



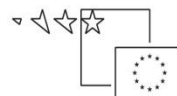
Kjer se srečata pomen in skladnja:

leksikalna baza za slovenščino kot vir podatkov za
pedagoško korpusno slovnico

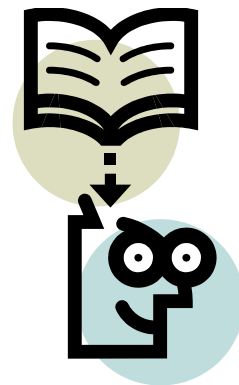
Polona Gantar
Inštitut za slovenski jezik FR
Apolonija.Gantar@guest.arnes.si



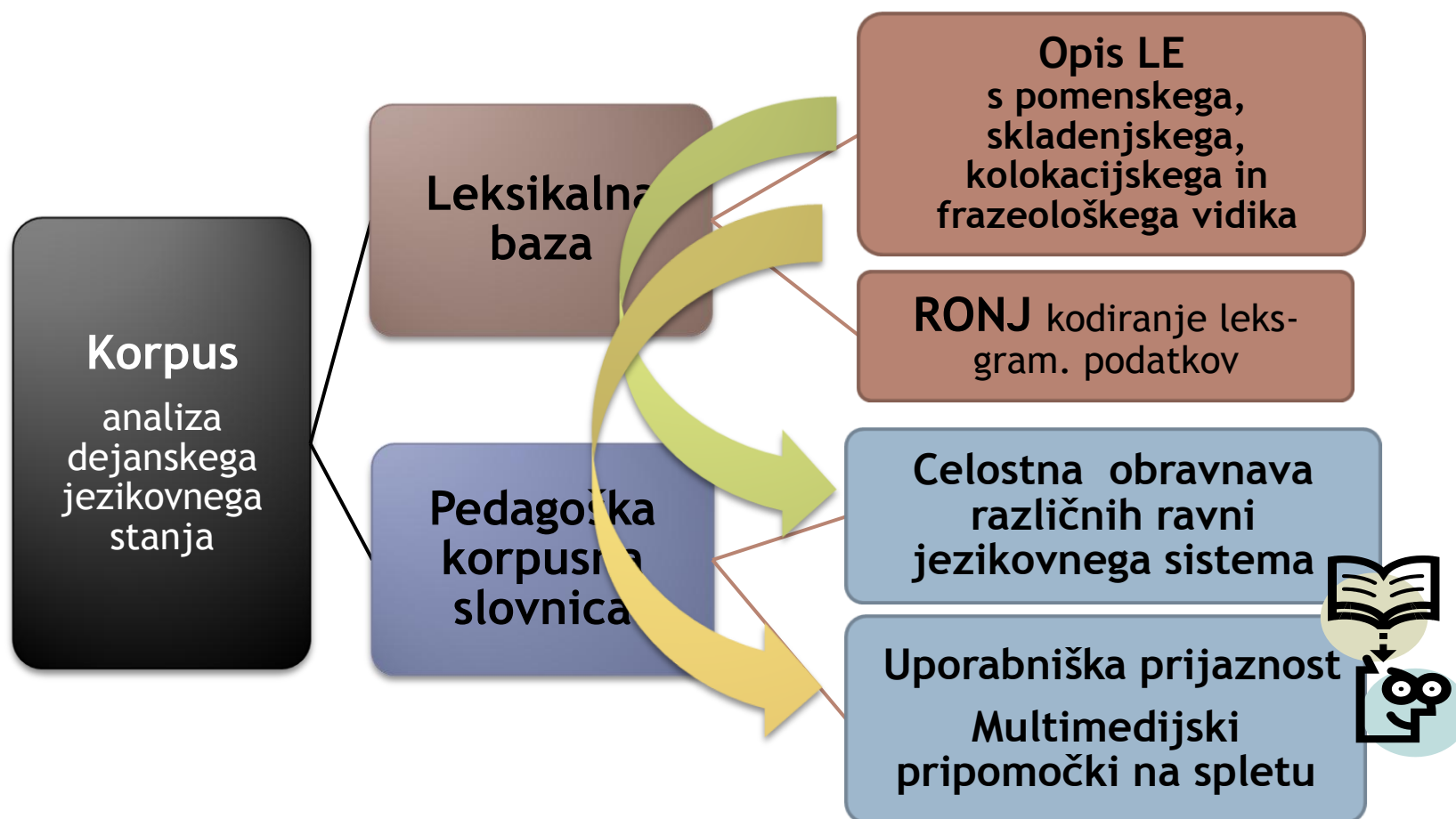
REPUBLIKA SLOVENIJA
MINISTRSTVO ZA ŠOLSTVO IN ŠPORT



Naložba v vašo prihodnost
OPERACIJO DELNO FINANCIRA EVROPSKA UNIJA
Evropski socialni sklad



Slovnica in LBS v projektu SSJ





ZA
KAJ

priročniki



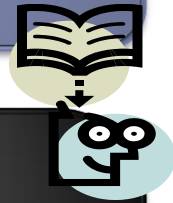
ZA
KOGA

splošni uporabnik
šolski uporabnik
učenec slovenščine kot TJ



KAJ

pomenski opis
kolokacije
korpusni zgledi



računalniška obdelava
jezika



računalnik
jezikoslovec



**Slovnični podatki v
LBS**

I. LEMA

- iztočnica
- besedna vrsta

svitati se
glagol



II. POMEN

- indikator

1. daniti se

2. dojemati

- pomenska shema

KO se svita **DAN**,
začne vzhajati sonce

če se **ČLOVEKU** začne svitati o nekem **DOGAJANJU**, začne dojemati, kar prej ni vedel, ali pa je bilo to pred njim skrito

III. SKLADNJA

- oznaka
- struktura
- vzorec
- sklad. zveza

v. 3. osebi
gbz Inf-GBZ
kaj se svita

rbz GBZ
komu se svita, zakaj
komu se svita o čem

IV. KOLOKAC.

- kolokacija

[začeti, pričeti] se svitati

[počasi, malo, malce] se svita

V. ZGLEDI

- zgled

Preden se začne zjutraj svitati, je najtemnejša noč.

Počasi se mi je začelo svitati, zakaj Jasni oči tako žarijo.

Na vzhodu se je že svital dan, ko sta se poslovila.

Petru se pričinja svitati o nekdanji zvezi ned Chadom in Heather.

- stalna zveza

VI. FRAZEOLOGIJA • frazeološka enota



besedna vrsta

- samostalnik, glagol, pridevnik, prislov
- konverzija BV

pri (pod)pomenu

debel prid.

1 ki ima preveliko telesno težo

1.1 kdor ima preveliko telesno težo

Oznaka: **v samostalniški rabi**

pri skladenjski strukturi

celofan sam.

1 ovojni papir
celofan [papir]

SBZ1<r>**v pridevniški rabi**</r>sbz0

iztočnica	slovnična (pod)kategorija po SS	LBS	
		besedna vrsta	oznaka ali restrikcija pri strukturi
<i>iskanje</i>	glagolnik	samostalnik	
<i>fitnes, kaki, latino</i>	neskl. prilastek		v pridevniški rabi
<i>blagor, balzam,</i>	povedkovnik		v povedni rabi
<i>vrh, konec, dno</i>	sam. v prislovni vlogi		v prislovni rabi
<i>stisnjen</i>	deležnik	pridevnik	
<i>pričakovan, debel</i>	posam. pridevnik		v samostalniški rabi
<i>všeč, rad</i>	povedkovnik	prislov	v povedni rabi
<i>naprodaj</i>			
<i>miže</i>	deležje		

Slovnični podatki na ravni *pomena*

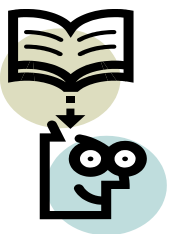
- Katere slovnične podatke upoštevamo v LBS
- Kako razumemo slovarski pomen (opis pomena)
- Kako razumemo pomen



- **Katere slovnične podatke upoštevamo v LBS**
 - Slovnične podatke, ki so neposredno povezani s pomenom
- **Kako razumemo formalizacijo slovarskega pomena (oblikovanje slovarskih razlag)**

DA: Govorci v situaciji vsakodnevne komunikacije ne razmišljamo o konceptu besednega pomena, o tem razmišljajo leksikografi (Rundell, 2010)

NE: Strukturiranost leksikalnega pomena odraža strukturiranost človekovih predstav o zunajjezikovni stvarnosti (Vidovič Muha, 2000)



• Kako razumemo “pomen”

- Ne verjamemo v besedni pomen (Sinclair, Atkins, Kilgarriff, Hanks idr.)
 - Pomen kot zaključena enota ne obstaja, obstajajo zgolj pomenske tendence.
- Verjamemo, da pomenske tendence besed vplivajo na izbire v svojem besedilnem okolju
 - Vsak pomen se oblikuje v konkretni govorni situaciji in je odvisen od vsakokratnega konteksta.





leksikogramatika

- »Kamalu je postalo jasno, da je oblika dejansko odločilna za pomen /.../ izbira pomena, kjerkoli v besedilu, ima odločilen vpliv na okoliške izbire.«

J. M. Sinclair, *Corpus Concordance Collocation*, 1991

probabilistično obnašanje jezika

- **FrameNet** (Filmore et. al 1976)
 - Pomen je najbolje interpretirati probabilistično, se pravi v smislu najboljšega približka prototipu in ne v smislu iskanja niza zadostnih ali potrebnih pogojev.
- **Corpus pattern Analysis** (Hanks 2004)
 - Pomeni so neposredno povezani z vzorci (ang. patterns), strukturami (ang. constructions) in frazeološkimi enotami.

PASTI 'izgubiti ravnotežje'
'znižati se na lestvici'

*motorist, na vozišče, na glavo, po tleh...
tečaj, cena, pod raven, za x odstotkov ...*

CVETETI 'poganjati cvetove'
'prinašati dobiček'
'izgledati zadovoljen in zdrav'

KAJ cveti + kako/kdaj
KAJ cveti in KAJ cveti KOMU
KDO cveti (od česa)



... kjer se srečata pomen in skladnja

IZTOČNICA

pomenski indikator



pomenski meni



pomenska shema

IMPROVIZIRATI

1. delati brez priprav

1.1 sproti ustvarjati

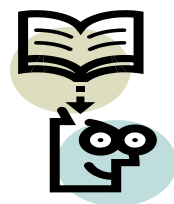
2. izdelati brez primerne orodja ali materiala

če ČLOVEK med nastopom ali igranjem improvizira, si sproti izmišlja ali ustvarja VSEBINO

Sem tiste vrste igralka, ki kdaj tudi malce improvizira na odru.

Vključitev skladnje v pomenski opis

- Empirične raziskave (Rozman et al. 2010, Lew in Dziemianko 2006) so pokazale, da:
- eksplicitno navajanje slovničnih podatkov v slovarju zaradi zahtevnega metajezika ne doseže naslovnikov oz. da naslovniki te podatke pogosto ignorirajo;
- definicije poleg zgledov rabe vsebujejo za uporabnika zelo pomemben vir slovničnih informacij in so za razliko od gramatičnih kod in slovničnih vzorcev za uporabnika veliko bolj dostopne in berljive.





Pomenska shema - 3 naloge

- zapis udeleženske zgradbe pomena
 - prototipični vzorec - “norma”
 - iztočnica + formalizacija udeleženskih mest (RONJ)
- zapis semantičnih tipov, ki zasedajo udeleženska mesta
- pomenski scenarij
 - stavčna razlaga, ki povezuje udeležence in okoliščine v konkretno situacijo, ki je za dani pomen značilna

SESTI 'ustrezati'

Po letih plesa imam izostren občutek za glasbo in ritem. Vem, kaj pali in **kaj** **mi** sede.

1. udeleženska zgradba

KAJ sede **KOMU**

2. semantični tipi

DEJAVNOST, STANJE ali **LASTNOST česa** sede **ČLOVEKU**

3. pomenski scenarij

če neka **DEJAVNOST, STANJE** ali **LASTNOST česa** **ČLOVEKU** sede, mu ugaja ali ustreza

NOTA 'značilnost'

Voda in svetloba
prineseta posebno
noto na vrt in
teraso.

1. udeleženska zgradba

KAJ daje noto ČEMU

2. semantični tipi

LASTNOSTI česa dajejo noto IZDELKU, KRAJU ali
DEJANJEM

3. pomenski scenarij

če LASTNOSTI česa dajejo IZDELKU, KRAJU ali DEJANJEM
svojo noto, se v njem izražajo in ga delajo posebnega

POZOREN 'ustrežljiv'

Ari je bil ves čas zelo pozoren do svoje žene, ona pa mu je s prisrčnim nasmehom vračala naklonjenost.

1. udeleženska zgradba

KDO je pozoren do KOGA

2. semantični tipi

ČLOVEK je pozoren do drugega ČLOVEKA

3. pomenski scenarij

če je ČLOVEK pozoren do drugega ČLOVEKA, je do njega ustrežljiv in mu izkazuje naklonjenost



norma - alternacija - eksploatacija

- P. Hanks, Theory of norms and exploitations, 2004
- Vsaka beseda, ki jo uporabi govorec, da bi z njo nekaj povedal, je uporabljena v skladu z neko **normo** “dogovorom”, ki jo **določa njen stavčni vzorec**.
- Orodja za analizo korpusov kažejo, da je jezikovna raba **vzorčena** in hkrati **variabilna**.
- Korpusna analiza posamezne besede pokaže, da je večina konkordanc realizacija prepoznavnega “**normalnega**” vzorca, druge konkordance so **alternacije tega vzorca**, malo pa je t. i. **eksploatacij**.

SPRATI - 'umiti z vodo'

če ČLOVEK spere PREDMET, OBLAČILO ali ŽIVILO s TEKOČO VODO, z njega odstrani umazanijo ali delce kake snovi

v
z
o
r
c
i

KDO spere KAJ s ČIM

KDO spere KAJ

KDO spere KAJ pod ČIM

KDO spere KAJ v ČEM

Vok pred prvo uporabo speremo z vročo vodo.

Školjke večkrat temeljito speremo in skrtačimo.

Namizni prt sperite pod mrzlo vodo, nato ga namočite.

Čopič dobro sperite v topli vodi.

??

KDO spere KAJ s ČESA s ČIM

- Po kopanju je treba kopalke sprati s sladko vodo, saj **sol** uničuje tkanino.
- Plenice dobro sperite, da odstranite vse **sledi pralnega praška**.

Slovnični podatki - *skladenjski nivo*

- **Vzorci**

- stavčna skladnja (udeleženci znotraj celotnega stavčnega vzorca)
- samo glagol

- **Skladenske strukture**

- besednozvezna skladnja (zapis besedne zveze, v kateri se pojavlja iztočnica, s simboli)
- vse besedne vrste
- število struktur je končno (več kot 290)
- vir: Besedne skice (SkE)



(a1) samostalniške	(a1) glagolske	(a1) pridevniške	(a1) prislovne
SBZ0 num2 petek [trinajstega]	rbz GBZ [nesrečno, hudo] pasti	Vez-gbz PBZ4 [delati se] pomembnega	rbz RBZ [nadvse] preprosto
Vez-gbz SBZ1 [izpasti] bedak	GBZ Inf-gbz dati se [pregovoriti]	PBZ0 sbz0 svetla [prihodnost]	sbz0 RBZ [odhod, pot] domov
pbz0 SBZ0 [boleč] spomin	GBZ Nam-gbz priti [spodbujat]	PBZ0 sbz2 poln [denarja]	RBZ sbz1 pasje [mraz]
pbz0 SBZ3 uporabniku [prijazen]	GBZ Odv-da ... vedeti, da ...	PBZ0 sbz4 dolžen [vsoto, denar]	num1 RBZ [dvakrat] tedensko
(a2) samostalniške	(a2) glagolske	(a2) pridevniške	(a2) prislovne
SBZ0 blizu sbz2 vir blizu [ministrstva]	GBZ ali gbz kupiti ali [najeti]	PBZ0 glede sbz2 negotov glede [izida]	RBZ do sbz2 najpozneje do [marca]
SBZ0 k/h sbz3 pripombe k [osnutku]	GBZ in gbz spotakniti se in [pasti]	PBZ0 do sbz2 kritičen do [udeležencev]	RBZ po sbz5 takoј po [sestanku]
SBZ0 med sbz6 ¹ razlika med [spoloma]	GBZ Inf-gbz sbz1 želeti postati [manekenka]	PBZ0 k/h sbz3 nagnjen k [samomorom]	RBZ s/z sbz6 skladno z [inflacijo]

gbz Inf-GBZ

kdo fotografira koga

kdo fotografira s čim

rbz GBZ

kdo fotografira koga

fotografirati (glagol)

Fida JS 620m (SLD sketch gramm) freq = 8858

GPRS-inf	1633	18.3
<input type="checkbox"/> želeči	151	42.37
<input type="checkbox"/> hoteti	136	42.09
<input type="checkbox"/> dovoliti	49	36.95
<input type="checkbox"/> smeti	64	34.11
<input type="checkbox"/> pustiti	130	31.21
<input type="checkbox"/> marati	11	26.63
<input type="checkbox"/> morati	72	23.93
<input type="checkbox"/> začeti	123	18.1
<input type="checkbox"/> mogoč	15	17.41
<input type="checkbox"/> uspeti	55	17.08
<input type="checkbox"/> moči	25	16.88
<input type="checkbox"/> biti	174	16.77
>>		

v_tožil	1781	9.6
<input type="checkbox"/> sošolka	18	34.16
<input type="checkbox"/> motiv	24	26.92
<input type="checkbox"/> fotografija	33	26.41
<input type="checkbox"/> človek	84	25.65
<input type="checkbox"/> gol	24	25.12
<input type="checkbox"/> predmet	24	22.02
<input type="checkbox"/> pokrajina	18	21.92
<input type="checkbox"/> prizor	15	21.85
<input type="checkbox"/> objekt	27	21.74
<input type="checkbox"/> otrok	47	21.45
<input type="checkbox"/> truplo	12	21.09
<input type="checkbox"/> slika	30	20.76
>>		

kako-kdaj?	1984	5.4
<input type="checkbox"/> rad	194	48.09
<input type="checkbox"/> lahko	322	40.12
<input type="checkbox"/> najraje	34	34.59
<input type="checkbox"/> navdušeno	20	33.18
<input type="checkbox"/> nerad	18	32.29
<input type="checkbox"/> skrivaj	16	30.05
<input type="checkbox"/> skupaj	45	27.77
<input type="checkbox"/> ponosno	13	27.19
<input type="checkbox"/> pridno	15	27.06
<input type="checkbox"/> največkrat	20	27.02
<input type="checkbox"/> večkrat	26	25.46
<input type="checkbox"/> zadnjič	14	24.95
>>		

z-d	530	5.2
<input type="checkbox"/> bliskavica	17	38.91
<input type="checkbox"/> fotoaparāt	24	35.26
<input type="checkbox"/> aparat	13	23.55
<input type="checkbox"/> kamera	12	23.51
<input type="checkbox"/> veselje	12	21.69
<input type="checkbox"/> stran	14	14.3
>>		

v_rodil	216	5.2
<input type="checkbox"/> gol	37	41.74
<input type="checkbox"/> žival	12	21.91
<input type="checkbox"/> leto	16	13.42
>>		

osebek_je	1211	4.0
<input type="checkbox"/> fotograf	45	42.11
<input type="checkbox"/> turist	47	37.62
<input type="checkbox"/> gol	21	25.25
<input type="checkbox"/> kamera	13	21.35
<input type="checkbox"/> zanimivost	10	20.76
<input type="checkbox"/> fotografija	14	18.44
<input type="checkbox"/> novinar	13	17.67
<input type="checkbox"/> obiskovalec	12	17.48
<input type="checkbox"/> kot	21	16.28
<input type="checkbox"/> človek	26	15.29
<input type="checkbox"/> oče	10	14.88
<input type="checkbox"/> moški	10	13.39
>>		

num0 SBZ2
num0 SBZ0

SBZ0 brez SBZ2

gbz SBZ4

sbz0 za SBZ4

krona (samostalnik)

Fida PLUS 620m (SLD sketch grammar) freq = 12072

unary_rels

☐ štev-pred [1490](#) 7.9

>>

pod-d_X 86 12.5

☐ biti [17](#) 11.96

>>

p-koga-česa 270 9.2

☐ diamanten [79](#) 74.1

☐ brusilen [79](#) 71.85

☐ katoliški [10](#) 24.59

>>

brez-d_X 60 8.6

☐ vladar [11](#) 34.85

☐ kralj [14](#) 33.24

>>

poleg-d_X 15 6.2

☐ biti [10](#) 14.13

>>

koga-kaj 1643 6.0

☐ osvojiti [198](#) 38.12

☐ nadeti [48](#) 32.43

☐ nositi [117](#) 31.67

☐ predati [61](#) 28.63

☐ sneti [23](#) 23.38

☐ posaditi [18](#) 19.42

☐ ubraniti [15](#) 19.02

☐ naslediti [13](#) 18.96

☐ prevzeti [46](#) 18.84

☐ odvzeti [20](#) 18.17

☐ braniti [20](#) 16.14

☐ obdržati [16](#) 14.41

>>

za-d_X 643 4.7

☐ boj [66](#) 40.46

☐ favorit [17](#) 29.85

☐ bankovec [14](#) 28.84

☐ potegovati [31](#) 28.17

☐ meč [11](#) 27.57

☐ dvoboj [13](#) 21.46

☐ kandidat [14](#) 19.72

☐ tekmeč [10](#) 19.62

☐ biti [85](#) 15.85

☐ pomeriti [11](#) 14.1

☐ vrsta [11](#) 13.27

☐ bati [11](#) 11.53

>>

s-koga-česa 1647 4.5

☐ prim. [367](#) 95.45

☐ milijarda [270](#) 51.5

☐ balirka [40](#) 51.35

☐ milijon [235](#) 41.54

☐ nakladalka [20](#) 40.62

☐ kosilnica [33](#) 39.46

☐ brana [23](#) 39.15

☐ samonakladalka [12](#) 34.16

☐ tovarna [44](#) 31.51

☐ penzion [14](#) 30.82

☐ zgrabljalnik [10](#) 28.99

☐ plug [15](#) 28.12

>>

povzetek

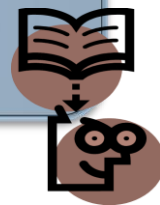


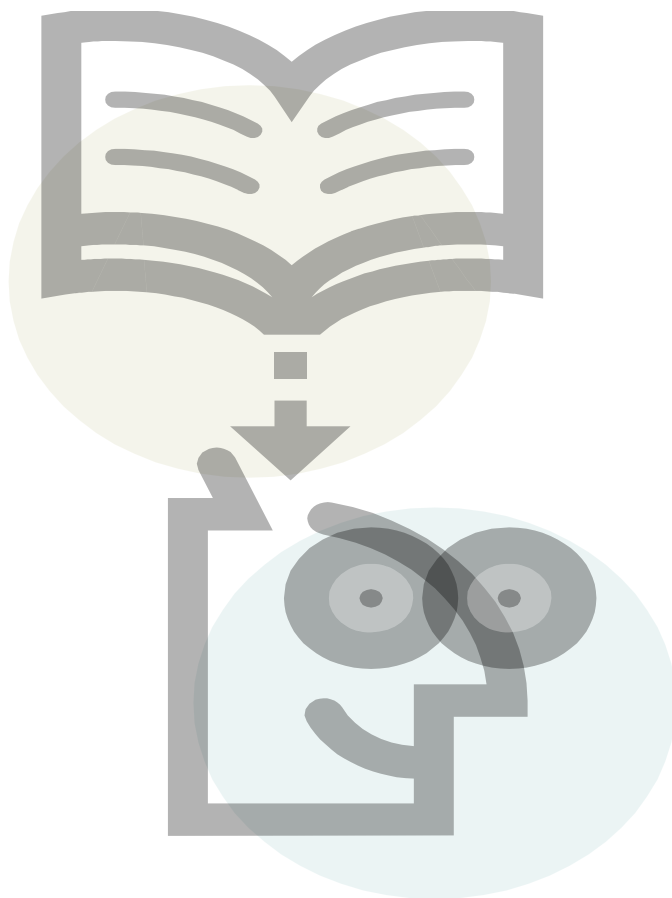
Leksikalna baza je v prvi vrsti namenjena slovarskim opisom besedišča, zato so slovnični podatki v njej podrejeni pomenskim lastnostim LE.

Uporabniška prijaznost in raba v šolske namene narekujeta vključitev slovničnih podatkov v pomenske opise LE na način, ki je uporabniku znan iz vsakodnevne komunikacije in mu je zato bolj razumljiv.

Formalizacija udeležencev v pomenski shemi in registriranje alternacij v obliki stavčnih vzorcev omogoča avtomatično povezavo pomena z njegovim pomensko relevantnim skladenjskim okoljem.

Menimo namreč, da mora biti elektronska leksikalna baza, če hoče zadostiti pedagoškim, računalniškim in drugim aplikacijam prihodnosti, uporabna kot vir za avtomatično pripisovanje pomena besedam v realni jezikovni rabi.





Hvala za pozornost!



REPUBLIKA SLOVENIJA
MINISTRSTVO ZA ŠOLSTVO IN ŠPORT



Naložba v vašo prihodnost
OPERACIJO DELNO FINANCIRA EVROPSKA UNIJA
Evropski socialni sklad