo v obliki kolokacij, na primer: [izguba, pešanje] kratkoročnega spomina, [izguba] kratkotrajnega spomina, [težave] s kratkotrajnim spominom. Kolokabilnost stalnih besednih zvez je v nekaterih primerih neposredno razvidna iz besednih skic, zlasti pri samostalniških iztočnicah, pri pridevniških iztočnicah pa so navadno potrebna dodatna poizvedovanja. Ker predvidevamo v drugi fazi izdelave LBS tudi kodifikacijo stalnih besednih zvez, v prvi fazi ne iščemo posebej tudi kolokabilnega okolja zveze, ker bo to mogoče narediti avtomatično potem ko bodo stalne besedne zveze dejansko registrirane.

Tako kot pri iztočnicah oz. lemah tudi pri stalnih besednih zvezah (in FE) navajamo zglede rabe za vsak pomen. Ker dojemamo stalne besedne zveze kot potencialne leme, zapisujemo v zgledih v elementu >i/< celotno stalno besedno zvezo.

2.2.7 Pomenska shema – splošno

Pomensko shemo si v LBS predstavljamo kot *opis pomenskega scenarija*, ki ga oblikujemo na podlagi analize konkordančnega niza, v katerem za vsak pomen (in podpomen) obravnavane besede v iztočnici natančno opišemo:

- (skladenjske) udeležence (število ter njihovo pomenska vloga ali semantični tip) ter
- okoliščine, zlasti kraj in smer ter način, pa tudi krajevne in časovne okoliščine, če so za opis pomenskega scenarija oziroma za evociranje ustreznega pomena ali podpomena ključne. S tovrstnimi informacijami predstavlja pomenska shema povezavo med pomenskim in skladenjskim opisom LE.

Slika 10: Analiza konkordančnega niza z vidika udeleženske zgradbe za glagol *stisniti*



še rasti . Zamenu je strah , nekak ga je	stisnu	ven , tko ko gnoj iz mozola . Seveda ,	iztisniti snowtekočino iz česa - preneseno
vzamemo iz lonca in ju s pomočjo vilice	stisnemo	iz ovoja . Ker sta tako česen kot čebula	iztisniti snowtekočino iz česa v drugo stvar
na cedilu . Če pa sem ter tja v črpalko	stisnete	kapljico silikonske masti , bo delovanje	iztisniti snowtekočino iz česa v drugo stvar
segrevajte marelično marmelado , v katero ste	stisnili	malo limoninega soka . Nato zmes pretlačite	iztisniti snowtekočino iz česa v drugo stvar
varuje pred prehladom . V ne pretopel čaj	stisnimo	sok limone . Prav tako prija topla limonada	iztisniti snowtekočino iz česa v drugo stvar
opozori , da ti je nekdo na ramo pljunil ali	stisnil	šampon . Zadnji trik sem doživel	iztisniti snowtekočino iz česa v drugo stvar
prebavni cikel , tako da mu moram občasno	stisniti	v danko (z nastavkom za modolex) nekaj	iztisniti snowtekočino iz česa v drugo stvar
pusitjo v posebnih sušilnicah , potem pa se	stisnejo	, in ko kvasovke še ne predelajo vseh	izžeti kaj
jabolka je treba najprej umiti , razrezati in	stisniti	. Jabolčni sok nato pustimo stati v nepredušno	izžeti kaj
slami sušili grozdje in ga po nekaj mesecih	stisnili	. Letos so s sušenjem grozdja poskusili	izžeti kaj
je začeti dan s svežim sadnim sokom .	Stisnite	si limono , pomarančo , grenivko , melono	izžeti kaj

2.2.7.1 Postopek oblikovanja pomenske sheme

V izhodišču je naša naloga ugotoviti mehanizme, po katerih se govorniki slovenščine med seboj razumemo oz. mehanizme, na podlagi katerih smo sposobni razločevati med pomenom glagola *stisniti* v zgledih (a) od pomena glagola *stisniti* v zgledih (b).

- (a) *Otipal je ženino roko in **jo** močno **stisnil**.*
***Anja** pa ni Marjana niti enkrat **stisnila** za **ramena**.*
***Eleanor** je zdravnica **zelišča** in **pismo** trdno **stisnila** v **dlan**.*
***Cezar** z **dlanjo** **stisne ročaj** meča.*
- (b) ***Stisnite** si **limono , pomarančo , grenivko , melono , ananas** itn.*
*Iz **grozdja** so **stisnili** 70 litrov ledenega **vina**.*
*Počasi segrevajte marelično marmelado, v katero ste **stisnili** malo limoninega **soka**.*
***Krompir** olupite in še vročega **stisnite** skozi **stiskalnico** za krompir.*

Če primerjamo pomene ali pomenske tendence glagola *stisniti* v zgledih (a) in (b) zgoraj, lahko ugotovimo, da se med seboj razlikujejo oziroma, da imajo pomeni glagola v zgledih (a) nekatere skupne lastnosti, ki se ločujejo od pomenov glagola v zgledih (b), ki imajo prav tako nekatere skupne lastnosti. Razlike med pomenoma (a) in (b) želimo v pomenski shemi opisati na način, s katerim bomo dosegli najmanjši skupni imenovalac razumevanja konkretnega pomena oz. skupni imenovalac razlikovanja med obema pomenoma oz. med drugimi pomeni/podpomeni istega glagola. Za pomen glagola *stisniti* v zgledih (a) smo zapisali pomenski scenarij v obliki pomenske sheme:

ČLOVEK *stisne* **PREDMET** z **DLANJO** (tako, da ga trdno oprime).

Za pomen glagola v zgledih (b) pa takole:

ČLOVEK s **PRIPOMOČKOM** *stisne* **SADJE** | ali | *stisne* **TEKOČINO** iz **SADJA** | ali | *stisne* **SADJE** v **TEKOČINO**

Za besede, ki so v pomenski shemi napisane z velikimi črkami, predvidevamo, da so obvezni shemski elementi za konkretni pomen glagola, različno obarvani deli stavka v zgledih zgoraj pa ustrezajo shemskim elementom v pomenskih shemah. Primerjava med zgledi skupine (a) in (b) ter pripadajočima pomenskima shemama kaže, da

- so za izražanje pomena 'oprijeti' potrebni najmanj trije udeleženci, ki zasedajo vezljivostna mesta, kot jih napoveduje glagol v konkretnem pomenu, medtem ko glagol v pomenu 'iztisniti tekočino' napoveduje štiri vezljivostna mesta in njihovo specifično zasedbo.
- se **DELOVALNIK** v zgledih (b), tj. za pomen 'iztisniti tekočino', za razliko od pomena 'oprijeti' redko neposredno izraža, je pa vedno razviden iz osebne glagolske oblike;
- se v zgledih (a), tj. za pomen 'oprijeti', lahko opušča **SREDSTVO**, s katerim navadno (o)prijemamo;
- v zgledih (b), tj. za pomen 'iztisniti tekočino', navadno ni opuščen element, ki vsebuje **TEKOČINO** ali **SNOV**;
- se običajno, vendar ne nujno, za pomen 'iztisniti tekočino', v zgledih (b) navaja **PREDMET**, iz katerega se stiska tekočino-/snov.
- **SREDSTVO**, s katerim izžemamo, se v pomenu 'iztisniti tekočino' običajno izpušča.

2.2.7.2 Elementi pomenske sheme

Za opis pomenskega scenarija glagola *vzeti* v pomenu 'ukrasti' moramo v pomenski shemi, poleg glagola, ki ga, za razliko od pomenskih indikatorjev vedno navajamo v glagolski obliki, ki ustreza elementu na osebkovem mestu, izpostaviti tri obvezne elemente ali udeležence: **KDO**, ki je navadno človek in ga je mogoče še natančneje opisali kot storilca, vzame **KOMU**, ki je prav tako človek, lahko tudi institucija, ki mu/ji je bilo nekaj odvzeto in je v tem smislu prizadeti ali

oškodovanec, in **KAJ**, ki predstavlja odtujeni predmet, ki je navadno lastnina ali vrednost. Okoliščine, v konkretnem primeru **od kod**, so v pomenskem scenariju navadno prisotne kot opcijski, tj. neobvezni element pomenske sheme.

Zgled 16: Elementi pomenske sheme za pomen 'ukrasti' glagola *vzeti*



Kot rečeno, navajamo v opisu pomenskega scenarija skladijsko in pomensko nujne elemente pomenske sheme z velikimi črkami, s čimer nakažemo osnovni vezljivostni potencial konkretnega pomena – na tej stopnji se torej odločamo tudi o tem, ali so okoliščine imanentna lastnost konkretnega glagolskega pomena. Če je odgovor pritrđen, mora biti njihova obveznost razvidna iz pomenske sheme z zapisom z velikimi črkami, se pa zaenkrat (na podlagi sistema CPA) odločamo za omejen nabor pomensko relevantnih prislovnodoločilnih elementov PS. Kot take določamo samo lokacijo, smer, način in čas. Poimenovanje obveznih elementov pomenske sheme na tej stopnji izdelave leksikalne baze ni poenoteno niti ne obstaja zaključen seznam elementov oz. njihovih semantičnih tipov, npr. ČLOVEK – storilec, prizadeti; PREDMET – lastnina, vrednost itd. V začetnih fazah izdelave iztočnic v LBS opisujemo udeležence izključno za potrebe konkretnih pomenov, saj se je pri pregledu tujejezičnih leksikalnih baz pokazalo, da so taki zaključeni sezname za urednike prej ovira kot pomagalo. Gre namreč v prvi vrsti za to, da ne želimo opisov udeležencev prilagajati vnaprej predvidenim možnostim, ampak nasprotno, dopuščamo vse opise, za katere menimo, da ustrezajo celotnemu pomenskemu scenariju obravnavanega pomena. Omenili smo tudi, da je tako poenotenje mogoče v zaključni fazi, ko bodo za vse pomene in podpomene v LBS vključenih glagolskih iztočnic pomenske sheme izdelane.



Z analizo konkordančnega niza lahko ugotovimo, da so obvezni elementi, kot je razvidno tudi iz zgornjega zgleada, v stavku lahko izraženi ali pa ne. Pri tem so **(a) izraženi** elementi pomenske sheme lahko razvidni **(a.1) znotraj** ali **(a.2) zunaj stavčne zgradbe** ali pa so **(b) neizraženi**. Neizraženost obveznih elementov je povezana s stavčno strukturo, ki omogoča izpuščanje (npr. brezosebna raba, splošni vršilec ipd.) ali pa obstaja splošno vedenje o nujni prisotnosti sicer neizraženega elementa, npr.: *Tomaž je (hrano)*. Neizraženost elementov na osebkovem mestu v udeleženski strukturi je v slovenščini pogosto, kot izkazujejo spodnji zgledi, zajeta v osebni obliki glagola.

Kako se v konkretnih zgledih iz korpusa kaže izraženost/neizraženost obveznih elementov pri opisu scenarija v pomenski shemi lahko ponazorimo z opisom pomenskega scenarija za pomen 'iztisniti tekočino' glagola *stisniti* in 'dati darilo' glagola *podariti*, ki ju v pomenski shemi zapišemo kot:

ČLOVEK s **PRIPOMOČKOM** *stisne* **SADJE** | ali | *stisne* **TEKOČINO** iz **SADJA** | ali | *stisne* **SADJE v** **TEKOČINO**

(a) izraženi

(a.1) znotraj stavka: *Iz grozdja so na domači stiskalnici stisnili 70 litrov ledenega vina.*

(a.2) zunaj stavka: *Jabolka je treba najprej umiti, razrezati in stisniti. Jabolčni sok nato pustimo stati.*

(b) neizraženi (tekočina in sredstvo): *Stisnite si limono, pomarančo, grenivko, melono.*

ČLOVEK ali **INSTITUCIJA** *podari* drugemu **ČLOVEKU** ali **INSTITUCIJI DARILO** za konkretno priložnost ali z določenim namenom

(a.2) Bil je tudi cerkveni **ključar**. **Cerkvi** je *podaril* večji nabožni **kip**.



(b) Priložnosti, ob katerih nazdravimo s **penečim vinom**, je vedno več. Lahko ga tudi *podarimo*, toda katerega izbrati?

2.2.7.3 Razmerje med pomenskimi indikatorji in pomenskimi shemami – primer glagola *pasti*

Glagol *pasti* ima v LBS registriranih več pomenov. V spodnji primerjavi je opisano razmerje med pomenskimi indikatorji jedrnih pomenov in podpomenov in med pripadajočimi pomenskimi shemami.

Zgled 17: Razmerje med pomenskimi indikatorji in pomenskimi shemami pri glagolu *pasti*

pasti

1 premikati se od zgoraj proti navzdol - o predmetu

1.1 kaj izpade, odpade

1.2 izgubiti, izpustiti

1.3 kaj prileti na človeka

1.4 o padavinah: prileteti na zemljo

1.5 o svetlobi: pojaviti se na površini

1.6 o delih telesa: spustiti se, zavzeti položaj

1.7 o oblačilih: zavzeti obliko

1.8 o udarcih

1.9 o denarju: dobiti

1.10 o drevesu: posekati

2 izgubiti ravnotežje - o človeku

2.1 ne biti integralni del česa

2.2 ne biti več v določenem stanju, razpoloženju

PREDMET zaradi sile težnosti pade KAM

PREDMET pade s PROSTORA ali

PREDMETA, kjer je postavljen ali pritrjen

PREDMET pade ČLOVEKU KAM |ali|

MU pade od KOD

PREDMET pade ČLOVEKU na DEL

TELESA

KOLIKO PADAVIN pade

SVETLOBA ali SENCA pade na PROSTOR

ali ga v celoti prekrije

DEL TELESA pade ČLOVEKU na drugi

DEL TELESA ali v določen POLOŽAJ

OBLAČILO ali TKANINA pade tako, da

zavzame določen POLOŽAJ

UDARCI padejo po ČLOVEKU

DENAR pade ČLOVEKU

DREVO pade, ker ga posekajo ali požagajo

ČLOVEK ali ŽIVAL zaradi izgube

ravnotežja pade KAM |ali| pade

ČLOVEK ali PREDMET pade iz CELOTE

med seboj povezanih stvari

ČLOVEK pade iz STANJA



3 znižati se na lestvici - o vrednosti

3.1 o količini ali vrednosti

3.2 o intenzivnosti

3.3 o statusu ali kakovosti

3.4 o socialnem položaju: propasti

4 zgoditi, dogajati se

4.1 o stanju: doživljati, občutiti

4.2 o prostoru, situaciji: pojaviti se

4.3 o dejavnosti: ukvarjati se

4.4 o dogodku: nastopiti

4.5 o besedah: izreči

4.6 o neprijetnostih: doleteti

4.7 napasti

5 prenehati obstajati

5.1 umreti

6 podleči

6.1 zaradi vpliva

6.2 zaradi prevlade

KOLIČINA, VREDNOST ali PREDMET, ki ima neko vrednost, pade na lestvici za določeno KOLIČINO ali VREDNOST STANJE, v katerem je ČLOVEK ali INSTITUCIJA, pade tako, da izgubi intenzivnost ali učinkovitost ČLOVEK ali INSTITUCIJA |ali| STATUS, KAKOVOST ali POLOŽAJ človeka ali institucije pade na lestvici na določeno TOČKO ali VREDNOST ČLOVEK pade v moralnem ali socialnem smislu tako, da izgubi svoj status na družbeni lestvici

NEKAJ, kar pade, se zgodi ali začne obstajati ČLOVEK, ki pade v določeno STANJE, se v njem nahaja ali ga doživlja ČLOVEK, ki pade v PROSTOR, OKOLJE ali SITUACIJO, se nenadoma znajde tam ČLOVEK, ki pade v neko DEJAVNOST, se z njo ukvarja

DOGODEK ali ČAS dogodka pade na določen ČAS ko BESEDE padejo, so izrečene na ČLOVEKA ali INSTITUCIJO pade NEKAJ neprijetnega ČLOVEK pade po drugem ČLOVEKU

če ČLOVEK, INSTITUCIJA ali DOSEŽEK pade, preneha obstajati oz. izgubi svojo pomembnost ČLOVEK pade v BOJU |ali| pade zaradi uporabe OROŽJA ali SILE s strani drugega ČLOVEKA

če ČLOVEK pade zaradi lastnosti drugega ČLOVEKA ali PREDMETA, se jim ne more upreti, ker ga privlačijo ali nanj vplivajo ČLOVEK, UTRDBA, MESTO ali INSTITUCIJA pade pod oblast, nadzor drugega ČLOVEKA ali INSTITUCIJE

6.3 zaradi neznanja, nesposobnosti

če ČLOVEK ali INSTITUCIJA pade pri preverjanju ZNANJA ali USPOSOBLJENOSTI, je pri tem neuspešen

2.3 Skladenjski nivo

Opis skladenjskih informacij v zvezi s posamezno LE oz. njenim pomenom je v LBS namenjen:

- (a) beleženju **tipičnih skladenjskih struktur** in pri glagolu **skladenjskih vzorcev**, s katerimi se uresničuje pomen ali podpomen LE, ki jo obravnavamo v iztočnici;
- (b) beleženju **skladenjskih zvez**, ki predstavljajo razširjene in /ali strukturno ustaljene dele jezika;
- (c) formalizaciji **skladenjskih struktur** in **skladenjskih vzorcev za računalniško obdelavo naravnega jezika**, kjer posamezne dele struktur in vzorcev označimo s predvidenimi elementi in atributi ter
- (d) **beleženju slovničnih omejitev in preferenc**, ki izhajajo iz skladenjskega okolja LE in jih beležimo ob posameznih delih struktur in vzorcev.

Skladenjska informacija je v LBS vezana na dejstvo, da se pomen LE v stavku lahko uresničuje na različne skladenjske načine: neskončno število pomenov je mogoče izraziti s končnim številom skladenjskih kombinacij. V ta namen predvidevamo dva tipa skladenjskih informacij:

(a) **skladenjske skupine**, kjer beležimo:

- skladenjske strukture in
- skladenjske vzorce (pri glagolih), ter

(b) **skladenjske zveze**, ki obsegajo:

- kolokacijsko razširjene dele jezika, ki jih ni mogoče zajeti z „navadnimi“ kolokacijami, ker presegajo meje registriranih skladenjskih struktur;
- strukturno ustaljene in hkrati pogoste dele jezika, ki nimajo frazeološkega pomena oz. je pomen predvidljiv iz njihove sestavinske zgradbe.



2.3.1 Skladenjske skupine

Opis skladijskih struktur v nadelementu *skladenjske skupine* je v prvi vrsti namenjen potrebam računalniške obdelave, čemur je v celoti podrejen tudi njihov zapis. V leksikalnih podatkovnih bazah so tovrstni podatki navadno zapisani v obliki kod, v našem primeru v obliki atributov, ki jih pripisujemo posameznim delom skladijskih struktur ali skladijskih vzorcev. V ta namen je redaktorjem na voljo zaključen seznam atributov, ki ga vključuje programski vmesnik DPS. V primerih, kjer je skladijska informacija pomembna tudi za slovarske in slovnične aplikacije, ohranjamo tudi zapis v obliki skladijskega vzorca. Na primer, skladijsko strukturo za enega od pomenov glagola *pasti*: *sbz1 GBZ na sbz4*, katere zapis je namenjen računalniški berljivosti, je mogoče zapisati tudi v obliki skladijskega vzorca: *kdo pade na kaj*. Vsaj posredno so taki vzorci uporabni za posredovanje informacij v tipu slovarja, namenjenega učenju (tujega) jezika.

Zaradi možne vsaj delne avtomatizacije formaliziranja skladijskih struktur za namene računalniške obdelave, predvidevamo dve fazi dela pri polnjenju LBS s skladijskimi podatki in posledično dve vzoredni leksikalni bazi: LBS-A in LBS-B.

- **V 1. fazi** bomo izdelali **LBS-A**, kjer bomo zapisovali skladijske strukture, npr. *rbz GBZ; pbz1 SBZ1*, ki se potrjujejo s kolokacijami: *močno boleti, huda rana*, in skladijske vzorce, npr. *kdo pade na kaj*, ki ustreza skladijski strukturi: *sbz1 GBZ na sbz4*, v ustrezne elemente DTD-ja.
- **V 2. fazi** bomo izdelali **LBS-B**, v kateri bomo posameznim delom skladijskih struktur in skladijskih vzorcev pripisali ustrezne kode oz. attribute. Skladijske informacije v LBS-A (tj. ob koncu 1. faze) bodo tako vezane predvsem na slovarske in druge priročniške aplikacije, medtem ko bodo skladijske informacije ob zaključku druge faze – v LBS-B – tudi računalniško berljive. Zapis skladijskih struktur (in nanje vezanih skladijskih vzorcev) se deloma razlikuje pri posameznih besednih vrstah, čemur bomo v nadaljevanju namenili posebno pozornost.



2.3.1.1 Glagol

V polju **skladenjske skupine** predvidevamo pri vsakem pomenu in podpomenu glagolske iztočnice informacijo o **megastrukturi**, ki združuje vse skladenjske strukture, ki predstavljajo formalizirano podstavo **skladenjskih vzorcev**.

V 1. fazi beležimo pri glagolskih iztočnicah skladenjske vzorce, v katerih obravnavani glagol izraža določen pomen. Skladenjski vzorci predstavljajo razvezan zapis skladenjskih struktur, združenih v mergastruktur, in sicer v obliki (leksikografsko) berljivih vzorcev, npr. *kdo stisne koga za kaj*. Formalizacijo vzorcev z atributi, ki bo omogočala računalniško berljivost, kot tudi zapis megastruktur, bomo izdelali v 2. fazi LBS. Neposredno za skladenjskimi vzorci beležimo pri glagolu tudi (posebne) glagolske strukture, njihovo kodiranje v smislu računalniške berljivosti pa bomo prav tako izvedli v 2. fazi izdelave LBS.

2.3.1.1.1 Megastrukture

Megastrukture glagolskih iztočnic bomo naknadno avtomatično generirali iz zabeleženih skladenjskih vzorcev (2. faza – LBS-B), zato jih v programskem vmesniku v 1. fazi ne zapisujemo. V nadaljevanju si kljub temu ogledamo njihovo zgradbo, zapis in informacije, ki jih vsebujejo.

Namen megastruktur je predstaviti vse njene dele, tj. besede oz. besedne zveze, vključno s predlogi, vezniki in odvisnimi stavki znotraj stavčne zgradbe, ki so potrebni za izražanje konkretnega pomena glagola v iztočnici. Poleg obveznih delov so iz megastrukture razvidni tudi možni (opsijski) deli in deli, ki so medsebojno zamenljivi (alternativni deli strukture). Na primer, pomen glagola *pobirati* 'zaračunavati plačilo' se realizira z megastruktur:

sbz1 GBZ **sbz4** (za **sbz4**) (od **sbz2**)

ki jo beremo takole:

Samostalniku ali samostalniški besedni zvezi v imenovalniku – **sbz1** sledi glagol oz. glagolska besedna zveza (**GBZ**) *pobirati* in samostalnik oz. samostalniška besedna zveza v tožilniku – **sbz4**.

Tej strukturi lahko (opcija), kar označimo z (), sledi bodisi predlog **za**, ki veže samostalnik ali samostalniško besedno zvezo v tožilniku – **za sbz4**, bodisi predlog **od**, ki veže samostalnik ali samostalniško besedno zvezo v roditelju – **od sbz2**.

Postopek lahko pogledamo tudi z vidika obrnjenega zaporedja, kjer za glagol *stisniti* v pomenu: 'oprijeti z dlanjo' predvidevamo najprej ločene skladienske strukture, s katerimi se konkretni pomen tipično izraža, nato pa jih združimo v megastrukturo, ki poleg osnovne, tj. za izražanje konkretnega pomena minimalne strukture, izraža še opcijske dele (v oklepaju) in alternativne oz. med seboj zamenljive dele (ločene s poševnico).

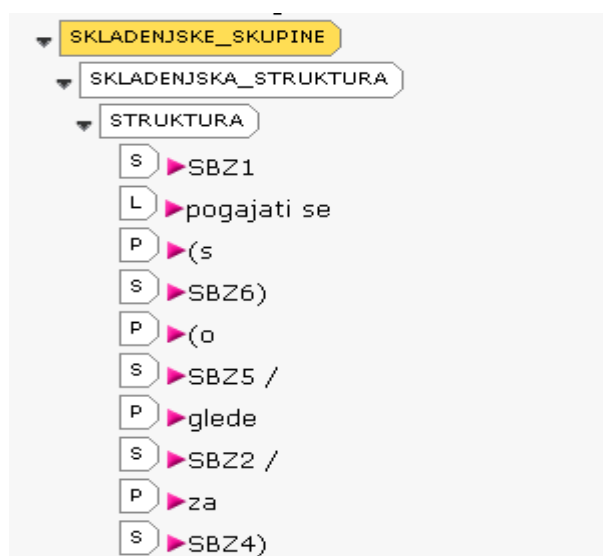
Zgled 18: Združevanje struktur v megastrukturo pri glagolu

sbz1 GBZ sbz4	}	sbz1 GBZ sbz4 (za sbz4 /v sbz4 / z/s sbz6)
sbz1 GBZ sbz4 za sbz4		
sbz1 GBZ sbz4 v sbz4		
sbz1 GBZ sbz4 z/s sbz6		

Če uporabimo način zapisa z skladienskimi vzorci, ki ga bomo podrobneje predstavili v nadaljevanju, beremo zgornjo megastrukturo takole: *kdo stisne kaj*, kar je izhodiščna minimalna struktura za izražanje tega pomena, lahko pa ji sledi še predložna zveza *za kaj* ali predložna zveza *v kaj* ali predložna zveza *z/s čim*. V konkretnem primeru se opcijske predložne zveze med seboj izključujejo, kar pomeni, da je mogoče reči: *kdo stisne kaj*, *kdo stisne kogal/kaj za kaj*, *kdo stisne kaj v kaj* in *kdo stisne kaj s čim*, ne pa tudi npr. *kdo stisne kogal/kaj za kaj v kaj* ali *kdo stisne kaj v kaj z/s čim*, pri čemer ne smemo pozabiti, da take omejitve veljajo samo za konkretni pomen glagola *stisniti*, tj. 'prijeti z roko'.

Predvidevamo, da bo zapis megastrukture, kot smo jo predstavili zgoraj, mogoče v 2. fazi avtomatično generirati iz kodiranih skladijskih vzorcev. Zapis megastrukture za glagol *pogajati se* v pomenu 'poskušati se dogovoriti' izgleda v programskem vmesniku tako, kot prikazuje slika.

Slika 11: Kodiran zapis megastrukture za glagol *pogajati se*



Pomembno: ko govorimo o glagolskih, samostalniških, pridevniških in prislovnih besednih zvezah, s katerimi zapisujemo posamezne dele struktur, imamo v mislih vse potencialne možnosti "zloženih" zvez. Predvidevamo torej, da so posamezni deli skladijskih struktur vedno lahko tudi sami na sebi skladijske zveze. Npr. skladijska struktura pbz1 sbz1: *gorsko jezero* je lahko notranje zložena kot v primeru: *kristalno čisto gorsko jezero* (= prisl+prid + prid+sam); struktura gbz na sbz5 (npr. *pasti na lestvici*) predvideva možnost zloženosti posameznih delov strukture kot v primeru: *drastično pasti na lestvici popularnosti*. Tak zapis smo, čeprav se zavedamo, da gre za sistemsko možnost, ki pa se v vseh primerih ne uresničuje sistemsko oz. so v nekaterih primerih take možnosti prej izjema kot pravilo (npr. *morati se znati počutiti nadvse srečnega*), sprejeli zaradi konsistentnosti, poenoteno kodifikacijo, pa zahteva tudi računalniški zapis.



2.3.1.1.2 Skladenjske strukture pri glagolih

V obliki skladenjskih struktur (in ne skladenjskih vzorcev) beležimo pri glagolskih iztočnicah:

(a) Glagolske zveze, ki bodisi ne izražajo glagolske vezljivosti v smislu pomensko nujnih delov stavka, ampak gre za modifikacije glagolske iztočnice oz. za notranjo strukturiranost glagolskih zvez bodisi za samostalniško in pridevniško vezljivost, kot v primerih: *(biti) pripravljen pomagati*, *(imeti) možnost pritožiti se* ipd.

Tabela 5: Skladenjske strukture pri glagolih – (a)

Vrsta	skladenjska struktura	kolokacijski zgled
glagolske	gbz Inf-GBZ	[znati] pasti , [nehati] boleti , [morati] paziti
	gbz Inf-gbz GBZ	[utegniti začet] snežiti
	pbz1 Inf-GBZ	[(biti) pripravljen] pomagati
samostalniške	sbz1 Inf-GBZ	[(imeti) možnost] pritožiti se
pridevniške	pbz1 GBZ	[vesel] razlagati
prislovne	rbz GBZ	[nesrečno] pasti , [neznosno] boleti
priredne	GBZ in gbz	pasti in se poškodovati
	gbz in GBZ	spotakniti se in pasti ; peči in boleti
	GBZ ali gbz	kupiti ali najeti

(b) Sem prištevamo tudi strukture z glagoli v pomožni funkciji in strukture z nedoločniki, kjer so določila pomensko nujni del obravnavanega glagola.

Tabela 6: Skladenjske strukture pri glagolih – (b)

Vrsta	Skladenjska struktura	Kolokacijski zgled
-------	-----------------------	--------------------

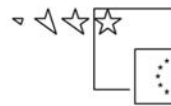


Glagolske		
- nedoločniške	GBZ Inf-gbz	uspeti [doseči], začeti [snežiti]
	GBZ inf-gbz gbz	utegniti [začeti snežiti]
	GBZ Inf-gbz sbz1	želeti biti [igralec]
	GBZ Inf-gbz pbz1	znati biti [nesramen]
- pomožniške	GBZ Inf-gbz rbz	začenjati postajati [nevarno]
	Vez-GBZ pbz1	izgledati [utrujen], izpasti [neumen], (biti) videti [vesel]
	Vez-GBZ pbz2	čutiti se [osamljenega]
	Vez-GBZ sbz1	izpasti [bedak], biti [lisica]

Glagole oz. glagolske besedne zveze zapisujemo v skladijskih strukturah drugače glede na to, ali gre za iztočnico (velike črke) ali ne (male črke). Poleg glagolskih zvez lahko tvorijo glagolske strukture tudi samostalniške, pridevniške in prislovne zveze, vendar tu samostalniki, pridevniki in prislovi nikoli ne nastopajo kot iztočnice, zato so v glagolskih strukturah vedno zapisane z malimi črkami. Vezniške dele struktur zapisujemo kot konkretne vezniške, v drugi fazi pa jih označimo z ustreznim elementom in atributom.

2.3.1.1.3 Skladijski vzorci

Medsebojno prepletanje skladijske in pomena pri glagolskih iztočnicah najboljše (poleg pomenskih shem) odražajo *skladijski vzorci*. Predstavljajo namreč tisti del LBS, ki daje informacije o vezljivosti glagola, obenem pa predstavlja posredniško fazo med skladijskimi strukturami (združenimi v megastrukturo) in njihovimi uresničitvami v konkretnem skladijskem vzorcu. Skladijski vzorec tako izraža vse tipične realizacije glagola v konkretnem pomenu v okviru stavka. To pomeni vse obvezne dele PS, razen, če se tipično ne opuščajo, in vse druge dele stavka, ki so za uresničevanja pomena tipični, zlasti okoliščine. Na primer, zapisu PS za glagol *roditi* v pomenu 'o rojstvu otroka': ŽENSKA rodi OTROKA ali OTROK se rodi KOMU, sledijo skladijski



vzorci, v katerih se ta pomen tipično realizira, zato izražajo tudi okoliščine, najpogosteje s predložnimi zvezami v prislovni funkciji. Konkretna zapolnitev predložnih zvez je zabeležena na kolokacijskem nivoju.

Zgled 19: Povezava med skladenjskimi vzorci in kolokacijami pri glagolu *roditi*

roditi – 'o rojstvu otroka'

PS: ŽENSKA rodi OTROKA ali OTROK se rodi KOMU

vzorec: kdo rodi koga

kolokacija: roditi [sina, hčerko, otroka, dvojčka, dvojčici, fantka, deklico, hčer, hčerkico, punčko]

vzorec: kdo rodi

kolokacija: [ženska, žena] rodi

vzorec: kdo se rodi

kolokacija: [otrok, dete, dojenček, deklica, novorojenček] se rodi

vzorec: kdo se rodi česa/sbz2 *KDAJ*

kolokacija: roditi se [aprila, maja]

vzorec: kdo rodi kaj/sbz4 *KDAJ*

kolokacija: roditi [vsako leto]

vzorec: kdo se rodi po čem *KDAJ*

kolokacija: roditi se po [vojni]

vzorec: kdo se rodi med čim *KDAJ*

kolokacija: roditi se med [vojno]

vzorec: kdo se rodi v čem *KJE*

kolokacija: roditi se v [Ljubljani, Chicagu, Nemčiji, Mariboru, Trstu, Celju, Argentini]

vzorec: kdo se rodi na čem *KJE*

kolokacija: roditi se na [Dunaju, Hrvaškem, Ptuju]

vzorec: kdo se rodi s čim

kolokacija: roditi se z/s [Downovim sindromom, okvaro, napako]

V primeru, da so okoliščine imanentni del konkretnega glagolskega pomena, so razvidne že na ravni pomenske sheme.

Zgled 20: Povezava med skladenjskimi vzorci in kolokacijami pri glagolu *pasti*

pasti – 'premikati se od zgoraj proti navzdol'

PS: PREDMET zaradi sile težnosti pade KAM

vzorec: kaj pade

kolokacija: [bomba, granata, drevo] pade

vzorec: kaj pade na kaj

kolokacija: pasti na [dno, Zemljo, zemljo, tla]

vzorec: kaj pade v kaj

kolokacija: pasti v [prepad, globino, morje]

vzorec: kaj pade kam

kolokacija: pasti [dol, globoko]



Zapis skladijskih vzorcev je v programskem vmesniku prilagojen leksikografskemu oz. uporabniškemu branju (*koga boli kaj*), hkrati pa bo s kodifikacijo posameznih delov vzorca s pomočjo atributov v 2. fazi izdelave LBS zagotovljena tudi računalniška berljivost in možnost procesiranja podatkov.

Podlaga vsakemu konkretnemu skladijskemu vzorcu je konkretna skladijska struktura. Skladijske strukture, ki smo jih ugotovili za pomen glagolov *stisniti* 'oprijeti z dlanjo' in *pobirati* 'zaračunavati plačilo' zgoraj, lahko prikažemo s pripadajočimi skladijskimi vzorci.

Zgled 21: Zapis skladijskih struktur z vzorci

skladijska struktura

(a) stisniti – 'oprijeti z dlanjo'

sbz1 GBZ **sbz4**
sbz1 GBZ **sbz4** za **sbz4**
sbz1 GBZ **sbz4** v **sbz4**
sbz1 GBZ **sbz4** z/s **sbz6**

skladijski vzorec

kdo stisne **kaj**
kdo stisne **kaj** za **kaj**
kdo stisne **kaj** v **kaj**
kdo stisne **kaj** z/s **čim**

(b) pobirati – 'zaračunavati plačilo'

sbz1 GBZ **sbz4**
sbz1 GBZ **sbz4** za **sbz4**
sbz1 GBZ **sbz4** od **sbz2**

kdo pobira **kaj**
kdo pobira **kaj** za **koga**
kdo pobira **kaj** od **koga**

Skladijske vzorce, kot tudi skladijske strukture, na katerih vzorci temeljijo, ugotavljamo iz pregleda konkordančnega niza in besednih skic. V vsaki od spodnjih konkordanc glagola *pasti* je mogoče pripisati skladijski vzorec.

Zgled 22: Zapis skladijskih vzorcev na podlagi konkordanc



kinkal naslonjen na oknico in skoraj bi **padel**, ko je na oni strani zaropotalo. Sam Bog
šestindvajsetega septembra 1996 je Kabul **padel** v roke talibanov, Nadžibulaha so obesili
natančno veliki in ravno prav težki, da ne **padejo** takoj na zemljo, temveč jih človek še
osvetljevala gozd stotih volkov: tedaj bodo idoli **padli** in kri bogatih bo oplaknila hiše siromakov
romale iz Aten. Med vsemi je Schliemannu **padla** v oči 16 - letna šolarka Sofija Engastromenos
se je sesedla pisarna, kjer je dremal, in **padla** pet etaž globoko. Njegovo truplo je pokrivala
šinilo v glavo. *</p><p>* Odprl je vrata in **padel** v objem otrokom in ženi ... *</p><p>* Sporočilo
kritična meseca sta bila junij, ko je obisk **padel** na le 62 odstotkov, in avgust, ko so v
se je nekoliko znižal tečaj. Tečaji so **padli**, v povprečju za pol odstotka, tudi drugim
Počutje je v vzhodnem delu države septembra **padlo** s 105,1 na 103,5 indeksne točke, na zahodni
avgustovskimi 103,7. Pogoji poslovanja so s 87,3 **padli** na 87,0. Objava nemških uvoznih cen z
presežek je v drugem četrtletju letos **padel** za 13,6 odstotka na isti ravni predvsem

- 1 KDO pade
- 2 KAJ pade v roke KOMU (FE)
- 3 KAJ pade na KAJ
- 4 KAJ pade
- 5 KDO pade KOMU v oči (FE)
- 6 KAJ pade KAM
- 7 KDO pade KOMU v objem (?FE)
- 8 KAJ pade na KAJ
- 9 KAJ pade
- 10 KAJ pade s ČESA na KAJ
- 11 KAJ pade s ČESA na KAJ
- 12 KAJ pade za KAJ

Iz zapisov skladijskih vzorcev na podlagi konkordančnega niza je razvidno, da se nekateri vzorci ponavljajo, ker bodisi izražajo isti pomen (prim. vzorca 10 in 11) ali pa se različni pomeni izražajo z istimi skladijskimi vzorci (prim. vzorca 4 in 9 in vzorca 3 in 8).

2.3.1.1.3.1 Zapisovanje skladijskih vzorcev

Zapisovanje besednozveznih delov skladijskih vzorcev z zaimki (sbz1 = kdo/kaj; sbz2 = koga/česa; sbz3 = komu/čemu; sbz4 = koga/kaj itd.) zahteva izražanje nekaterih dodatnih slovničnih informacij, npr. zaimka KDO in KAJ razločujeta podspol človeškosti. V primeru, da kategorija podspola človeškosti ni pomensko razločevalna, navajamo oba zaimka in ju ločimo s poševnico. Zapis konkretnega predloga in veznika v skladijskem vzorcu nam omogoča



naknadno analizo vseh skladijskih vzorcev, v katerih se pojavlja določen predlog ali veznik. Na ta način predvidevamo, da bo mogoče ugotavljati pomenske vloge in skladijske funkcije predložitnih zvez. Variantne veznike in predloge, ki nastopajo v strukturah, kot na primer: *prek/preko; okrog/okoli, zunaj/izven* ipd. zapisujemo ločeno v posameznih vzorcih; variantni zapis ohranjamo samo pri predlogih *z/s, k/h* in *prek/preko*.

Pomembno: Pri zapisu skladijskih vzorcev predložitnih samostalniških besednih zvez v prislovni funkciji ne zapisujemo glede na njihovo prislovnodoločilno vlogo, ampak glede na njihovo besednozvezno zgradbo. Skladijskega vzorca za 3 primer zgoraj: ... *ne padejo takoj na Zemljo* ne zapišemo kot: *kaj pade kam*, ampak kot: *kaj pade na kaj*. Prislovne zaimke *kam, kako* ipd. uporabljamo samo v primeru, ko se na tem mestu (tj. na mestu prislovnega določila) pojavlja prislov ali prislovna besedna zveza. Npr. za zgled: ... *in kaj, če helikopter pade dol* in za zgornji zgled 6: ... *in padla pet etaž globoko*, zapišemo skladijski vzorec: *kaj pade kam*.

Poskusno v fazi izdelave LBS-1 LBS-A uvajamo tudi zapis funkcije v vzorcih zlasti pri tistih (predložitnih) besednih zvezah, kjer želimo poudariti, da določen del vzorca nastopa v konkretni prislovnodoločilni vlogi, npr. *kaj pade na kaj *KAM* kraj, kdo leti v kaj *KAM/smer**. Zaenkrat se zdijo taka predvsem krajevna prislovna določila (lokacija in smer), npr. *leteti nad oblaki; leteti čez ocean*, in ponekod načinovna in časovna prislovna določila, npr. *ravnati s previdnostjo – 'previdno': kdo ravna s čim *KAKO/način**.

2.3.1.1.3.2 Razmerje med skladijskimi vzorci in pomensko shemo

Skladijski vzorci tako kot pomenske sheme (od tu dalje PS), predstavljajo stično točko med pomenskimi in skladijskimi informacijami v LBS. Ker PS obvezno izraža vse udeležence, ki so nujni za izražanje vsakega konkretnega glagolskega pomena in podpomena – potrdilo se je, da spremembe v številu in vlogi udeležencev, ki jih v stavčnem okviru predvideva glagol, vplivajo na spremembe glagolskega pomena – je nujno, da skladijski vzorci vedno izražajo tudi v pomenski shemi izražene obvezne elemente, npr. za glagol *boleti* v pomenu 'čutiti fizično bolečino' velja PS:



ČLOVEKA ali ŽIVAL boli DEL TELESA, skladijski vzorca: *koga boli kaj* in *koga boli v čem* izražata vse obvezne dele pomenske sheme. Do odstopanj prihaja, ko se v besedilu kateri od obveznih elementov običajno izpušča, npr. v zgledu: *stisnite limono* (tj. sok iz limone – PS: ČLOVEK s PRIPOMOČKOM stisne SADJE |ali| stisne TEKOČINO iz SADJA |ali| stisne SADJE v TEKOČINO), zapis skladijskega vzorca: *kdo stisne kaj* ne izraža elementa TEKOČINA. Razlika med elementi pomenske sheme in skladijskim vzorcem nastaja tudi takrat, kadar s skladijskim vzorcem kažemo tipična vezljivostna mesta, ki pa niso nujna za izražanje konkretnega pomena, zaradi česar jih tudi pomenska shema ne izkazuje. Npr. pomenska shema za glagol *podariti* vsebuje tri obvezne shemske elemente: KDO podari KOMU KAJ, raba pa pogostokrat izkazuje tudi namen, ki se izraža v skladijskem vzorcu: *kdo podari komu kaj za kaj*. Iz tega sledi, da skladijski vzorci izražajo vse nujne elemente pomenske sheme (razen pri opuščanju in elipsi) in da hkrati lahko izražajo tudi elemente, ki se v stavčni strukturi preučevanega glagola ne kažejo kot pomensko nujni, ampak kot tipični – gl. tudi spodaj.

Pomembno: V PS ne beležimo splošnih okoliščin oz. tistih delov stavka, ki za izražanje pomena konkretnega glagola niso pomensko relevantni, z velikimi črkami, čeprav jih omenjamo v pomenskem scenariju, na primer okoliščin (kje) pri glagolu *pobirati* – 'prostiti za denar': ČLOVEK s prošnjo pobira DENAR od DRUGIH LJUDI in na ta način služi, ali pri glagolu *igrati* v pomenu 'izvajati glasbo pred skupino ljudi': ČLOVEK igra na DOGODKU. V omenjenih primerih gre za (sistemsko) lastnost glagola: dejanje, stanje ipd., ki ga izraža glagol, je vedno mogoče opredeliti glede na čas, kraj, način ipd. Če so okoliščine in način njihovega izražanja – ob sicer pomenski neobveznosti – vendarle tipični, kar je pogosto praksa, so razvidni iz skladijskih vzorcev, ki so v prvi vrsti odraz realizacije konkretnega pomena v stavku, npr. *kdo igra na čem* (na festivalu, proslavi, odprtju ...), *kdo igra po čem* (po festivalih, veselicah, hotelih).



2.3.1.2. Samostalnik

S skladijskimi informacijami v elementu **skladijske skupine** želimo tudi pri samostalniških LE beležiti načine, po katerih se pomeni samostalniških LE uresničujejo v okviru stavčne zgradbe. Znotraj skladijskih struktur zato beležimo:

- strukture, kjer je samostalnik oz. samostalniška besedna zveza jedro zveze – t. i. **jedrne skladijske strukture** in
- strukture, v katerih se tipično pojavlja preučevani samostalnik ne glede na skladijsko vlogo v njej. Za namene konkretnih specifikacij smo jih poimenovali **nejedrne skladijske strukture**.

Pomembno: Tudi pri samostalnikih beležimo skladijske strukture v obliki linearnih kombinacij besednozveznih delov (tj. samostalniških, pridevniških, glagolskih in prislovnih besednih zvez) in ne glede na njihovo stavčno funkcijo (osebek, predmet, prislovno določilo). Z ločevanjem na jedrne in nejedrne strukture pa vzpostavljamo razliko med notranjo zgradbo samostalniške besedne zveze, pri čemer je tudi ta lahko sama na sebi zložena (npr: *kristalno čisto gorsko jezero* = pbz SBZ1), in tipičnim skladijskim okoljem samostalniške besedne zveze. Nejedrne skladijske strukture so navadno tipične tako za pomen samostalnika ali samostalniške besedne zveze in za njeno ustrezno rabo v besedilu, zato so bolj zanimive za slovarske opise in učenje jezika.

Vse skladijske strukture, ki jih beležimo v nadelementu skladijske skupine, bodo v drugi fazi izdelave LBS označene z elementi in atributi, tako da jih bo mogoče uporabiti tudi v aplikacijah za računalniško obdelavo naravnega jezika.

2.3.1.2.1 Jedrne samostalniške strukture

Skladijske strukture, kjer samostalnik ali samostalniška besedna zveza nastopa kot skladijsko jedro, so dveh vrst:



(a) Strukture s samostalnikom/samostalniško besedno zvezo kot iztočnico, ki jo določa pridevniška, samostalniška, prislovna, glagolska besedna zveza ali odvisni stavek. Posebej navajamo samostalnike/besedne zveze s količinskimi modifikacijami.

Tabela 7: Jedrne strukture pri samostalnikih – (a)

Skladenjska struktura	Kolokacijski zgled
pbz1 SBZ1	[boleč, lep] spomin
pbz2 SBZ2	[šibkega] zdravja
SBZ1 Vez-gbz pbz	spomin je [boleč, živ]
SBZ1 Inf-gbz	možnost [pritožiti se]
sbz1 SBZ1	[tenis, golf] igrišče
SBZ1 sbz2	iskanje [zaposlitve]
SBZ2 sbz2	meseca [aprila]
SBZ1 rbz	potovanje [domov]
Kol-rbz SBZ2	[nekaj, precej] časa
Kol-pbz SBZ4	[vsako, prejšnje] leto , [tri] metre , [dva] milijona
Kol-pbz SBZ2	[deset, petnajst] milijonov , [pet] metrov , [22., drugega] aprila
SBZ1 Kol-pbz2	petek [trinajstega]
SBZ1 Odv-ali	vprašanje , ali...
SBZ1 Odv-ko	čas , ko...
SBZ1 Odv-da	slutnja , da ...

(b) Strukture s predložno samostalniško besedno zvezo, ki določa samostalnik/samostalniško besedno zvezo, ki vsebuje iztočnico.



Tabela 8: Jedrne strukture pri samostalnikih – (b)

Skladenjska struktura	Kolokacijski zgled
SBZ1 blizu sbz2	vir blizu [ministrstva]
SBZ1 brez sbz2	otrok brez [staršev]
SBZ1 čez sbz4	most čez [reko]
SBZ1 do sbz2	čas do [volitev]
SBZ1 glede sbz2	pričakovanja glede [zaposlovanja]
SBZ1 iz sbz2	marmelada iz [pomaranč]
SBZ1 izven sbz2	cesta izven [naselja]
SBZ1 k sbz3	pripombe k [osnutku]
SBZ1 med sbz6	razlika med [spoloma]
SBZ1 mimo sbz2	financiranje mimo [zakona]
SBZ1 na sbz4	spomin na [otročstvo, mladost]
SBZ1 nad sbz6	skrbništvo nad [otrokom]
SBZ1 o sbz5	film o [ljubezni]
SBZ1 ob sbz5	incident ob [meji]
SBZ1 od sbz2	čas od [sprejetja]
SBZ1 okoli sbz2	ograja okoli [hiše]
SBZ1 okrog sbz2	razprava okrog [zakona]
SBZ1 po sbz5	potreba po [varnosti]
SBZ1 pod sbz6	rezultat pod [pričakovanji]
SBZ1 pred sbz6	čas pred [volitvami]
SBZ1 prek sbz2	nakup prek [interneta, spleta]
SBZ1 preko sbz2	nakup preko [interneta, spleta]
SBZ1 pri sbz5	alergija pri [otročih, odraslih]
SBZ1 proti sbz3	boj proti [korupciji]



SBZ1 skozi sbz4	potovanje skozi [čas]
SBZ1 sredi sbz2	otok sredi [morja]
SBZ1 v sbz5	rana v [naravi, okolju]
SBZ1 v sbz4	metanje v [zrak]
SBZ1 z sbz6	čaj z [limono, mlekom]
SBZ1 za sbz4	čas za [premislek, igro]
SBZ1 zaradi sbz2	smrt zaradi [mamil]
SBZ1 znotraj sbz2	usklajevanje znotraj [ministrstva]
SBZ1 zoper sbz4	ovadba zoper [zdravnika]
SBZ1 zunaj sbz2	cesta zunaj [naselja]

2.3.1.2.2 Nejedrne samostalniške strukture

Kot nejedrne obravnavamo tiste samostalniške skladijske strukture, v katerih nastopa samostalniška iztočnica kot del glagolske ali nejedrni del samostalniške besedne zveze. Čeprav obravnavani samostalniki v teh strukturah ne predstavljajo skladijskih organizatorjev besedne zveze in jih posledično te zveze tudi pomensko ne določajo, jih beležimo zato, ker predstavljajo tipično besedilno okolje obravnavanega samostalnika. Zaradi tega so tako parsersko (tj. računalniško) kot leksikološko zanimive. Samostalniške strukture tega tipa lahko razdelimo v tri skupine:

(a) Strukture, v katerih je samostalniška iztočnica bodisi osebek, predmetno dopolnilo ali povedkovo določilo.

Tabela 9: Nejedrne strukture pri samostalnikih – (a)

Skladijska struktura	Kolokacijski zgled
SBZ1 gbz	spomini [oživijo]



gbz SBZ4	[zapravljati, izgubljeni] čas
Vez-gbz SBZ1	[izpasti] bedak , [biti] lisica
gbz Inf-gbz SBZ1	[utegniti postati] strokovnjak
gbz SBZ3	[predajati se] spominom
gbz SBZ2	[manjkati] denarja
Neg-gbz SBZ2	[ne izgubljeni, ne zapravljati] časa
SBZ1 sbz1	tenis [igrišče, lopar]
sbz2 SBZ2	[meseca] aprila
sbz1 SBZ2	[iskanje] zaposlitve
pbz1 SBZ2	[poln] denarja
pbz1 SBZ3	[nevaren] državi
pbz1 SBZ4	[dolžen] vsoto

(b) Strukture z glagolom/besedno zvezo v kombinaciji s predložno samostalniško besedno zvezo (konkretni predlog in samostalnik v ustreznem sklonu), ki vsebuje samostalnik iz iztočnice. Funkcijsko gledano, nastopajo predložne zveze v vlogi predložnih dopolnil, npr. *[ustaviti se] pred semaforjem*, ali prislovnih določil, npr. *[zadrževati se] zunaj hiše*, *[govoriti] izven konteksta*.

Tabela 10: Nejedrne strukture pri samostalnikih – (b)

Skladenjska struktura	Kolokacijski zgled
gbz blizu SBZ2	[stati] blizu odra
gbz brez SBZ2	[shajati] brez denarja
gbz čez SBZ4	[stopiti] čez lužo
gbz do SBZ2	[stopiti] do soseda
gbz glede SBZ2	[odločati se] glede barve
gbz iz SBZ2	[teči, curljati] iz rane



gbz izven SBZ2	[govoriti] izven konteksta
gbz k SBZ3	[privezati] k postelji
gbz med SBZ6	[pomešati se] med gledalce
gbz mimo SBZ2	[peljati se] mimo trgovine
gbz na SBZ4	[pomisliti] na otročstvo
gbz nad SBZ6	[planiti] nad sovražnika
gbz o SBZ5	[spregovoriti] o nasilju
gbz ob SBZ5	[nasloniti se] ob steno
gbz od SBZ2	[kričati] od bolečin
gbz okoli SBZ2	[teči] okoli stadiona
gbz okrog SBZ2	[teči] okrog stadiona
gbz po SBZ5	[politi se, popackati se] po obleki
gbz pod SBZ6	[zlesti] pod odejo
gbz pred SBZ6	[ustaviti se] pred semaforjem
gbz prek SBZ2	[pogledati] prek ograje
gbz preko SBZ2	[stopiti] preko potoka
gbz pri SBZ5	[ostati] pri življenju
gbz proti SBZ3	[glasovati] proti zakonu
gbz skozi SBZ4	[pogledati] skozi okno
gbz sredi SBZ2	[zbuditi se] sredi noči
gbz v SBZ5	[ostati] v spominu
gbz v SBZ4	[vtisniti se] v spomin
gbz z SBZ6	[praznovati] s sorodniki
gbz za SBZ4	[glasovati] za zakon
gbz zaradi SBZ2	[umreti] zaradi rane
gbz znotraj SBZ2	[nahajati se] znotraj obzidja
gbz zoper SBZ4	[pritožiti se] zoper sklep



gbz zunaj SBZ2	[zadrževati se] zunaj hiše
-----------------------	-----------------------------------

(c) Strukture s samostalnikom/besedno zvezo in predložno samostalniško besedno zvezo, ki vsebuje samostalnik iz iztočnice

Tabela 11: Nejedrne strukture pri samostalnikih – (c)

Skladenjska struktura	Kolokacijski zgled
sbz1 blizu SBZ2	[vir] blizu ministrstva
sbz1 brez SBZ2	[otrok] brez staršev
sbz1 čez SBZ4	[most] čez reko
sbz1 do SBZ2	[čas] do volitev
sbz1 glede SBZ2	[pričakovanja] glede zaposlovanja
sbz1 iz SBZ2	[krvavitev] iz rane
sbz1 izven SBZ2	[cesta] izven naselja
sbz1 k SBZ3	[pripombe] k osnutku
sbz1 med SBZ6	[razlika] med spoloma
sbz1 mimo SBZ2	[financiranje] mimo zakona
sbz1 na SBZ4	[spomin] na otročstvo
sbz1 nad SBZ6	[skrbništvo] nad otrokom
sbz1 o SBZ5	[film] o ljubezni
sbz1 ob SBZ5	[incident] ob meji
sbz1 od SBZ2	[čas] od sprejetja
sbz1 okoli SBZ2	[ograja] okoli hiše
sbz1 okrog SBZ2	[razprava] okrog zakona
sbz1 po SBZ5	[potreba] po varnosti
sbz1 pod SBZ6	[rezultat] pod pričakovanji



sbz1 pred SBZ6	[čas] pred volitvami
sbz1 prek SBZ2	[nakup, prodaja] prek spleta
sbz1 preko SBZ2	[nakup, prodaja] preko interneta
sbz1 pri SBZ5	[alergija, motnja] pri otrocih
sbz1 proti SBZ3	[boj] proti korupciji
sbz1 skozi SBZ4	[potovanje, sprehod] skozi čas
sbz1 sredi SBZ2	[otok] sredi morja
sbz1 v SBZ5	[rana] v okolju
sbz1 v SBZ4	[metanje] v zrak
sbz1 z SBZ6	[čaj, kava] z mlekom
sbz1 za SBZ4	[čas, minuta] za premislek
sbz1 zaradi SBZ2	[smrt] zaradi mamil
sbz1 znotraj SBZ2	[usklajevanje, spor] znotraj ministrstva
sbz1 zoper SBZ4	[ovadba, pritožba] zoper zdravnika
sbz1 zunaj SBZ2	[cesta] zunaj naselja

2.3.1.2.3 Zapisovanje samostalniških struktur

Zapis skladenjskih struktur pri samostalniških iztočnicah (1. faza: LBS-A) sledi logiki zapisa glagolskih skladenjskih struktur, kar pomeni, da v vseh strukturah enotno zapisujemo besednozvezne dele struktur, in sicer: samostalnike/samostalniške besedne zveze v ustreznem sklonu: sbz1-6/SBZ1-6, glagole/glagolske besedne zveze: gbz/GBZ; pridevnike/pridevniške besedne zveze: pbz/PBZ in prislove/prislovne besedne zveze: rbz/RBZ. Z zapisom z velikimi črkami opozarjamo, da konkretna beseda/besedna zveza v strukturi nastopa kot iztočnica. Predložne besedne zveze zapisujemo glede na njihovo notranjo zgradbo, pri čemer je zapis predloga vezan na konkretni predlog, ki se v strukturi pojavlja, sledi pa mu beseda/besedna zveza v določenem sklonu, npr. čas za počitek – SBZ1 za sbz4.



V programskem vmesniku zapisujemo najprej jedrne in nato nejedrne skladijske strukture. Med nejedrnimi najprej navajamo skladijske strukture, ki vključujejo samostalnik/samostalniško besedno zvezo kot iztočnico, ki jo določa pridevniška, samostalniška, prislovna, glagolska besedna zveza ali odvisni stavek. Sledijo strukture s samostalnikom/samostalniško besedno zvezo kot iztočnico, ki jo določa predložna samostalniška besedna zveza. Pri nejedrnih skladijskih strukturah navajamo zaporedoma glagolske besedne zveze, v katerih je samostalniška iztočnica bodisi osebek ali predmetno dopolnilo in nepredložne samostalniške besedne zveze. Sledijo strukture z glagolom/glagolsko besedno zvezo ali samostalnikom/samostalniško besedno zvezo v kombinaciji s predložno samostalniško besedno zvezo, ki vsebuje samostalnik iz iztočnice.

2.3.1.3 Pridevnik

Pridevniške skladijske strukture smo tako kot pri glagolih in samostalnikih razdelili na strukture, v katerih pridevnik ali pridevniška besedna zveza nastopa kot jedro besedne zveze in na take, v katerih je pridevniška iztočnica nejedrni del glagolske ali samostalniške besedne zveze.

2.3.1.3.1 Jedrne pridevniške strukture

S skladijskimi strukturami, kjer pridevnik ali pridevniška besedna zveza nastopa kot skladijsko jedro, beležimo pridevniške besedne zveze. Ločimo dva tipa:

(a) Strukture, kjer pridevnik/pridevniška besedna zveza, ki vsebuje iztočnico, določa samostalnik/samostalniško besedno zvezo v različnih sklonih, prislov/prislovno besedno zvezo, glagol/glagolsko besedno zvezo ali odvisni stavek.

Tabela 12: Jedrne strukture pri pridevnikih – (a)

Skladijska struktura	Kolokacijski zgled
*PBZ2 sbz2	bledega [obraza], [spoštovanja] vreden
*PBZ1 sbz2	poln [denarja]



PBZ1 sbz3	zvest [prijatelju], nevaren [državi]
PBZ1 sbz4	vreden [milijardo], dolžen [vsoto]
rbz PBZ1	[kristalno] čist , [neznansko] vesel
Kol-sbz PBZ1	[meter] dolg , [uro] trajajoč
PBZ1 Inf-gbz	pripravljen [oditi], sposoben [izpeljati]
PBZ1 Odv-da ...	vesel , da je kaj

***Pomembno:** Za pridevniške strukture PBZ2 sbz2 in PBZ1 sbz2 je značilno, da predvidevajo vsaj tri tipe pričakovanega besedilnega okolja, ki določa njihove tipične stavčnočlenske vloge: nastopajo kot povedkovo določilo, tj. ob glagolu *biti*: *je spoštovanja vreden*, ob polnopomenskem glagolu opravljajo vlogo povedkovega prilastka: *sprejel jih je odprtih rok*; ob samostalniku pa vlogo desnega prilastka, npr. *boginja domačega ognjišča*; *leta neuspešnega vztrajanja*. Ob tem je očitno, da se vse tri systemske možnosti v zvezi s posamezno strukturo ne uresničujejo systemsko, tj. nekatere tipično ja, druge sploh ne. Ker želimo take realizacije, kadar so očitne, prav tako beležiti, premaknemo informacijo še na raven razširjene skladišne zveze, npr.

Zgled 23: Razširitev pridevniških struktur z elementi besedilnega okolja

skladišne skupine

struktura: PBZ sbz2

kolokacija: vesele [narave, značaja]

kolokacija: veselega [obraza, srca]

skladišne zveze

zveza: vesele narave/značaja

vzorec: biti vesele narave/značaja

vzorec: narediti kaj vesele narave/značaja

vzorec: kdo vesele narave/značaja

skladišne zveze

zveza: veselega obraza/srca

vzorec: biti veselega obraza/srca



vzorec: narediti kaj veselega obraza/srca

vzorec: kdo veselega obraza/srca

Za samostalniški del struktur PBZ1 sbz2 in PBZ2 sbz2 je tudi značilno da gre pogostokrat za strukturno in pomensko ustaljene samostalniške besedne zveze z metaforičnim oz. prenesenim pomenom, kar jih uvršča med frazeološke zveze, npr.: *odprtih rok* (velikodušen), *dolgega nosu* (razočaran) ipd. Pogosti so tudi primeri, kjer ob strukturni ustaljenosti ni mogoče govoriti tudi o frazeološkem pomenu, kar je navadno znak, da zvezo obravnavamo med skladenjskimi zvezami, npr. *dobrega srca*, *kislega obraza* (*s kislim obrazom*) ipd. – zadnji primer pod pomenskim indikatorjem 'čustvo'.

(b) Strukture, kjer pridevnik/pridevniška besedna zveza skupaj s predlogom določa (modificira) samostalnik ali samostalniško zvezo.

Tabela 13: Jedrne strukture pri pridevnikih – (b)

Skladenjska struktura	Kolokacijski zgled
PBZ1 sbz2	negotov glede [meril]; enoten glede [sodelovanja]
PBZ1 do sbz2	ravnodušen do [odzivov]; radodaren do [otrok]; sovražen do [ZDA]; aboten do [konca]
PBZ1 iz sbz2	narejen iz [lesa, školjk, sladkorja]
PBZ1 med sbz6	priljubljen med [ljudmi]
PBZ1 na sbz4	pogolten na [denar]; prijeten na [pogled];
PBZ1 od sbz2	zaripel od [napora]; bled od [jeze]; odrezan od [sveta]
	razburljivejši od [bojev] trši od [betona], manjši od [svinčnika]
PBZ1 ob sbz5	postavljen ob [zapori]; pritisnjen ob [zid]
PBZ1 k sbz3	nagnjen k [samomorom]
PBZ1 kot sbz1	omenjan kot [kandidat]



PBZ1 okrog sbz2	zavezan okrog [pasu]
PBZ1 po sbz5	krojen po [meri]; razpoznaven po [cvetovih]; arabski po [veri]; urejen po [barvi]
PBZ1 pod sbz5	znan pod [nazivom]
PBZ1 pred sbz6	izdelan pred [letom 1955]; enak pred [zakonom]
PBZ1 prek sbz2	pridobljen prek [interneta]
PBZ1 pri sbz5	glaven pri [prodaji]; priljubljen pri [ljudeh]
PBZ1 proti sbz3	odporen proti [vodi]
PBZ1 spričo sbz2	slep spričo [dogodkov]
PBZ1 v sbz5	združen v [šop]; dober v [šoli]; napisan v [arabščini]
PBZ1 z/s sbz6	določen z [zakonom]; bogat s [cvetovi]; zavarovan s [poroštvom]; zasvojen z [nakupovanjem]; pokrit s [slamo]; sprt z [logiko]
PBZ1 za sbz4	dosegljiv za [komentar]; namenjen za [televizijo]; smrtonosen za [ljudi]; bogatejši za [izkušnjo]; motiviran za [igro]; krščena za [Katarino]; idealna za [ravnateljico]
PBZ1 zaradi sbz6	užaljen zaradi [očitkov]

2.3.1.3.2 Nejedrne pridevniške strukture

Kot nejedrne obravnavamo tiste pridevniške skladienske strukture, v katerih je pridevniška iztočnica nejedrni del glagolske ali samostalniške besedne zveze. Pri glagolskih zvezah posebej ločimo zveze s pomožnim glagolom *biti*, kjer nastopa pridevnik v vlogi povedkovega določila, pri čemer prištevamo sem tudi vse glagole, ki nastopajo v pomožniški vlogi, npr. *čutiti se*, *izgledati*, *zdeti se*, *postati*, *ostati* ipd. *vesel*; *počutiti se samozavestnega*, in zveze s polnopomenskimi glagoli, kot npr. *razlagati*, *poslavljati se*, *vračati se* ipd. *vesel*. Pri pomožnih glagolih upoštevamo tudi strukture, kjer je pomožniški del sam na sebi zložen: (biti + glagol v nedoločniku) = (Vez-gbz) pbz, npr. *je bil videti*, *je bil slišati* ipd. *vesel*.

Tabela 14: Nejedrne strukture pri pridevnikih



Tip	Skladenjska struktura	Kolokacijski zgled
pridevniške/prilastkovne	PBZ1 sbz1 gbz PBZ1	vesel [obraz] [vstati] neprespan , [razlagati] vesel
pomožniške	Vez-gbz PBZ1 Vez-gbz PBZ2 gbz Inf-gbz PBZ1	<ul style="list-style-type: none"> • [biti] vesel; [čutiti se] vesel, [ostati, postati] vesel; [zdeti se] vesel • [biti videti, biti slišati] vesel [čutiti se] veselega [znati biti] nesramen , [začeti postajati] nor , [poskušati biti] sproščen
količinske	Kol-PBZ sbz4 Kol-PBZ2 sbz2	vsako [leto], tri [metre], vsako [nedeljo] tretjega [marca]

2.3.1.3.3 Zapisovanje pridevniških struktur

Ker je primarna narava pridevnika določanje samostalniških, prislovnih in glagolskih besednih zvez, si pridevniške strukture, ko jih zapisujemo v programski vmesnik, ne sledijo nujno glede na jedrnost oz. nejedrnost pridevniške iztočnice, saj so najbolj tipične za pridevnik navadno tiste strukture, kjer pridevnik NI jedro besedne zveze. Za razliko od samostalniških struktur si zato pridevniške strukture v programskem vmesniku sledijo glede na tipičnost/pogostnost.

2.3.1.4 Prislov

Tudi pri prislovnih iztočnicah beležimo jedrne in nejedrne prislovne skladijske strukture.

2.3.1.4.1 Jedrne prislovne strukture

V nadelementu **skladijske skupine** beležimo pri prislovnih iztočnicah med jedrnimi strukturami tiste, kjer prislov/prislovna zveza ali odvisni stavek določa prislovno jedro. Kot jedrne prislovne strukture obravnavamo tudi prislov/zveze v vlogi povedkovnika.



Tabela 15: Jedrne strukture pri prislovih

Skladenjska struktura	Kolokacijski zgled
rbz RBZ	[nadvse, izjemno] preprosto , [čisto] blizu
Vez-gbz RBZ	lepo je; dobro je: (kaj) je lepo
RBZ Odv- da, ... kot	preprosto , da (bolj ne more biti); (nekaj) je lepo , kot ... da, ker ...

2.3.1.4.2 Nejedrne prislovne strukture

Med nejedrnimi prislovnimi strukturami ločimo dva podtipa:

(a) Strukture, kjer prislov/besedna zveza določa samostalnik/zvezo, pridevnik/zvezo, prislov/zvezo ali glagol/zvezo. Tu beležimo tudi prislove/besedne zveze v vlogi prislovnih določil ob glagolskih iztočnicah in zveze s količinskimi prislovnimi določili.

Tabela 16: Nejedrne strukture pri prislovih – (a)

Struktura	Kolokacijski zgled
sbz1 RBZ	[pot] domov
RBZ gbz	močno [deževati], drastično [se znižati]
RBZ pbz1	bledo [rdeč], hudo [zapit]; gotovo [najpomembnejši]
RBZ rbz	hudo [dobro]
gbz RBZ	[zmagati] doma
gbz Inf-gbz RBZ	[zna biti] hudo , [utegniti postati] zabavno
Kol-RBZ sbz2	nekaj [časa], precej [dela]
Kol-RBZ sbz4	najmanj [deset]



(b) Strukture s prislovom/besedno zvezo kot iztočnico v kombinaciji s predložno samostalniško besedno zvezo. Kolokacijske zglede za posamezne strukture smo v spodnji preglednici razdelili še glede na izražanje vrste okoliščin.

Tabela 17: Nejedrne strukture pri prislovih – (b)

Struktura	Kolokacijski zglede		
	kraj	čas	količina/vzrok/način ...
RBZ čez sbz1		očitno čez [noč]	
RBZ do Kol-sbz4			najmanj do [deset]
RBZ do sbz2		najpozneje do [februarja]	
RBZ kljub sbz3			mogoče kljub [vsemu]
RBZ na sbz5	najverjetneje na [Hrvaškem], nikjer na		
RBZ na sbz4	[svetu]	trikrat na [leto]	
RBZ ob sbz5		lani ob [200. letnici] zjutraj ob [pol šestih]	
RBZ od sbz4	dlje od [koga] zahodno od [Bagdada]		
RBZ od sbz2		danes od [štirinajstih]	najmanj od [dveh]
	povsod po [svetu]	takoj po [seji], prvič po [vrhu]	
RBZ pred sbz6		popoldne pred [sejo]	
RBZ s sbz6		hkrati s [predlogom]	skladno z [inflacijo], vključno s [50 milijoni]; skupaj z [DDV]



RBZ v sbz5	drugje v [Evropi], nekje v [sredini]	pozneje v [življenju] najpozneje v [aprilu]	
RBZ za sbz4			posebej za [Delo]
RBZ zaradi sbz2			prav zaradi [tega]

2.3.1.5 Kodiranje skladijskih struktur in skladijskih vzorcev

V drugi fazi izdelave LBS predvidevamo kodiranje skladijskih informacij za namene računalniške obdelave. Postopek predvideva dva načina:

- označevanje posameznih delov skladijskih struktur in skladijskih vzorcev s predvidenimi elementi ter pripis ustreznih atributov;
- avtomatično generiranje megastruktur iz označenih skladijskih vzorcev

2.3.1.5.1 Glagol

Elementi, s katerimi označujemo posamezne dele skladijskih struktur in skladijskih vzorcev pri glagolu so naslednji:

S – označuje besednozvezni del skladijskega vzorca, tj. samostalnik oz. samostalniško besedno zvezo, prislov oz. prislovno besedno zvezo ali odvisni stavek. Pri skladijskih strukturah lahko posamezne dele struktur zapolnjujejo še pridevniki/pridevniške besedne zveze, prislovi/prislovne besedne zveze in glagoli oz. glagolske besedne zveze.

L – označuje lemo; v primeru glagolskih struktur je to obravnavani glagol, kar označuje zapis z velikimi črkami. V skladijskih vzorcih je to vedno konkretni glagol v tretjeosebni edninski obliki: koga boli kaj, kdo pade s česa.

P – označuje del strukture oz. skladijskega vzorca, ki mu pripišemo konkretne oblikoskladijske vrednosti v atributu, npr. predlog v ustreznem sklonu ali veznik. Kadar gre za podredni veznik, ga v skladijskem vzorcu vedno zapišemo kot konkretni veznik, npr. *spominjati se, ko ...; vprašanje, ali ...* Kadar je za vzorec tipičnih več veznikov, jih zapisujemo vsakega v svojem vzorcu. Priredni vezniki pridejo pri glagolskih iztočnicah v poštev samo pri prirednih strukturah, kjer vedno izpišemo konkretni veznik, npr. *spotakniti se in pasti: gbz in GBZ*, v atributu pa ima oznako Vp.



Tabela 18: Elementi v skladenjskih vzorcih, njihov opis, vrednost atributa in zapis

Element	Opis elementa	Vrednost atributa	Zapis v skladenjskem vzorcu
S – struktura	Označuje besednozvezne (frazne) dele vzorca. Omogoča povezavo s skladenjskim razčlenjevalnikom.	<ul style="list-style-type: none"> • sbz1-6 • rbz • Odv • BrezO 	<ul style="list-style-type: none"> • kdo/kaj • kam, kako ... • konkretni veznik • >Rbrezosebno/R< za delom vzorca, na katerega se nanaša
L – lema	Označuje osnovno obliko obravnavanega glagola v vzorcu. Omogoča povezavo z leksikonom.	<ul style="list-style-type: none"> • lema 	<ul style="list-style-type: none"> • konkretni glagol
P – MSD opis	Označuje dele vzorca, ki jih določajo oblikoskladenjske oznake. Omogoča povezava z MSD oznakami.	<ul style="list-style-type: none"> • D1-6 • Vd 	<ul style="list-style-type: none"> • konkretni predlog • Odv-konkretni podredni veznik

Tabela 19: Elementi v glagolskih strukturah, njihov opis, vrednost atributa in zapis

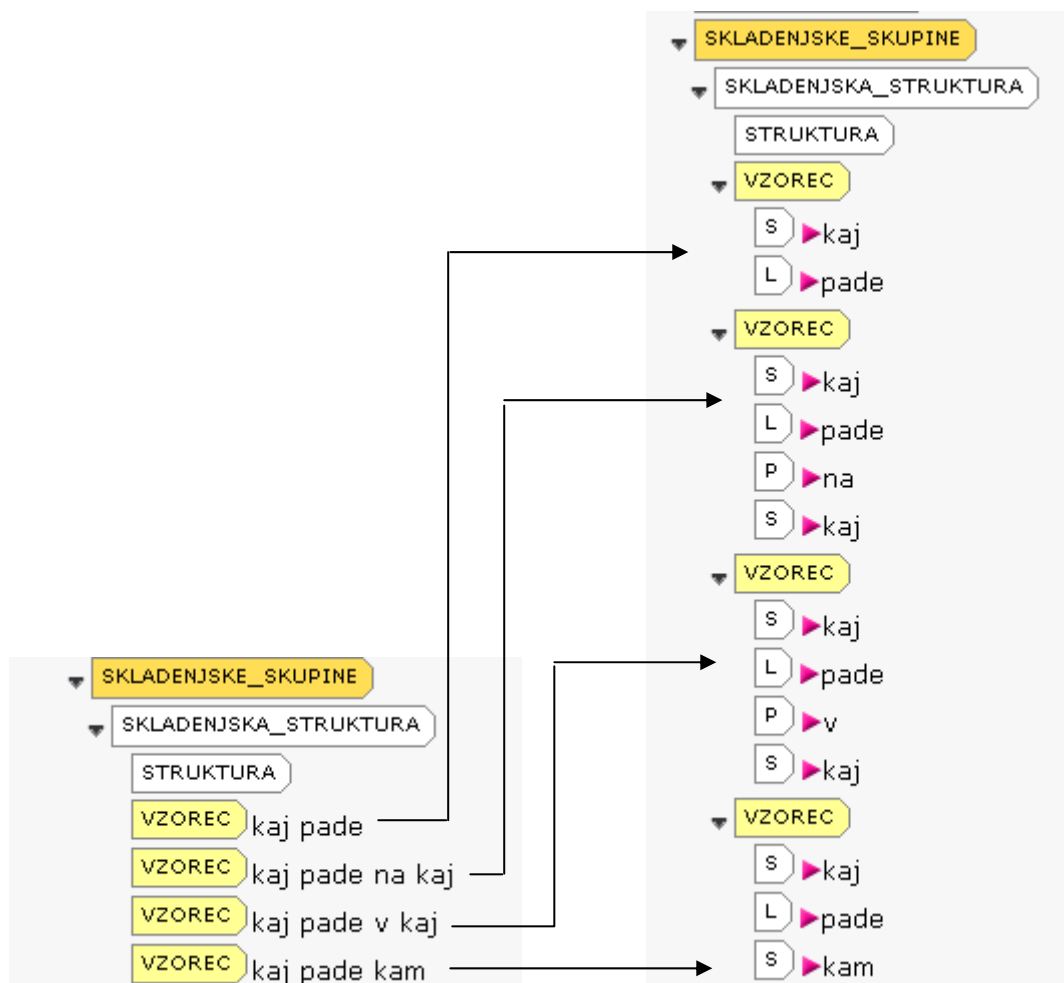
Element	Opis elementa	Vrednost atributa	Zapis v skladenjski strukturi
S – struktura	Označuje besednozvezne (frazne) dele strukture. Omogoča povezavo s skladenjskim razčlenjevalnikom.	<ul style="list-style-type: none"> • gbz • rbz • pbz 	<ul style="list-style-type: none"> • gbz • rbz • pbz • Vez-gbz • Inf-gbz



L – lema	Označuje osnovno obliko obravnavanega glagola v vzorcu. Omogoča povezavo z leksikonom.	• lema	• GBZ • Vez-GBZ • Inf-GBZ
P – MSD opis	Označuje dele strukture, ki jih določajo oblikoskladenjske oznake. Omogoča povezava z MSD oznakami.	• Vp	• konkretni priredni veznik

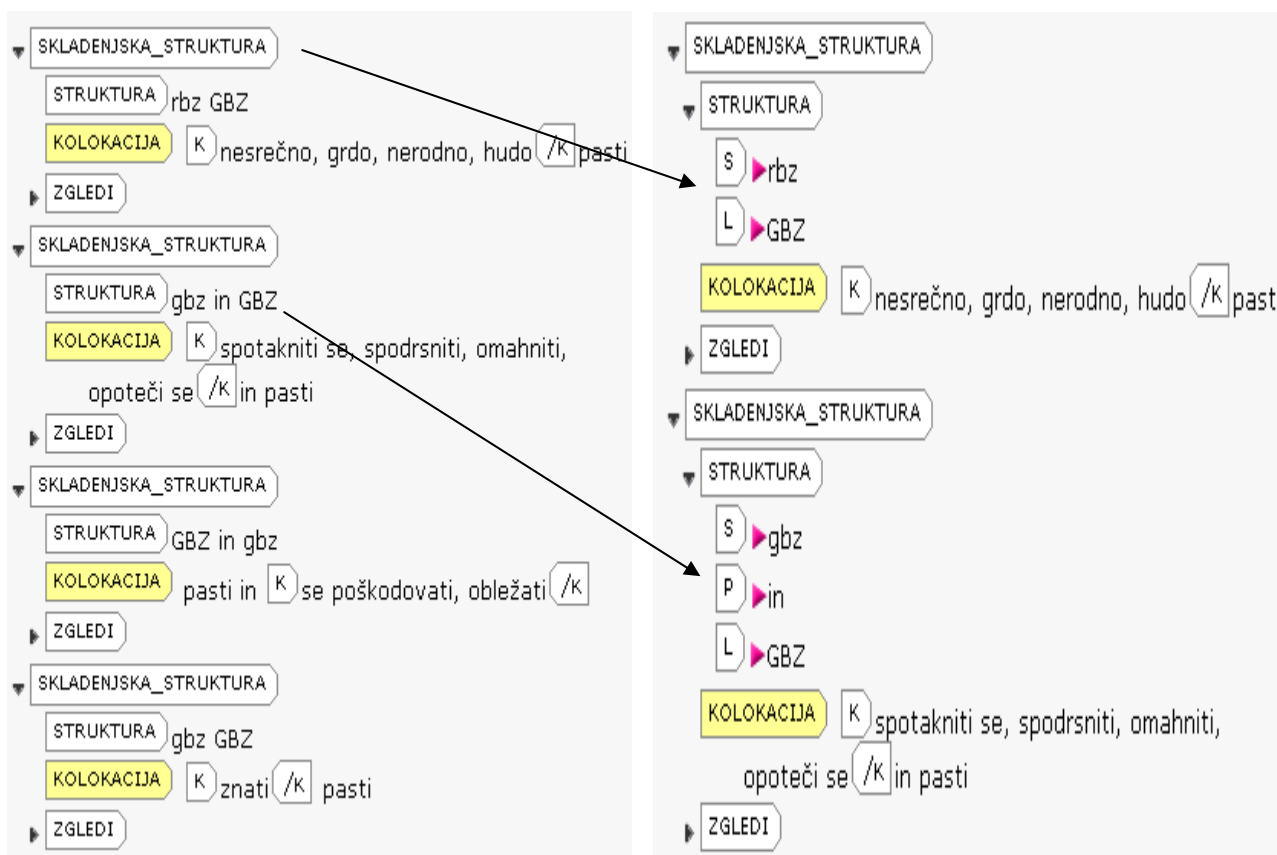
Če zgornje informacije prenesemo na konkretni primer, potem skladenjski vzorec: *kdo pade na kaj* (ki ustreza skladenjski strukturi: : sbz1 GBZ na sbz4) v prvi fazi izdelave LBS-A zapišemo v polje struktura, v drugi fazi pa pripišemo posameznim delom skladenjskega vzorca elemente in pripadajoče atributne vrednosti, kot kaže spodnja slika.

Slika 12: Zapis skladenjskih vzorcev v 1. in 2. fazi izdelave LBS



Enako kot pri skladijskih vzorcih tudi zapis skladijskih struktur poteka v dveh fazah. V prvi fazi zapišemo »gole« skladijske strukture v polje »struktura«, kot je razvidno iz tabele, tj. brez označenih elementov in pripisanih atributov. V drugi fazi opredelimo posamezne dele struktur z ustreznimi elementi in izberemo pripadajoče atributne vrednosti. Rezultat druge faze oblikovanja skladijskih struktur pri glagolu prikazuje spodnja slika.

Slika 13: Zapis glagolskih struktur v 1. in 2. fazi izdelave LBS



2.3.1.5.2 Samostalnik

Elementi, s katerimi označujemo posamezne dele skladijskih struktur pri samostalniku (2. faza: LBS-B) so naslednji:

S – označuje besednozvezni del skladijske strukture, in sicer: samostalnik oz. samostalniško besedno zvezo, glagol/glagolsko besedno zvezo, pridevnik/pridevniško besedno zvezo, prislov oz. prislovno besedno zvezo ali odvisni stavek. Poleg omenjenih besednih zvez nekatere dele samostalniških skladijskih struktur bolj natančno določamo, in sicer količinske prislove, glagol v pomožni vlogi in nedoločnik.

L – označuje lemo oz. samostalnik, ki ga obravnavamo v iztočnici. Pri jedrnih skladijskih strukturah je iztočnica vedno samostalnik/samostalniška besedna zveza v imenovalniku, zapisana z velikimi črkami. Pri nejedrnih skladijskih strukturah so samostalniške iztočnice vedno v odvisnih sklonih, zapisujemo pa jih prav tako z velikimi črkami.



P – označuje del skladenjske strukture, ki mu pripišemo oblikoskladenjske vrednosti v atributu, pri samostalniških strukturah so to na ravni zapisa konkretni predlogi, ki jim ustreza oznaka D (predlog) v ustreznem sklonu.

Tabela 20: Elementi v samostalniških strukturah, njihov opis, vrednost atributa in zapis

Element	Zapis strukturnega dela (1. faza)	Atribut (2. faza)	*Restrikcije (1. faza)
S	sbz: samostalnik/samostalniška besedna zveza	sbz1-6	
	gbz: glagol/glagolska besedna zveza	gbz	
	pbz: pridevnik/pridevniška besedna zveza	pbz	
	rbz: prislov/prislovna besedna zveza	rbz	
	Odv-konkretni veznik/zaimek: odvisni del skladenjske strukture, ki ga uvaja konkretni veznik	Odv	
	Kol-rbz: količinski prislov	rbz*	*>R količinski prislov /R<
	Vez-gbz: vezni (pomožni) glagol	gbz* ali Vez-gbz	*>R vezni glagol /R<
	Inf-gbz: nedoločnik	gbz* ali Inf-gbz	*>R nedoločnik /R<
L	SBZ1 (jedrne skladenjske strukture)	samostalnik v	
	SBZ2-6 (nejedrne skladenjske strukture)	osnovni obliki	
P	konkretni predlog	D1-6	

2.3.1.5.3 Pridevnik in prislov

Kodiranje skladenjskih struktur predvideva pri pridevniških in prislovnih iztočnicah iste elemente, kot pri glagolskih in samostalniških iztočnicah. Spodnji tabeli prikazujeta, kateri deli skladenjskih struktur se lahko pojavljajo v posameznih elementih pri pridevniških in prislovnih skladenjskih strukturah in katere attribute jim pripišemo v 2. fazi izdelave LBS.



Tabela 21: Kodiranje pridevniških struktur

Element	Zapis strukturnega dela (1. faza)	Atribut (2. faza)	*Restrikcije (1. faza)
S	sbz: samostalnik/samostalniška besedna zveza	sbz1-6	
	gbz: glagol/glagolska besedna zveza	gbz	
	pbz: pridevnik/pridevniška besedna zveza	pbz	
	rbz: prislov/prislovna besedna zveza	rbz	
	Odv-konkretni veznik – odvisni del skladijske strukture, ki ga uvaja konkretni veznik	Odv	
	Kol-rbz: količinski prislov ali količinska prislovna besedna zveza	rbz* ali Kol-rbz	*>R količinski prislov/zveza/R<
	Kol-sbz – količinska samostalniška besedna zveza	pbz* ali Kol-pbz	*>R količinska samostalniška besedna zveza/R<
	Kol-pbz – količinski pridevnik (števniki) ali količinska/števniška besedna zveza	pbz* ali Kol-pbz	*>R števniki ali števniška besedna zveza /R<
	Vez-gbz: vezni (pomožni) glagol	gbz* ali Vez-gbz	*>R vezni glagol /R<
Inf-gbz: nedoločnik	gbz* ali Inf-gbz	*>R nedoločnik /R<	
L	PBZ (jedrne skladijske strukture)	pridevnik v	
	PBZ2 (nejedrne skladijske strukture)	osnovni obliki oz. v ustreznem sklonu	
P	konkretni predlog	D1-6	



Tabela 22: Kodiranje prislovnih struktur

Element	Zapis strukturnega dela (1. faza)	Atribut (2. faza)	*Restrikcije (1. faza)
S	sbz: samostalnik/samostalniška besedna zveza	sbz1-6	
	gbz: glagol/glagolska besedna zveza	gbz	
	pbz: pridevnik/pridevniška besedna zveza	pbz	
	rbz: prislov/prislovna besedna zveza	rbz	
	Odv-konkretni veznik – odvisni del skladenjske strukture, ki ga uvaja konkretni veznik	Odv	
	Kol-rbz: količinski prislov ali količinska prislovna besedna zveza	rbz* ali Kol-rbz	*>R količinski prislov/zveza/R<
	Kol-sbz – količinska samostalniška besedna zveza	pbz* ali Kol-pbz	*>R količinska samostalniška besedna zveza/R<
	Kol-pbz – količinski pridevnik (števnik) ali količinska/števniška besedna zveza	pbz* ali Kol-pbz	*>R števnik ali števniška besedna zveza /R<
	Vez-gbz: vezni (pomožni) glagol	gbz* ali Vez-gbz	*>R vezni glagol /R<
	Inf-gbz: nedoločnik	gbz* ali Inf-gbz	*>R nedoločnik /R<
L	RBZ	prislov	
P	konkretni predlog	D1-6	

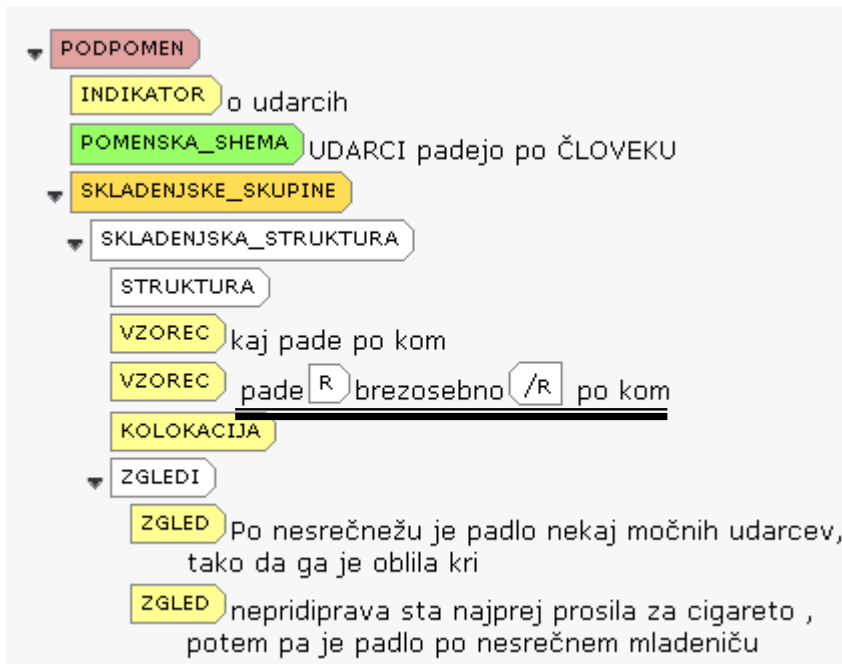


2.3.1.6 Slovnčne omejitve in preference

Že v prvi fazi izdelave LBS beležimo ob posameznih delih strukture ali vzorca tudi slovnčne omejitve ali t. i. koligacije (ang. colligational preferences/restrictions),²⁹ s katerimi izražamo tendenco nekaterih LE, da izbirajo ali pa se izogibajo določenim slovnčnim oblikam ali položajem. Take omejitve so razvidne iz rabe LE in izhajajo iz njenega skladskega okolja. Informacije o slovnčnih omejitvah in preferencah so zanimive tako za slovarske in slovnčne aplikacije kot tudi za računalniško procesiranje jezika, zato jih v programskem vmesniku zapisujemo v ločene elemente glede na to, ali želimo opozoriti, da se omejitev nanaša na celotni pomen/podpomen LE (element >oznaka< neposredno za pomenskim indikatorjem), ali zgolj na posamezni del strukture/vzorca (element >restrikcija<) pri skladskih strukturah (elementa S in L). Ker take tendence navadno niso absolutne v smislu, da veljajo za vse primere, v zvezi z njimi uporabljamo oznako »navadno«. Ker je prepoznavanje teh omejitev namenjeno tudi računalniškemu procesiranju, jih v programskem vmesniku označimo z elementom >R< (restrikcija), kot prikazuje spodnja slika, že v 1. fazi izdelave LBS.

Slika 14: Označevanje restrikcij pri glagolskih vzorcih v programskem vmesniku

²⁹ Hoey (1998) izhodiščno Firthov termin koligacije definira z naslednjimi merili: tipična gramatična družba, ki jo navadno ima (ali nima) določena beseda; tipična slovnčna funkcija, ki jo ima (ali nima) beseda; določeno mesto v nizu, ki ga tipično ima ali nima določena beseda.



V elementu >R< ob posameznem delu skladenjske strukture označujemo naslednje slovnične omejitve ali preference:

- Navadno v 3. osebi, če se omejitev nanaša na del strukture, sicer v elementu oznaka neposredno za pomenskim indikatorjem;
- brezosebno, npr. **pade >R< brezosebno /R<** po kom;
- vezni glagol, npr. spomin **je** boleč – SBZ1 Vez-gbz >R vezni glagol/R<;
- zanikani glagol, npr. **ni** časa za kaj – Neg-gbz SBZ2;
- glagol v nedoločni obliki (nedoločnik), npr. možnost **pritožiti se >R<** nedoločnik/R<;
- navadno množina, npr. [obujati, obuditi, zbudjati, priklicati, prebuditi] **spomine** – Glag SBZ4 >R nav. množina /R<. To možnost, če je mogoče, izrazimo tudi na ravni kolokacij.
- navadno ednina, npr. [zgodovinski, kolektiven, spoštljiv] **spomin** – Prid SBZ1 >R nav. ednina /R<

Pomembno: Kadar se take omejitve (npr. raba glagola izključno v 3. osebi, števnost, neštevnost samostalnika ipd.) nanašajo na celotni pomen ali podpomen LE, informacijo zabeležimo v

elementu **oznaka** neposredno za pomenskim indikatorjem (gl. pogl. 2.2.2.4 Slovnice, funkcijske in pragmatične omejitve pomena).

Pri skladijskih strukturah, ki vključujejo omejitve posameznih delov, npr. vezni glagol, nedoločnik itd., te dele na poseben način zapisujemo že na ravni strukture, kot je razvidno iz kodirnih tabele. V 2. fazi jim v postopku kodifikacije pripišemo ustrezen atribut in restrikcijo.

2.3.2 Skladijske zveze

Za razliko od skladijskih struktur in skladijskih vzorcev predstavljajo skladijske zveze tisto vrsto skladijskih informacij, ki vzpostavlja stično točko med kolokacijami in stalnimi besednimi zvezami na eni strani in frazeološkimi enotami na drugi. Oznaka »polfrazeološke« enote, ki se je uveljavila zanje, implicira:

- (a) kolokacijskost, ki presega okvire klasičnih dvobesednih kolokacij, npr. *pade [nekaj, malo, veliko, dosti] padavin*;
- (b) strukturno trdnost, v kateri je vsaj en del skladijske zveze stalen oz. zgolj variantno zamenljiv, npr. samostalniške predložne zveze: *v spomin, pred časom, na kose*; prislovne: *vsak čas, prej ali slej*; glagolske: *stisniti se v klobčič* ipd.

Hkrati skladijske zveze ne sodijo v razdelek stalnih besednih zvez ali frazeološki razdelek, saj nimajo od sestavin neodvisnega (metaforičnega) pomena, npr. *pasti v vodo* – 'propasti'; *vrteti se v krogu* – 'biti v brezizhodnem položaju' ipd.

Ker so skladijske zveze vedno bodisi s konkretnimi kolokatorji razširjeni deli bodisi ustaljeni deli jezika, jih v LBS v 2. fazi izdelave ne nameravamo formalizirati z atributi. Manj strog v smislu kombiniranja splošnoleksikalnih (*koga/komu* ipd.) in konkretnih kolokatorjev v zvezi, npr. (*breme, skrb*) *pade na (plača, rame/ramena) koga/komu*, je zato tudi njihov zapis v programskem vmesniku, kot bomo videli v nadaljevanju.



2.3.2.1 Skladenjske zveze pri glagolih

2.3.2.1.1 Kolokacijsko razširjeni segmenti jezika

Najpogosteje kot skladenjske zveze pri glagolskih iztočnicah obravnavamo **razširjene kolokacije**.

S tem imamo v mislih naslednje možnosti:

(a) V prvem primeru želimo opozoriti na tipičnost **širše okolice že registriranih kolokacij**, ki vključujejo obravnavani glagol v iztočnici. V takih primerih navadno posegamo v skladenjsko področje drugih besednih vrst, pri glagolih zlasti v samostalniške zveze in obratno. Ker pa so skladenjske zveze v prvi vrsti namenjene slovarskim aplikacijam, manj pa slovničnim analizam in računalniškemu procesiranju, predvidevamo možnost, da tudi pri glagolskih iztočnicah opozorijo na pomembnost širšega okolja, ki ni več zgolj v skladenjski pristojnosti glagola. Spodnji zgledi prikazujejo razširitve predhodno zabeleženih kolokacij, ki so značilne za konkretni pomen glagola.

Zgled 24: Razširitev glagolskih kolokacij v skladenjske zveze – (a)

Kolokacije

[temperatura] pade
pasti na [pleča, ramena]
boli v [prsih, križu ...]

Skladenjske zveze

temperatura pade pod [ničlo, ledišče]
breme pade na [pleča, ramena] komu
boli [v predelu] česa

(b) V drugem primeru gre za razširitev kolokatorjev **znotraj že registriranih glagolskih kolokacij**.

Zgled 25: Razširitev glagolskih kolokacij v skladenjske zveze – (b)

Kolokacije

[dež, sneg] pade
[sum, breme, senca] pade
[padavine, kaplje ...] padejo

Skladenjske zveze

[nekaj, nekoliko, toliko] kapelj, kapljic [dežja] pade
senca [dvoma, krivde] pade na koga
[debele, dežne] kaplje padejo



Poleg klasičnih razširitev, kot so pridevniške razširitve samostalniških jeder ter prislovne razširitve glagolskih jeder, obravnavamo tu tudi količinske zveze tipa: količinski prislov in samostalnik/samostalniška besedna zveza v rodilniku in razširitve s števnimi oz. količinskimi izrazi, npr.:

Zgled 26: Razširitev glagolskih kolokacij s količinskimi izrazi

pade [nekaj, malo veliko...] dežnih kapelj
pasti pod kotom [x] stopinj
pasti z višine [x] metrov
obiskati [enkrat letno]
praznovati [vsako leto]

[ponovno, še enkrat, znova] se roditi
roditi [x] otrok
roditi se ob [6h zjutraj],
roditi se [22. aprila]
roditi se pred [x] leti

2.3.2.1.2 Strukturno ustaljene skladijske zveze

Med skladijskimi zvezami pri glagolih obravnavamo tudi strukturno ustaljene segmente jezika, ki niso pomensko osamosvojeni oz. je njihov pomen mogoče razbrati iz posameznih sestavin, npr.

Zgled 27: Strukturno ustaljene glagolske zveze

pasti

zveza: pasti si v objem
zveza: pasti komu v objem
zveza: pasti komu okrog/okoli vratu
zveza: pasti pod pritiskom česa
zveza: pasti pod vpliv česa

pomešati

zveza: pomešati kaj med seboj
zveza: kaj se pomeša med seboj

stisniti

zveza: stisniti koga k sebi
zveza: stisniti se v klobčič

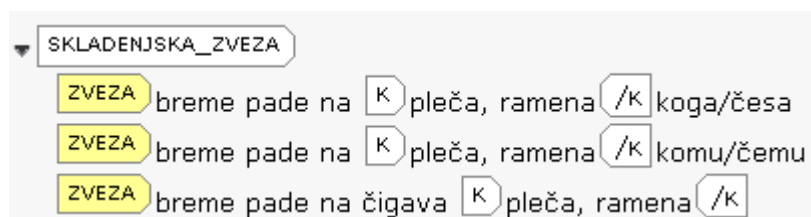
roditi

zveza: roditi po naravni poti

2.3.2.1.3 Zapis skladijskih zvez

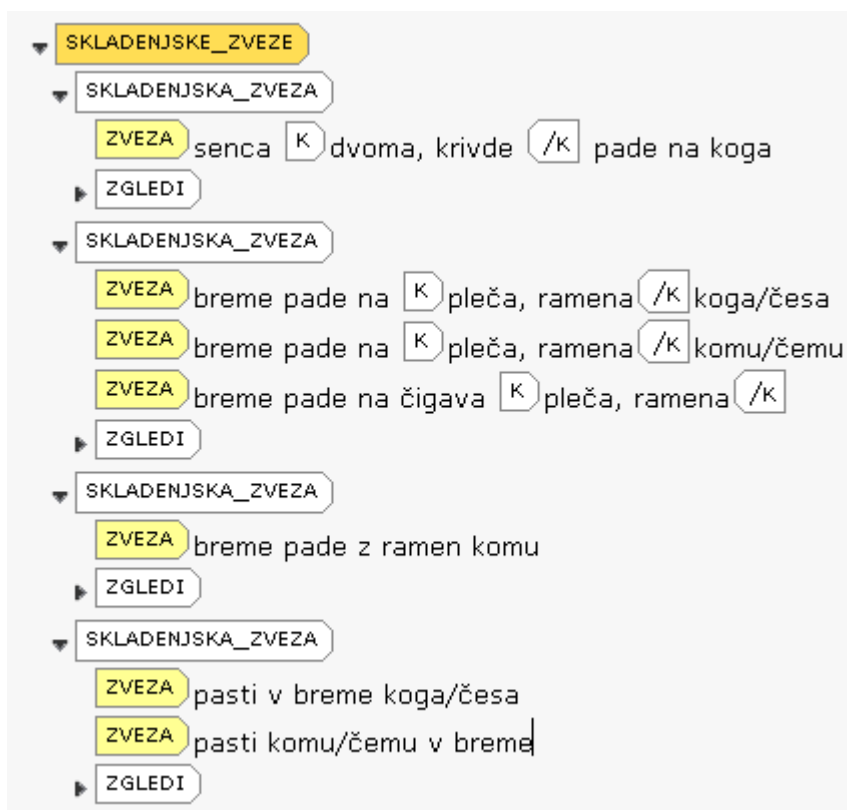
Skladijske zveze v programskem vmesniku vpisujemo v element skladijske zveze in nadelement skladijska zveza, pri čemer si znotraj skladijskih zvez (element zveza) lahko zaporedoma sledijo skladijske zveze, ki predstavljajo variante, različne oblike rabe ali sistemske možnosti (npr. tip: od koga – čigava, in se nanašajo na eno in isto zvezo, kot prikazuje spodnja slika.

Slika 15: Zapis zaporednih skladijskih zvez znotraj enega nadelementa



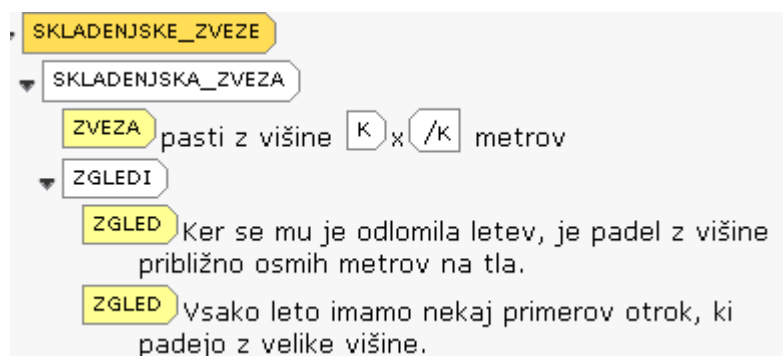
Ločeno, tj. vsako skladijsko zvezo v svoj nadelement, zapisujemo takrat, ko prihaja do sprememb v elementih zveze ali v širšem besedilnem okolju, kot kaže spodnja slika.

Slika 16: Zapis samostojnih skladijskih zvez znotraj ločenih nadelementov



Znotraj skladijskih zvez zapisujemo kolokabilne dele v elementu >K< (gl. slike zgoraj), vendar samo tisti del, ki je glede na že registrirano kolokacijo, nov. Številčne podatke pa bodisi s konkretnimi kolokatorji, če kateri izstopa, bodisi s splošno oznako x, kadar ne moremo izpostaviti tipične besede, kot prikazuje spodnja slika.

Slika 17: Zapis številčnih delov zvez v skladijskih zvezah



Skladenjske zveze lahko izkazujejo tudi svoje lastno kolokabilno okolje – zgledi kolokacij za glagolske skladenjske zveze

2.3.2.2 Skladske zveze pri samostalnikih

Tako kot pri glagolskih iztočnicah obsegajo skladske zveze pri samostalnikih:

- razširjene in hkrati pogoste dele jezika, ki jih ni mogoče zajeti s kolokacijami,
- v sebi zaključene (strukturno trdne) in hkrati pogoste dele jezika, ki nimajo lastnega pomena oz. je pomen predvidljiv iz sestavin.

Tako kot pri glagolih predstavljajo **skladske zveze** pri samostalnikih, pridevniki tisto vrsto informacij, ki presega strukturno »čiste« dele jezika, za katere želimo, da so tudi računalniško berljivi. Pri skladskih zvezah gre za razširjene in v jeziku izpostavljene ustaljene zveze, ki so nasprotno za računalniško formalizacijo manj zanimive (razen nekaterih izjem, kjer dopuščamo formalizacijo v obliki vzorcev, kot bomo pokazali v nadaljevanju), so pa nepogrešljiva slovarska informacija. Ker torej skladskih zvez ne bomo kodirali za namene računalniškega procesiranja, imajo redaktorji v programskem vmesniku več različnih možnosti obdelave. Odločitev za način predstavitve skladske zveze pa je odvisna od tega, ali gre zgolj za razširitev že evidentirane skladske strukture, za osamosvojeno skladsko zvezo brez tipičnega (kolokabilnega) besedilnega okolja, za relativno osamosvojeno skladsko zvezo z več skladsko različnimi realizacijskimi možnostmi ali pa za osamosvojeno skladsko zvezo, ki ima tudi lastno tipično besedilno okolje.

Pomembno: Pri samostalnikih (enako tudi pri pridevniških in prislovnih iztočnicah) za razliko od glagolov, kjer priredne in primerjalne zveze beležimo med glagolskimi strukturami, med zveze uvrščamo tudi vse priredne, npr. *noč in dan, čas in denar, veselje in smeh; kupiti ali prodati; kisel ali sladek* ipd. zveze, ter vse primerjalne zveze, ne glede na izraženo večjo ali manjšo stopnjo frazeološke ustaljenosti, npr. *vesel, kot ptiček na veji*.



2.3.2.2.1 Vrste in zapis skladenjskih zvez pri samostalnikih

Ločimo tri osnovne tipe skladenjskih zvez, za katere predvidevamo različne redakcijske možnosti.

(a) strukturno ustaljene zveze, ki ne izkazujejo tipičnega (kolokabilnega) besedilnega okolja

- (a) **zveza:** *pred časom*
kolokacija
zgled

(b) leksikalizirane (tj. strukturno in pomensko ustaljene) zveze, ki jih obravnavamo različno glede na to, ali:

- predvidevajo različne širše skladenjske vzorce, ki so sicer tipični, ne pa nujno tudi kolokabilni v smislu konkretnih leksikalnih enot na posameznih mestih razširjenega vzorca. V teh primerih uredniki navedejo bodisi zveze v izkazanih realizacijah, ki jim lahko dodajo še vzorec s splošnim glagolskim delom (*narediti kaj*) – (b1) ali pa splošni glagol vključijo v različne realizacije zveze same – (b2)

- (b1) **zveza:** *v spomin na koga*
zveza: *v spomin komu*
zveza: *v spomin koga*
vzorec: *narediti kaj v spomin*
kolokacija
zgled

- (b2) **zveza:** *narediti kaj v spomin na koga*
zveza: *narediti kaj v spomin komu*
zveza: *narediti kaj v spomin koga*

kolokacija
zgled

- želimo leksikalizirano zvezo posebej izpostaviti, različne realizacijske variante pa navesti v vzorcih, bodisi z izkazanim splošnim glagolskim delom – (b3) ali brez njega – (b4). Zapisovanje vzorcev v teh primerih sledi logiki zapisovanja skladenjskih vzorcev pri glagolih, pri čemer namesto konkretnega glagola tu izpisujemo ustaljeno zvezo.

- (b3) **zveza:** *v spomin*
vzorec: *v spomin komu*
vzorec: *v spomin koga*

- (b4) **zveza:** *v spomin*
vzorec: *narediti kaj v spomin komu*
vzorec: *narediti kaj v spomin koga*



vzorec: v spomin na koga
kolokacija
zglede

vzorec: narediti kaj v spomin na koga
kolokacija
zglede

(c) razširjene že evidentirane skladišne strukture

- (c) **zveza:** [zapravljati, izgubljeni] čas s čim
zveza: [imeti] boleč, lep spomin na koga

Kot kolokacijski niz, tj. v elementu >K<, navedemo znotraj zveze tiste kolokatorje, ki predstavljajo razširitev že registrirane kolokacije z »novimi« kolokatorji, npr: iskanje [primerne, ustrezne] zaposlitve. Kadar so ostali deli take razširjene kolokacije tudi kolokabilni, kar pomeni, da ne želimo izpostaviti samo enega, jih prav tako navedemo v elementu >K<, npr.: iskanje [primerne, ustrezne] [zaposlitve, rešitve]. V primeru, da je izpostavljen samo en kolokator, npr. iskanje [nove] zaposlitve – zapisujemo v elementu >K/< samo kolokabilno razširjeni del.

2.3.2.2 Primeri skladišnih zvez (samostalni čas)

Spodnji zgledi prikazujejo zapis skladišnih zvez pri iztočnici **čas** s pripadajočimi vzorci, kjer se želi izpostaviti različne tipične realizacijske možnosti, ki jih izkazuje raba, in kolokacijami, kjer skladišna zveza izkazuje tudi tipično lastno besedilno okolje.

Zgled 28: Skladišne zveze pri samostalnemu času

čas

1 dimenzija obstoja

skladišna zveza

zveza: skozi čas

skladišna zveza

zveza: čez čas

skladišna zveza

zveza: pred časom

2 omejeno trajanje, obdobje

skladišna zveza

zveza: za kratek/krajši čas

skladišna zveza

zveza: za dolgo/dlje/dalj časa

skladišna zveza

zveza: po dolgem/daljšem času



skladenjska zveza

zveza: ves čas

vzorec: imeti (kaj) ves čas pred očmi

skladenjska zveza

zveza: vsake toliko časa

skladenjska zveza

zveza: čez nekaj časa

skladenjska zveza

zveza: pustiti kaj času

zveza: pustiti času, da

skladenjska zveza

zveza: stiska s časom

skladenjska zveza

zveza: občutek za čas

skladenjska zveza

zveza: čas in trud

zveza: trud in čas

kolokacija: [veliko, ogromno, precej] časa in truda

kolokacija: [veliko, ogromno, precej] truda in časa

skladenjska zveza

zveza: čas in ...

kolokacija: čas in [energija, prostor, kraj, denar]

zveza: ... in čas

kolokacija: [energija, prostor, kraj, denar] in čas

skladenjska zveza

zveza: imeti časa na pretek

skladenjska zveza

zveza: zadnji čas

zveza: v zadnjem času

zveza: v zadnjih časih

zveza: zadnje čase

skladenjska zveza

zveza: največji kdo vseh časov

skladenjska zveza

zveza: čas trajanja

vzorec: čas trajanja česa

kolokacija: čas trajanja [90 minut]

kolokacija: čas trajanja [akumulatorja]

kolokacija: čas trajanja [pogodbe, postopka]

vzorec: v času trajanja česa

kolokacija: v času trajanja [mandata, pogodbe]

skladenjska zveza

zveza: v času pred čim

kolokacija: v času pred [volitvami, osamosvojitvijo]

zveza: v času pred začetkom česa

kolokacija: v času pred začetkom [prireditve, volilne kampanje]

skladenjska zveza

zveza: v času po čem

kolokacija: v času po [volitvah, osamosvojitvi, vojni, revoluciji]

kolokacija: v času po [upokojitvi, porodu]

skladenjska zveza

zveza: čas od česa do česa / čas od ... do ...

kolokacija: čas od [okužbe] do [pojava bolezni, obolenja]

kolokacija: čas od [ponedeljka] do [petka]

skladenjska zveza

zveza: v času od česa

kolokacija: v času od [začetka,



nastanka]

skladenjska zveza

zveza: časa je/ni na pretek

skladenjska zveza

zveza: ni časa za kaj

skladenjska zveza

zveza: duh časa

skladenjska zveza

zveza: tok časa

skladenjska zveza

zveza: čas bo pokazal, če/ali ...

skladenjska zveza

zveza: kaj je vprašanje časa

kolokacija: kaj je [le, samo] (še) vprašanje časa

skladenjska zveza

zveza: v odvisnosti od časa

skladenjska zveza

zveza: v časih pred prihodom koga

skladenjska zveza

zveza: v času ob čem

kolokacija: v času ob [uvedbi, menjavi, vstopu]

kolokacija: [segati, spadati, biti postavljen] v čas ob čem

skladenjska zveza

zveza: v času do česa

kolokacija: v času do [volitev, leta x, uveljavitve, spremembe, sprejetja]

skladenjska zveza

zveza: v času/časih

vzorec: v času/časih koga

vzorec: v času česa

skladenjska zveza

zveza: v času okrog česa

kolokacija: v času okrog [leta, dneva, ovulacije, volitev, božiča]

skladenjska zveza

zveza: v času med čim

kolokacija: v času med [vojno, vojnama, letoma, leti]

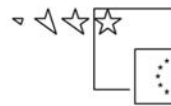
skladenjska zveza

zveza: preložiti kaj za nedoločen čas

2.3.2.3 Skladenjske zveze pri pridevnikih in prislovih

Pri pridevnikih in prislovih evidentiramo strukturno ustaljene in pomensko prozorne skladenjske zveze.

Zgled 29: Skladenjske zveze pri pridevnikih in prislovih



Prislovne skladijske zveze

zveza: od nekdanj

zveza: enkrat za vselej

zveza: (biti) več kot kakšen/kdo

kolokacija: več kot razumljiv

kolokacija: več kot poštenjak

zveza: na kratko

zveza: ne glede na

zveza: glede na

zveza: tik pred (čem)

zveza: za nazaj

zveza: še posebno

Pridevniške skladijske zveze

zveza: po svoje

zveza: tako rekoč

Tu registriramo tudi vse notranje kolokacijske razširitve, zlasti s količinskimi izrazi, npr. *pasti [150, nekaj ...] metrov stran* ipd.

2.4 Kolokacijski nivo

V tem segmentu LBS beležimo očitno tendenco nekaterih besed, da se sopojevajo z drugimi besedami: *kolokacije* tako razumemo kot pogoste kombinacije besed, v katerih določena beseda kaže očitno težnjo po sopojevljanju z drugo besedo, pri čemer je pogostost večja od naključne sopojevjitve. Kolokacija nam zato predstavlja v prvi vrsti informacijo o:

(a) tipičnih delčkov jezika:

pasti pod (ničlo, vpliv, streli, mejo, kolo ...)

pasti na (pamet, tla, izpitu, bojišču ...)

(b) ki so pogostokrat tudi pomensko relevantni:

pasti pod (ničlo, minus, ledišče) – *znižati se*

pasti pod (streli, krogli) – *umreti*

pasti pod vpliv – *podleči*

pasti pod (mejo, raven, nivo) – *znižati se*

pasti pod (kolo, traktor) – *izgubiti ravnotežje*

pasti na (pamet, misel, um) – *FE*

pasti na (tla, vozišče, pločnik) – *izgubiti ravnotežje*

pasti na (izpitu, testu) – *neizdelati*

pasti na (bojišču, fronti) – *umreti*

2.4.1 Povezava med skladijskim in kolokacijskim nivojem v LBS

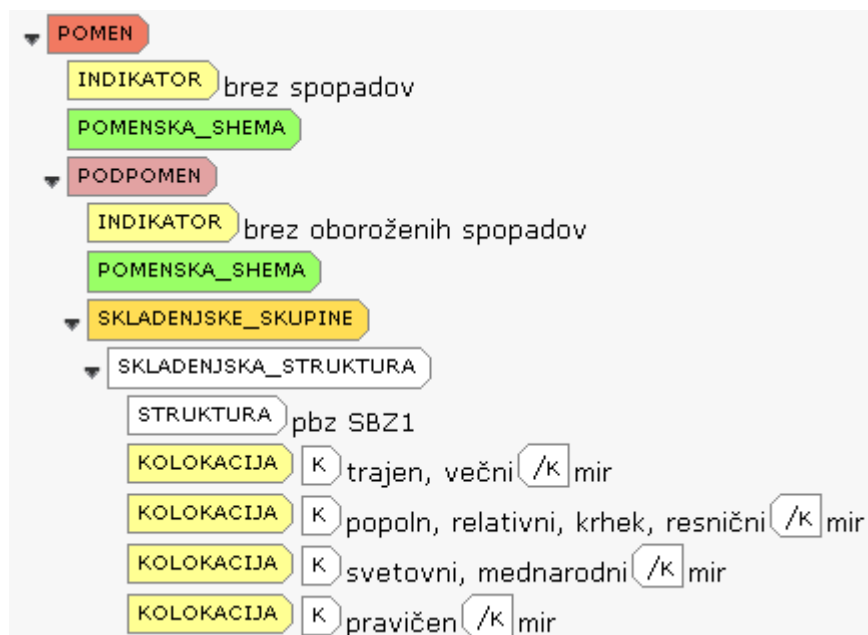
Tipične kolokacije, v katere vstopa obravnavana beseda v iztočnici, beležimo v LBS glede na registrirane pomene in skladijske strukture. Z upoštevanjem statističnih izračunov besedne povezovalnosti, kot so razvidni iz programa za izdelavo besednih skic (gl. Kazalnik 5), vključujemo v LBS informacijo o sintagmatskem povezovanju besed, s podrejanjem kolokacij pomenski členitvi pa upoštevamo tudi paradigatski vidik besedne povezovalnosti, kjer se kaže pomenska soodvisnost besed, ki se tipično sopoljavljajo. To pomeni, da ob vsakem pomenu ali/in podpomenu, ki ga za obravnavano besedo registriramo v LBS, beležimo v kolokacijah samo tiste dele struktur in skladijskih vzorcev, ki so kolokabilni:

Slika 18: Kolokabilni deli skladijskega vzorca za glagol *pogajati se*



- CLANEK
 - GLAVA
 - OBLIKA
 - ZAPIS pogajati se (SIKR)
 - IZTOCNICA pogajati se
 - ZAGLAVJE
 - BESVRS glagol
 - GESLO
 - POMEN
 - INDIKATOR poskušati se dogovoriti
 - POMENSKA_SHEMA OSEBA se pogaja z DRUGO OSEBO glede VSEBINE dogovora
 - SKLADENJSKE_SKUPINE
 - SKLADENJSKA_STRUKTURA
 - STRUKTURA SBZ1 pogajati se (s SBZ6) (o SBZ5 / glede SBZ2 / za SBZ4)
 - VZOREC kdo se pogaja
 - VZOREC kdo se pogaja s kom
 - VZOREC kdo se pogaja s kom o čem
 - VZOREC kdo se pogaja s kom glede česa
 - VZOREC kdo se pogaja s kom za kaj
 - KOLOKACIJA [K] vlada, država, vodstvo [/K] se pogaja
 - KOLOKACIJA [K] sindikati, delodajalci, partnerji [/K] se pogajajo
 - KOLOKACIJA pogajati se s/z [K] sindikati, EU, vlado, Brusljem, delodajalci [/K]
 - KOLOKACIJA pogajati se s/z [K] ugrabitelji, uporniki, teroristi [/K]
 - KOLOKACIJA pogajati se o [K] ceni, nakupu, odkupu, prodaji; pogojih, spremembah [/K]
 - KOLOKACIJA pogajati se o [K] članstvu, vstopu, sporazumu [/K]
 - KOLOKACIJA pogajati se o [K] vdaji, izpustitvi [/K]
 - KOLOKACIJA pogajati se za [K] ceno, plačo [/K]
 - KOLOKACIJA pogajati se za [K] vstop, članstvo, podaljšanje, izpustitev [/K]
 - KOLOKACIJA [K] trdo, tajno, intenzivno, neuspešno, resno [/K] se pogajati
 - KOLOKACIJA [K] biti pripravljen [/K] pogajati se
 - KOLOKACIJA [K] začeti, pričeti, nehati, nameravati, znati [/K] se pogajati

Slika 19: Kolokabilni deli skladske strukture pri samostalniku *mir*



2.4.1.1 Pri glagolu

Kot smo pokazali v predprejšnjem zgledu, si v nadelementu skladijske strukture pri glagolskih iztočnicah v elementu struktura sledijo skladijski vzorci. Če se posamezni deli skladijskega vzorca potrjujejo s kolokacijami, jih navajamo zaporedno v elementu kolokacija na koncu skladijskih vzorcev. Skladijski vzorec se lahko potrjuje z več kolokacijami (v spodnjem zgledu to označuje enaka barva vzorca in kolokacije) glede na to, kateri del vzorca se izkazuje kot kolokabilen, kot prikazuje spodnji zgled.

Zgled 30: Zapis kolokacij pri glagolskih iztočnicah

skladijske skupine

skladijska struktura

struktura sbz1 pade (sbz3) (pod sbz4/ na sbz4 /s sbz2 na sbz4 / za sbz2)

vzorec kaj pade

vzorec komu/čemu pade kaj

vzorec kaj pade pod kaj

vzorec kaj pade na kaj

vzorec kaj pade s/z česa na kaj

vzorec kaj pade za kaj



kolokacija [temperatura] pade
kolokacija [tečaj, indeks, cena, vrednost] pade
kolokacija pasti pod [raven, vrednost, nivo, povprečje, prag, ceno]
kolokacija pasti na [raven, ničlo, minimum]
kolokacija: pasti za [x] odstotkov

Če posamezni skladijski vzorec ali del vzorca, kot npr. zgoraj *komulčemu pade kaj* ni kolokabilen in ga s kolokacijo ni mogoče potrditi, navedemo v programskem vmesniku zgolj ustrezen zgled, npr.

vzorec *komu/čemu pade kaj*

zgled *Trgovine in gostilne bodo lahko kmalu zaprle svoja vrata , saj jim bo prodaja pijače padla na ničlo .*

Za skladijskimi vzorci navajamo v elementu skladijska struktura posebne glagolske strukture. Za razliko od skladijskih vzorcev zapisujemo posebne glagolske strukture vsako v svoj element. Vsako skladijsko strukturo, če se pokaže kolokabilna, potrdimo s kolokacijo in zgledi, kot v spodnjih primerih.

Zgled 31: Potrjevanje glagolskih struktur s kolokacijami

skladijska struktura

struktura rbz GBZ

kolokacija [nesrečno, grdo, nerodno, hudo] pasti

zgledi

zgled Na srednji skakalnici je hudo padel mladi skakalec.

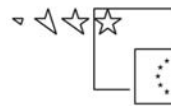
skladijska struktura

struktura GBZ in gbz

kolokacija pasti in [se poškodovati, obležati]

zgledi

zgled Kolesarka je padla po cestišču in se hudo poškodovala.



2.4.1.2 Pri pridevnikih in prislovih

Pri samostalnikih, pridevnikih in prislovih navajamo kolokacije v programskem vmesniku podobno kot pri glagolskih strukturah: kolokacije potrjujejo registrirane skladijske strukture, zato jih navajamo neposredno za njimi. V programskem vmesniku si znotraj skladijskih skupin torej sledijo skladijske strukture, pri čemer je posamezna struktura potrjena s kolokacijami, če obstajajo, in korpurnimi zgledi, kot kaže spodnji primer.

Zgled 32: Kolokacije pri pridevnikih in prislovih

skladijske skupine

skladijska struktura

struktura pbz1 SBZ1

kolokacija [dirkalni, jahalni, vlečni, vprežni, tekmovalni, kasaški] konj

zgledi

zglede

skladijska struktura

struktura sbz1 SBZ2

kolokacija [ljubitelj, pasma, rejec, reja, vzreja, lastnik] konj

kolokacija [blagoslov, žegnanje, razstava] konj

kolokacija [jahanje] konj

zgledi

zglede

V primeru, da za določeni pomen samostalnik, pridevnik ali prislov ne izkazuje kolokabilnega okolja, hkrati pa ne želimo posebej izpostaviti strukture, ker ni tipična, kot v spodnjih primerih, se pomen potrjuje zgolj v korpurnih zgledih.

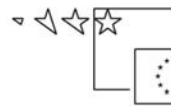
Zgled 33: Potrditev pomena z zgledom (ob odsotnosti tipičnih struktur in kolokacij)

pomen

indikator merska enota

skladijske skupine

skladijska struktura



struktura
kolokacija
zgledi

zglede S sedmimi sedeži v avtu je tako kot s 500 konji pod motornim pokrovom - skoraj nikoli jih ne potrebujemo, jih je pa lepo imeti.

zglede Menda naj bi ta avtomobil čez dve leti že zapeljal na ceste, napovedujejo pa zmogljiv motor z okrog 250 konji.

pomen

indikator šahovska figura

skladenjske skupine

skladenjska struktura

struktura
kolokacija
zgledi

zglede Črni bi rad pregnal ali zamenjal belega konja, ki se šopiri na polju e6.

Če pa se pomen iztočnice izraža v produktivni strukturi, ki ni kolokacijsko izpolnjena, želimo to strukturo vseeno zabeležiti, zato uporabimo možnost zapisa v obliki vzorca. Zapis vzorca sledi pravilom, kot smo jih pojasnili pri zapisu skladenjskih vzorcev pri glagolskih iztočnicah.

Zgled 34: Potrditev pomena z vzorcem (ob odsotnosti tipičnih kolokacij)

pomen

indikator

skladenjske skupine

skladenjska struktura

struktura PBZ1 zaradi sbz2
vzorec vesel zaradi česa
zgledi
zglede

2.4.2 Zapis kolokacij v programskem vmesniku

Kolokacije predvidevamo na vseh tistih segmentih LBS, kjer se deli predhodno registriranih vzorcev, struktur in skladenjskih zvez kažejo kolokabilni in tam, kjer večbesedne LE, tj. stalne



besedne zveze in frazeološke enote, izkazujejo lastno kolokabilnost. V programskem vmesniku zapisujemo kolokacije oz. kolokacijski niz v elementu >K/<. Kolokacijski niz sestavljajo tiste besede (kolokatorji), ki so glede na obravnavano besedo (iztočnico – podčrtano) spremenljivi, kar je zlasti pomembno pri **zapisu predlogov**: marmelada iz [pomaranč, jagod, sliv], čas pred [volitvami, prazniki], odporen proti [vodi, vlagi], znan pod [nazivom, imenom], leteti na [jug, sever] itd.

Pri zapisovanju kolokacij, ki zapolnjujejo skladijske vzorce glagolov, potrjujemo prisotnost osebkovega mesta z navajanjem glagola (iztočnice) v nedoločniku, npr. pomešati z/s [pecilnim praškom, vodo, sokom, sladkorjem, smetano, oljem]; stisniti [cvet, čebulico, ličinke] med prsti. Glagol – kot iztočniško besedo – navajamo v nedoločniku tudi v primeru modifikacij: samostojno, prosto, svobodno [leteti], [začeti, znati, naučiti se] leteti. Kadar je kolokabilen predmetni ali osebkov del skladijskega vzorca, navedemo glagolsko iztočnico v 3. osebi ednine oz. množine, če so kolokatorji tipično množinski: [glava, noga, trebuh, hrbet, želodec, grlo, zob, koleno] boli, [ptice, štoklje] letijo.

Pri zapisovanju kolokatorjev znotraj kolokacijskega niza sledimo v prvi vrsti berljivosti. Kolokacijske nize zato razbijamo oz. združujemo glede na **pripadnost** kolokatorjev istemu **pomenskemu polju** (znotraj pomensko že podrejenih kolokacij).

Zgled 35: Pomensko združevanje kolokatorjev znotraj kolokacijskega niza

leteti

padati kam

[asteroid, zvezdni utrinek] leti

[žoga] leti

[kamen] leti

potovati

leteti z/s [letalom, helikopterjem]

leteti z/s [padalom, balonom, zmajem]

čas

dimenzija obstoja

[vzeti si, rezervirati si, imeti] čas

[zapravljati, tratiti, izgublјati, krasti] čas



[nizko, visoko] leteti
[redno, množično] leteti

Glagolske kolokatorje pri samostalniških in pridevniških iztočnicah združujemo v kolokacijske nize tudi glede na **povratnost glagolov**, npr. [kljubovati, pripadati] času, [prilagoditi se, prilagajati se] času; rana [boli, skeli], rana [se celi, se zaraste].

Poleg pomenskega kriterija upoštevamo zaradi berljivosti pri razvrstiti v kolokacijskem nizu tudi slovnično ujemanje, zato razvrščamo kolokatorje pri samostalniških, pridevniških in prislovnih iztočnicah tudi glede na **ujemanje v spolu in številu pa tudi sklonu**.

Zgled 36: Združevanje kolokatorjev glede na slovnične kategorije

glasben

glasbena [oddaja, televizija, tv, program], glasbena [vzgoja, pedagogika]
glasbeni [večer, festival, dogodek, nastop], glasbeni [pouk]
glasbeni [izvajalec, umetnik, ustvarjalec], glasbeni [talent, zvezda, gost]

spomin

[lepi, nostalgični, prijetni, nepozabni, mladostni] spomini; [obujati, obuditi, zbujati, priklicati, prebuditi] spomine
[zgodovinski, kolektiven, spoštljiv] spomin, [trajen, živ, večen, neizbrisen] spomin

šola

[poučevati, učiti] na šoli
[vpisati se] na šolo

V kolokacijskem nizu navajamo zaporedoma tudi **variante oblike**, npr. stisniti za [ramo, ramena], kamor sodi tudi zapis dovršnih in nedovršnih oblik kolokatorjev: [obujati, obuditi] spomin, [ohranjati, ohraniti] spomin, [skrivati, skriti] pred javnostjo; [obujati, obuditi, zbujati, prebuditi] spomine.



Stalne besedne zveze razumemo kot LE, zato zanje posebej navajamo tipično besedilno okolje v obliki kolokacij, če obstaja (gl. poglavje o stalnih besednih zvezah), podobno tudi v primeru skladenjskih zvez, zlasti ko gre za strukturno ustaljene dele, navajamo v obliki kolokacij njihovo tipično besedilno okolje (gl. poglavje o skladenjskih zvezah).

Ni redek primer, ko tudi **kolokator** sam nastopa **kot stalna besedna zveza**, npr. v spodnjih zgledih.

Zgled 37: Večbesedni kolokatorji

pasti na [pokrov motorja]
[berlinski zid, železna zavesa] pade
[minuta molka] v spomin
roditi se z/s [Downovim sindromom]

Izstopanje kolokatorja po pogostnosti v različnih skladenjskih relacijah, ki ga izkazujejo besedne skice, je navadno znak, da nastopa kolokator pri različnih pomenih preučevane besede. Take »večpomenske« kolokatorje navajamo v kolokacijkem nizu pri vsakem pomenu. V primeru glagola *stisniti* izstopajo kolokatorji *roka, dlan, prsti* v različnih skladenjskih strukturah znotraj različnih pomenov, pri samostalniku *spomin* se v različnih pomenskih vlogah pojavljajo glagolski kolokatorji kot: *obujati, priklicati, ohranjati, vtisniti* ipd.

Zgled 38: Večpomenski kolokatorji

stisniti

prijeti

stisniti [*dlan*, *vrat*, *roko*]

stisniti v [*dlan*, *roko*]

stisniti s/z [*prsti*, *roko*, *dlanjo*]

stisniti za [*ramo*, *ramena*, *roko*, *prste*, *vrat*, *boke*]

oblikovati pest

stisniti [*pest*, *prste*]



stisniti [*dlan*, *roko*] v [*pest*]
poškodovati
stisniti komu [*roko*, *nogo*]
priviti k telesu
stisniti [*roke*, *noge*] k [*sebi*, *prsim*]
dati
stisniti v [*roko*, *žep*, *dlan*]

2.5 Frazeologija in idiomatika

V samostojnem razdelku LBS (tj. neodvisno od pomenske členitve) beležimo frazeologijo – frazeološke in idiomatske enote). Gre za besedne zveze in stavke, ki izkazujejo glede na sestavne dele samostojni, neodvisni pomen. Ker se za FE navadno zahteva tudi prenesenost pomena, ki jo izkazuje zveza kot celota, kar ji daje glede na nefrazeološka večbesedna poimenovanja v besedilu ekspresivno oziroma zaznamovano noto, govorimo v zvezi s FE o t. i. *frazeološkem pomenu* (Gantar 2007: 133). Zadnje iz FE izloča vsa večbesedna terminološka poimenovanja, ne glede na sicer izkazan pomenski prenos, npr. *črna skrinjica* – 'naprava'. Z navajanjem frazeoloških in idiomatskih enot na samostojnem nivoju LBS se namenoma izognemo pomenskemu členjenju FE glede na povezanost njenih sestavin z registriranimi pomeni iztočnice, saj take povezave temeljijo na osebnih jezikovnih in kulturnih izkušnjah vsakega posameznega govorca (Cowie 1998: 215). Pomensko podrejanje FE glede na pomen njihovih sestavin bi bilo tako za hierarhizacijo FE znotraj leksikalne baze pa tudi znotraj slovarskih priročnikov nezanesljivo in nujno nekonsistentno, posledično pa zavajajoče tudi za uporabnika.³⁰ S tem ko FE in idiomatskih enot ne uvrščamo pod posamezne pomene besed v iztočnici, ohranjamo tudi razliko glede na stalne besedne zveze, ki so

³⁰ Kot je na eni strani očitno, da obstaja določena semantična povezava med pomenom sestavine in celostnim frazeološkim pomenom zveze, je na drugi strani jasno, da je status take pomenske udeležbe lahko različen in ga je težko natančno določiti, posledično pa tudi zanašanje na semantično analizo pri pomenski opredelitvi FE in njihovi klasifikaciji na pomensko motivirane in nemotivirane ni zanesljivo (Čermák 2001: 11).



vsaj deloma pomensko prozorne (s čimer upravičujemo njihovo pomensko podrejanje) in hkrati ne izkazujejo celostne ekspresivne vloge v besedilu.

Delitev frazeologije na frazeološke in idiomatske enote je v LBS prilagojena namenu leksikalne baze in ne sledi v celoti prevladujočim frazeološkim teorijam (prim. Gantar 2007: 72 in dalje), kjer izraz idiom, kadar se uporablja v odnosu do drugih FE oz. frazema, navadno označuje skrajno točko pomensko osamosvojene leksikalne in ne vedno nujno večbesedne enote.³¹ Pri sestavljanju LBS ne težimo k opredeljevanju FE glede na stopnjo pomenske prozornosti oz. neprozornosti, glede na obseg frazeološkega področja in posledično glede na frazeologijo v ožjem ali širšem smislu, pač pa smo večbesedne, pomensko neprozorne zveze z izkazano metaforično oz. ekspresivno vlogo v besedilu ločili glede na to, ali opravljajo v stavku vlogo stavčnega člana (t. i. FE – pretežno, a ne nujno z zgradbo nestavčne besedne zveze – za izjeme gl. v nadaljevanju) ali pa so same na sebi »zaključena« besedila, ki imajo navadno strukturo stavka. Sem (med t. i. idiomatske enote) prištevamo pregovore in reke (gl. primere v nadaljevanju), izraze s pretežno pragmatično vlogo, ki se kaže v vrednotenju ali nanašanju bodisi na vsebino sporočanega (na t. i. ubesedeno predmetnost), npr. *če me spomin ne vara*, bodisi v vrednotenju udeležencev, okoliščin, splošnih nazorov in prepričanj, vezanih na pripadnike določene jezikovne skupnosti (tj. na zunajjezikovno predmetnost), npr. *boli koga kurac (za koga/kaj)*. Poleg omenjenega prištevamo sem tudi ustaljene konverzijske formule in izraze, kot so pozdravi, opozorila, kletvice ipd (prim. Columas 1981).

2.5.1 Pomenska členitev frazeoloških in idiomatskih enot

Vsako frazeološko in idiomatsko enoto v LBS opredelimo s pomenskim indikatorjem, ki ima enako vlogo kot pri enobesednih iztočnicah, kar pomeni, da na kratko opredeli pomen FE, npr. *stisniti zobe* – 'potrpeti', *stisniti koga v kot* – 'spraviti v brezizhoden položaj', *stisniti glave* – 'sestati se in skupaj razmisliti', *stisniti pas* – 'varčevati', *stisniti ustnice* – 'gesta kot znak jeze, nestrinjanja ali

³¹ Zlasti ameriška generativna frazeološka teorija ne izključuje tudi enobesednih idiomov (prim. tudi Čermák 2001: 8).



kljubovanja', oz. vzpostavlja razliko med pomeni večpomenske FE, npr. *stisniti pest/pesti* – 1. 'gesta kot znak upora, pripravljenosti na boj, 2. – 'gesta kot znak podpore, upanja na uspeh', *dol pasti* – 1. biti presenečen, začuden, 2. – 'biti utrujen', 3. – 'omagati, izgubiti zavest', *zastor/zavesa pade* – 1. 'predstava se zaključi', 2. – 'dogodek, proces itd. se konča'.

Opredelitev pomena se pri idiomatskih enotah, ki so strukturno zaključena besedila – navadno z izkazanimi nanašalnimi besedami tipa, *ta je pa bosa*, *to/kaj ni mačji kašelj* ipd. – razlikuje od pomenskih indikatorjev, kot jih uporabljamo pri pomenskem razdvoumljanju LE, stalnih besednih zveze in FE nestavčnega tipa, saj zahteva pojasnitev okoliščin in vrednotenja, na katerega se idiomatske enote nanašajo, npr. *kocka je padla* – 'izraža da je bila odločitev dokončno sprejeta'; *kdor visoko leta, nizko pade*; 'izraža, da imajo velika pričakovanja ali precenjevanje lastnih zmožnosti negativne posledice', *če me spomin ne vara* – 'izraža omilitev, dvom v trditve'.

2.5.2 Zapisovanje frazeoloških enot

FE, ki vsebuje besedo iz iztočnice, navajamo v frazeološkem razdelku pri vseh predmetnopomenskih besedah, ki jih vsebuje, npr. FE *princ na belem konju*, navajamo pri iztočnicah *princ*, *konj* in *bel*, če so del jedrnega besedišča oz. geslovnika. Pri sestavinah, ki so nejedrni ali variantni deli FE, npr. *čakati na (svojega) princa na belem konju*, *sanjati o princu na belem konju* ipd. navedemo FE pri glagolskih sestavinah le, če pride na statistično površje v leksikalno-gramatičnem profilu, ki ga izkazujejo besedne skice.

Za številne FE je značilno, da izkazujejo *variante* bodisi **(a)** sestavin, npr. *luč na koncu tunela/predora*, *zastor/zavesa pade*, *stopicati/capljati na mestu*, **(b)** besednih oblik, npr. *stisniti pest/pesti*, ali **(c)** zunanjih udeležencev, npr. *biti trn v peti komu/koga*, *pasti v kremplje komu/čemu/koga*. Te vrste variante navajamo znotraj ene same frazeološke enote za poševnico.

Nekoliko drugače obravnavamo pri zapisu FE v programskem vmesniku t. i. *variantne oblike rabe* in *pretvorbene možnosti* FE. V prvem primeru gre za variantnost besednih oblik, ki so najpogosteje



posledica oblikoslovnega prilagajanja znotraj konkretne skladijske kombinacije v FE. Tipični primeri so, ko katera od sestavin FE, npr. variantni predlog, glagol ipd., zahteva drugačno sklonsko obliko določila, npr. *stisniti skozi zobe* in *stisniti med zobmi*, *stisniti koga v kot* in *stisniti koga ob zid*, *zaiti v slepo ulico* in *znajti se v slepi ulici*, ali ko variantnost oblike vpliva na razporeditev udeleženskih vlog, ki jih predvideva FE, kot v primeru: *srce se stisne komu* in *kaj stisne srce komu*.

O pretvorbenih možnostih FE govorimo v primerih, ko gre za:

(a) različne skladijske oz. stavčnočlenske funkcije FE, npr. *beraška palica – spraviti koga na beraško palico*; *zdrava pamet – biti zdrave pameti/pri zdravi pameti – (narediti kaj) po zdravi pameti*. Pri tem se zavedamo, da pri glagolskih rabah oz. FE v vlogi povedka na pomenske odtenke vplivajo konkretni glagoli, npr. *princ na belem konju – čakati na princa na belem konju*, *iskati*, *najti (svojega) princa na belem konju – sanjati o princu na belem konju*;

(b) različne realizacijske možnosti,³² ki jih omogoča zgradba FE, zlasti posamostaljenje, npr. *prati možgane – pranje možganov*, *pihati na dušo – pihanje na dušo*, *stopicati na mestu – stopicanje na mestu*, in potrpnjenje, npr. *stresati/zbijati šale – šale se stresajo/zbijajo*; *dvigati prah – prah se dviga*, *dati besedo – beseda je dana* ipd.;

(c) različno število in razporeditve »zunanjih« udeležencev, npr. *prepustiti se na milost in nemilost komu/čemu – prepustiti koga na milost in nemilost komu/čemu – biti prepuščen na milost in nemilost komu/čemu*;

(d) tipične realizacije v kateri od upovedovalnih oz. naklonskih možnosti, kot so zanikanje, velelnost, vprašalnost ipd., npr. *kaj (ne) pade na pamet/misel/glavo/um komu* in *kaj ne pade niti na kraj pameti komu*, *saj nisem na glavo padel/la* ipd. Tu beležimo tudi tipične realizacije FE v določenem slovničnem času, npr. *čas je povozil koga*;

³² V primeru, ko se katera od pretvorb izkaže kot edina možna, npr. *kopja se lomijo* (*lomiti kopja), navajamo samo to obliko.



(e) fakultativne sestavine FE: *pasti kot (zrela) hruška in pasti komu (v naročje) kot (zrela) hruška, podajati si kaj (iz rok v roke) kot vroč krompir, biti (vedno) korak pred časom*; tudi kadar so opsijske sestavine »zunanji« udeleženci, npr. *pasti na limanice (komu)* ipd.

Pretvorbene možnosti in vseh vrst variante seveda beležimo le, če se izkazujejo kot tipične in za posamezno FE relevantne.

Pri tem se držimo pravila, da kot variantne oblike rabe ter pretvorbe beležimo znotraj enega nadelementa frazeološka enota zaporedoma v elementih enota, če ne prihaja do bistvenih pomenskih sprememb, kar pomeni, da gre za dve ali več realizacijskih možnosti ene same FE. Večpomenske FE, npr. *pasti iz nebes/neba/vesolja* – 'nenadoma se pojaviti' in 'biti še neuveljavljen, nepoznan' ter *kaj ne pade z/liz neba* – 'kaj ne pride samo od sebe in brez truda', ter sinonimne FE, npr. *pasti na rit* in (skoraj) *pasti s stola* – 'biti presenečen'; *sesti skupaj in stisniti glave* – 'sestati se in skupaj razmisliti', kjer gre v obeh primerih za isti pomen, beležimo kot samostojne FE, torej v samostojnem nadelementu, saj predvidevamo, da lahko tudi večpomenske in sinonimne FE izkazujejo lastne variante in pretvorbene možnosti.

Na enak način kot FE beležimo v samostojnem razdelku LBS – tj. neodvisno od pomenske členitve – idiomatske enote. Kot rečeno, prištevamo sem FE s stavčno zgradbo, t. i. *pregovore* in *reke* (razen glagolskih FE tipa: *zob časa načne kogalkaj, mrak pade na oči komu, sekira pade v med komu, kaj pade v vodo* ipd., ki se v stavku obnašajo kot povedki).³³ Za pregovore in reke velja, da so »besedila že sama na sebi«, hkrati pa je zanje značilno, da se nanašajo na konkretno temo ali referenta, izraženega v širšem (so)besedilu, npr. *iz te moke ne bo kruha, jabolko ne pade daleč od drevesa, čas zaceli vse rane* itd.

³³ V slovenističnem jezikoslovju se je za FE tega tipa uveljavil izraz »nepravi« glagolski frazemi (prim. Gantar 2007: 105 in dalje).



2.6 Primeri rabe

S primeri rabe, ki v celoti temeljijo na korpusu FidaPLUS in jih v programski vmesnik prenašamo s pomočjo aplikacije Sketch Engine (SkE: podrobneje v Kazalniku 5) potrjujemo vse informacije, ki smo jih na višjih nivojih vnesli v LBS: za vsak zabeležen pomen LE morajo primeri, ki odražajo realno jezikovno rabo, potrditi s pomenskim indikatorjem opredeljen pomen, pomensko shemo z obveznimi in opcijskimi udeleženci, skladišne strukture in skladišne vzorce pri glagolskih iztočnicah. Najbolj se primerom rabe približajo kolokacije, ki so tipične potrditve skladišnih struktur ali skladišnih vzorcev. S primeri rabe vedno potrdimo tudi realno obnašanje stalnih besednih zvez ter frazeoloških in idiomatskih enot v besedilu. V hierarhično urejenih nivojih LBS tako predstavljajo primeri rabe najkonkretnjšo informacijo in skupaj s pomenskimi, skladišnimi in kolokacijskimi podatki o LE zaključujejo celoto, znotraj katere so informacije med seboj neposredno povezane in se medsebojno dopolnjujejo.

Zgled 39: Leksikalno-gramatične informacije v LBS: od abstraktne pomenske členitve do konkretnih primerov rabe

pasti

POMENSKI INDIKATOR: 1.2 izgubiti izpustiti

POMENSKA SHEMA: PREDMET pade ČLOVEKU KAM |ali| MU pade od KOD

MEGASTRUKTURA: sbz1 GBZ sbz3 (iz sbz2) v sbz4/na sbz4/pod sbz4

VZOREC: kaj pade komu v kaj

VZOREC: kaj pade komu na kaj

VZOREC: kaj pade komu pod kaj

VZOREC: kaj pade komu iz česa

KOLOKACIJA: [pištola, cigareta] pade

KOLOKACIJA: pasti v [jarek, jašek]

KOLOKACIJA: pasti na[tla]

KOLOKACIJA: pasti pod [mizo]

KOLOKACIJA: pasti iz [rok, roke, ust]

KOLOKACIJA: pasti iz [žepa]



- ZGLED:** Vžigalice so ji padle v cestni jarek in so zdaj vse premočene.
ZGLED: Fantku je plastični tovornjaček padel v jašek.
ZGLED: Pištola mu je med prerivanjem padla na tla.
ZGLED: Če vam hrana pade na tla, je lahko okužena v petih sekundah.
ZGLED: Vse kaže, da mi je padlo pod mizo deset šilingov.
ZGLED: Poberem knjigo, ki mi je padla iz rok, in nadaljujem branje.
ZGLED: Sneguljčici zastrupljeni košček jabolka pade iz ust in dekle oživi.
ZGLED: Med begom mu je ukradena denarnica padla iz žepa.

Pri izbiri primerov stremimo k relativni kratkosti. V primeru, da je izbrani primer mogoče krajšati (npr. izpustiti obsežen vrinjen stavek ipd.) krajšanje ustrezno označimo. Če je mogoče, vendar ne za vsako ceno, se izogibamo lastnih imen, predvsem pa smo pozorni na to, da izbiramo primere, ki ne vnašajo dvoma v informacije, ki jih želimo s primeri potrditi. Izogibamo se denimo primerom rabe, v katerih je obravnavana LE uporabljena bodisi v katerem od bolj ali manj individualnih pomenskih odtenkov oz. primerom, v katerih se obravnavana LE pojavlja v več pomenih. S primeri rabe želimo zajeti tudi tipično besedilno rabo (npr. diskurz, premi govor ipd.), tip besedila (npr. osmrtnice), področje rabe (šport) in skladiščno okolje, če ga ne moremo zajeti s skladišnimi strukturami. Pri navajanju števila primerov nismo omejeni, pomembno pa je, da je vsako skladiščno strukturo in kolokacijo mogoče videti v njenem tipičnem življenjskem okolju.

V programskem vmesniku označimo besedo, ki jo obravnavamo v iztočnici z elementom >i<.

Slika 20: Zgledi rabe za podpomen samostalnika *rana* v programskem vmesniku



POMEN

INDIKATOR psihična ali čustvena bolečina

POMENSKA_SHEMA

SKLADENJSKE_SKUPINE

SKLADENJSKA_STRUKTURA

STRUKTURA pbz SBZ1

KOLOKACIJA (K) duševna, srčna (/k) rana

KOLOKACIJA (K) boleča (/k) rana

ZGLEDI

ZGLED Da bi se ji po ločitvi od Brada Pitta zacelile srčne (I) rane (/i), se je zatekla k budizmu.

ZGLED Žaljive besede njenega ljubega so zadale Francki hudo duševno (I) rano (/i).

ZGLED Po njegovem prepričanju ne bi smeli brskati po bolečih (I) ranah (/i) preteklosti, ker bi to lahko razdelilo narod, ki mora biti v novih časih čim bolj složen.

3 LITERATURA

- ARHAR, Š., GORJANC, V., 2007: Korpus FidaPLUS: nova generacija slovenskega referenčnega korpusa. *Jezik in slovstvo* 52 (2). 95–110.
- ATKINS, S., B. T., RUNDELL, M., 2008: *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- BAKER, M., FRANCIS, G. in TOGNINI-BONELLI, E., 1993. *Text and Technology: In Honour of John Sinclair*. Amsterdam. Philadelphia: Benjamins.
- BRIDGE: *Collins Cobuild Angleško slovenski slovar BRIDGE*, 2000. Ljubljana: DZS.
- CHURCH, K. W., HANKS, P., 1990: Word Association Norms, Mutual Information and Lexicography. *Computational Linguistics* 1. 22–29.
- COLUMAS, F., 1981: *Conversational Routine*. The Hague: Mouton.
- COWIE, A., P., 1981a: The treatment of collocations and idioms in learners' dictionaries. A. P. Cowie (ur.): *Lexicography and its Pedagogical Applications*. Applied Linguistics 2/3 (thematic issue).
- COWIE, A., P., 1999: *English Dictionaries for Foreign Learners*. Oxford: Oxford University Press.



- ČERMÁK, F., 1985: Frazeologie a idiomatika. František Čermák, Josef Filipec: *Česká lexikologie*. Praha: Academia. 166–248.
- ČERMÁK, F., 2001: Substance of idioms: perennial problems, lack of data or theory? *International Journal of Lexicography* 14 (1). Oxford: Oxford University Press. 1–20
- ČERMÁK, F., 2006: Collocations, Collocability and Dictionary. Corino et. al (ur): *Proceedings of 12th EURALEX International Congress, EURALEX 2006*. Alessandria: Edeizioni Dell'Orso. 431–438.
- ČERMÁK, F., 2009: *Lexicographer's Notes Related to a Corpus-Based Dictionary*. Predavanje na posvetu: Slovarji, več kot le besede, Ljubljana, 6. 2. 2009; <http://www.slovenscina.eu/Vsebine/En/Dogodki/Slovarji/Program.aspx>; JiS – v tisku).
- FidaPLUS: *Korpus slovenskega jezika FidaPLUS*: <http://www.fidaplus.net/>.
- FILLMORE, Ch., J., JOHNSON, C., R., in PETRUCK, M., R. L., 2003: Background to Framenet. *International Journal of Lexicography* 16/3. 235–250.
- FIRTH, R. 1957. *Papers in Linguistics*. London: Oxford University Press.
- FRANCIS, G., HUNSTON, S. and MANNING, E. 1998. *Collins Cobuild Grammar Patterns 1: Verbs*. London: HarperCollins.
- FRANCIS, G., HUNSTON, S. and MANNING, E. 1998. *Collins Cobuild Grammar Patterns 2: Nouns and Adjectives*. London: HarperCollins.
- GANTAR, P., 2007: *Stalne besedne zveze v slovenščini: korpusni pristop*. Ljubljana: Inštitut za slovenski jezik Frana Ramovša ZRC SAZU.
- GEERAERTS, D., 1990: The Lexicographical Treatment of Prototypical Polysemy. Tsohatzidis, S. L. (ur.): *Meanings and Prototypes*. London. Routledge. 195–210.
- GRABNAR, K., ŠORLI, M., 2003: Novi veliki angleško-slovenski slovar Oxford-DZS. *Jezik in slovstvo* 48 (3/4). 126–133.
- GORJANC, V., KREK, S. in GANTAR, P., 2005: Slovenska leksikalna podatkovna zbirka. *Jezik in slovstvo* 50 (2). 3–19.
- HANKS, P., 2000: Contributions of Lexicography and Corpus Linguistics to a Theory of Language Performance. Heid, U., Evert, S., Lehmann, E. in Roher, C. (ur.): *Proceedings of the Ninth EURALEX Congress*. Stuttgart: University of Stuttgart. 3–14.



- HANKS, P., 2004: *Corpus Pattern Analysis*. Williams, G. in Vessier, S. (ur.): *EURALEX 2004 Proceedings*. Lorient: Université de Bretagne-Sud.
- HANKS, P., 2006: *The organization of the Lexicon: Semantic Types and Lexical Sets*. Corino, E.; Marello, C.; Onesti, C. (ur.): *Atti del XII congresso internazionale di lessicografia: proceedings XII EURALEX international congress, Torino, Italia September 6th-9th, 2006*. Torino: Edizioni dell'Orso: Università di Torino: Accademia della Crusca. 1165–1168.
- HANKS, P., 2008: *Mapping meaning onto use: a Pattern Dictionary of English Verbs*. *AAFL 2008, Utah*.
- HANKS, P., 2009: *Compiling a Monolingual Dictionary for Native Speakers*. Predavanje na posvetu: Slovarji, več kot le besede, Ljubljana, 6. 2. 2009; <http://www.slovenscina.eu/Vsebine/En/Dogodki/Slovarji/Program.aspx>; JiS – v tisku).
- HANKS, P., PUSTEJOVSKY, J., 2004: *Common Sense About Word Meaning: Sense in Context*. TSD 004. Brno.
- HANKS, P., PUSTEJOVSKY, J., 2005: *A Pattern Dictionary for Natural Language Processing*. *Revue Française de Linguistique Appliquée* 10 (2). 63–82.
- HAUSMANN, F., J., 1999: *Le dictionnaire de collocations – Critères de son organisation*. N. Greiner et al. (ur.): *Texte und Kontexte in Sprachen und Kulturen. Festschrift für Jorn Albrecht*. Trier: Wissenschaftlicher Verlag. 121–140.
- HAUSMANN, F., J., 2003: *Was sind eigentlich Kollokationen?* Steyer, K. (ur.): *Wortverbindungen – mehr oder weniger fest*. (Jahrbuch des Instituts für deutsche Sprache.) Berlin: De Gruyter. 309–334.
- HOEY, M., 1998: *Introducing Applied Linguistics: 25 Years on*. Plenary Paper in the 31st BAAL Annual Meeting: *Language and Literacies*. University of Manchester, September 1998.
- HOEY, M., P., 2000: *A world beyond collocation: new perspectives on vocabulary teaching*. Michael Lewis (ur.): *Teaching collocation: Further Developments in the Lexical Approach*. London: LTP. 224–244.
- HUNSTON, S. in FRANCIS, G., 1999: *Pattern Grammar: a corpus-driven approach to the lexical grammar of English*. *Studies in Corpus Linguistics* 4. Amsterdam: John Benjamins.



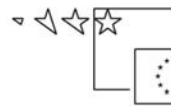
- JACKENDOFF, R., 1995: 138 : *Languages of the Mind: Essays on Mental Representation*. Cambridge: MIT.
- JOHANSSON, S., 1998: Goran Kjellmer. A Dictionary of English Collocations, Based on the Brown Corpus. 3 vols. Oxford: Clarendon Press, 1994. *International journal of Corpus Linguistics* 3 (2). 338–348.
- KILGARRIFF, A., 2006: Collocationality (and how to Measure it). Corino, E.; Marello, C.; Onesti, C. (ur.): *Atti del XII congresso internazionale di lessicografia: proceedings XII EURALEX international congress, Torino, Italia September 6th-9th, 2006*. Torino: Edizioni dell'Orso: Università di Torino: Accademia della Crusca.
- KILGARRIFF, A., 2006a: Word Senses. E. Agirre, P. Edmonds (ur.): *Words Sense Disambiguation: Algorithms and Applications*. New York: Springer. 29–45.
- KJELLMER, G. 1994. *A Dictionary of English Collocations*. Oxford: Clarendon Press.
- KREK, S. (ur.): 2005-6: *Veliki angleško-slovenski slovar Oxford*. Ljubljana: DZS.
- KREK, S., 2008: FrameNet in slovenščina. *Jezik in slovstvo*, 53/5. 37–54.
- KREK, S. in KILGARRIFF, A., 2006: Slovene Word Sketches. Erjavec, Tomaž in Gros, Jerneja (ur.): *Jezikovne tehnologije 5*. Ljubljana: IJS. 62–65.
- KOCJANČIČ, P. , KREK, S. in ŠORLI, M., 2008: The Funny Mirror of Language: the Process of Reversing the English-Slovenian Dictionary to Build the Framework for Compiling the New Slovenian-English Dictionary. Elisenda Bernal, Janet DeCesaris (ur.): *Proceedings of the XIII EURALEX International Congress* (Barcelona, 15-19 July 2008). Barcelona: Documenta Universitaria: Institut universitari de lingüística aplicada, Universitat Pompeu Fabra.
- LDOCE: *Longman Dictionary of Contemporary English*, 2003⁴. Harlow: Pearson Education Ltd.
- MEER, G. van der, 2006: It's about Time: On Coherence and Simplicity in Dictionary Entries. *English Studies*. Routledge. 602–616.
- MED: *Macmillan English Dictionary for Advanced Learners*, 2007². Oxford: Macmillan Publishers Ltd.
- MEL'ČUK, I. 1998: Collocations and Lexical Functions. A. Cowie (ur.): *Phraseology, Theory, Analysis and Applications*. Oxford: Clarendon Press. 23–53.



- MOON, R. E., 1987: *The Analysis of Meaning*. Sinclair, J. M. (ur.): *Looking up: A Account of the COBUILD Project in Lexical Computing*. London: Collins.
- NODE: *The Oxford Dictionary of English*, 1998². Oxford: Oxford University Press.
- OCD: *Oxford Collocations: Dictionary for students of English*, 2003⁴. Oxford: Oxford University Press.
- OOI, B. Y., V. 1998: *Computer Corpus Lexicography*. Edinburgh: Edinburgh University Press.
- PUSTEJOVSKY, J. 1995. *The Generative Lexicon*. Chicago: The MIT Press.
- SIEPMAN, D., 2005: Collocation, colligation and encoding dictionaries. Part I: Lexicological Aspects. *International Journal of Lexicography* 18/4 . 409–444.
- SIEPMAN, D., 2006: Collocation, colligation and encoding dictionaries. Part II: Lexicographical Aspects. *International Journal of Lexicography* 19/1 .1–39.
- SINCLAIR, Mc., J., COULTHARD, M., 1975: *Towards an Analysis of Discourse: The English Used by Teachers and Pupils*. Oxford: Oxford University Press.
- SINCLAIR, J. M., 1987: Collocation: a progress report. R. Steele and T. Threadgold (ur.).
- SINCLAIR, J. M. (ur.), 1987a: *Looking up. An Account of the COBUILD Project in Lexical Computing*. London and Glasgow: Collins.
- SINCLAIR, J. M., 1991: *Corpus, Concordance, 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.
- ŠORLI, M. et al., 2006: *The Oxford-DZS Comprehensive English-Slovenian Dictionary*. Corino, E.; Marello, C.; Onesti, C. (ur.): *Atti del XII congresso internazionale di lessicografia: proceedings XII EURALEX international congress*, Torino, Italia September 6th-9th, 2006. Torino: Edizioni dell'Orso: Università di Torino: Accademia della Crusca. 631–637.
- TOPORIŠIČ, J., 1982: *Nova slovenska skladnja*. Ljubljana: DZS.
- TOPORIŠIČ, J., 1992: *Enciklopedija slovenskega jezika*. Ljubljana: Cankarjeva založba.

Spletne strani

<http://www.sil.org/linguistics/GlossaryOfLinguisticTerms/Index.htm>.



[http://www.thefreedictionary.com/lexical+database.](http://www.thefreedictionary.com/lexical+database)
<http://www.slovenscina.eu/Vsebine/En/Dogodki/Slovarji/Program.aspx>
<http://www.owid.de/elexiko /index.html>
<http://www.ilc.cnr.it/clips/quattrolivelliweb.ppt>
<http://www2.let.vu.nl/oz/clt/cornetto/index.html>
<https://ilt.kuleuven.be/blf/>
<http://www.kuleuven.be/alfalex/index.php?id=&ng=0>
<http://www.cst.dk/cgi-bin/sto/defisto>
<http://adesse.uvigo.es/>
<http://grial.uab.es/>
<http://www.staff.amu.edu.pl/~zlisi/projects/ceglex/index.en.html>
<http://spraakbanken.gu.se/>
<http://deb.fi.muni.cz/install.php>
<http://nl.ijs.si/jos/>
<http://www.natcorp.ox.ac.uk/>
<http://nlp.fi.muni.cz/projekty/cpa/>
<http://nl.ijs.si/ME/>