

UDK 811.163.6'373.7

*Polona Gantar*

Inštitut za slovenski jezik Frana Ramovša v Ljubljani

## KORPUSNI PRISTOP V FRAZELOGIJI IN SLOVARSKÉ APLIKACIJE

V članku soočamo poskus identifikacije frazeološke enote na podlagi stopnje pomenske motivacije besednozveznih elementov z izhodišči v ruski frazeološki literaturi in različne vidike besedne povezovalnosti, kot se odkrivajo v korpusnem okolju. Relativizacija razmerij med eno- in večbesednimi leksikalnimi enotami ter pomensko transparentnimi in netransparentnimi besednimi zvezami širi predmet frazeološke problematike na različne tipe jezikovnega vzorčenja in hkrati ponuja slovarske rešitve, ki temeljijo na kontekstualni obravnavi leksikalnega elementa.

This paper compares an attempt to identify the phraseological unit on the basis of the degree of semantic motivation of phrasal elements, originating in the Russian phraseological tradition, with various aspects of word combining, as revealed in the corpus environment. The relativisation of relations between single word and multiword lexical units on the one hand and the semantically transparent and opaque phrases on the other broadens the subject-matter of phraseology to different types of language patterning and also offers dictionary solutions based on the contextual treatment of the lexical element.

**Ključne besede:** frazeološka enota, večbesedna enota, stalna besedna zveza, kolokacija, frazem, idiom, idiomatičnost, idiomatični pomen, skladijski vzorci, leksikalna enota, leksikografski vidik, slovarska struktura

**Key words:** phraseological unit, multiword unit, fixed expression, collocation, phraseme, idiom, idiomaticity, idiomatic meaning, syntactic patterns, lexical unit; corpus-based approach, lexicographical aspects, dictionary framework

### 1 Frazeologija – zamejitev področja

Frazeološka teorija si že od nekdaj prizadeva čim natančneje zamejiti področje frazeoloških raziskav in s tem tudi osnovno frazeološko enoto (FE). S tem namenom so se oblikovala merila, po katerih bi bilo mogoče določiti njene temeljne in glede na druge leksikalne enote razlikovalne lastnosti. Dejstvo, da je frazeološka teorija najbolj zapletena tam, kjer je večina pravil, ki so prepoznana in preverjena v sorodnih jezikovnih vzorcih t. i. pravičnega jezika<sup>1</sup> zabrisanih, se kaže v tem, da so lastnosti kot *večbesednost*, *kolokabilnost*, *ustaljenost*, *variantnost*, *idiomatičnost*, *konotativnost*, *pretvorbenost* itd. obravnavane z različnih zornih kotov, kar vodi v nasprotujoča si pojmovanja tega, kaj je bistveno za obstoj FE.

Zdi se, da tako stanje na področju večbesednih enot ni naključno. Na eni strani je posledica dejstva, da gre za kompleksen jezikovni pojav, kjer se prepoznavne lastnosti uresničujejo v različni meri, na drugi strani pa je vzrok njihove neujetosti v sistemsko predvidene skladijske in pomenske procese v tem, da zaradi svoje idiosinkratičnosti

---

<sup>1</sup> V frazeoloških razpravah, ki temeljijo na ravninskem pojmovanju jezikovne zgradbe, je pravilni jezik tisti, kjer za razliko od sistemsko nepričakovanih realizacij (t. i. anomalij), značilnih za FE, delujejo splošna skladijska in semantična pravila (prim. Čermák 1985: 167).

ne dopuščajo ločenega skladenjskega in pomenskega obravnavanja posameznih sestavin. Tradicionalno pojmovanje FE je tako izpostavljalo zgolj določen tip večbesednih enot, ki so ustrezale specifičnim zahtevam, kot so strukturna in pomenska ustaljenost, konotativnost, neterminološkost itd., medtem ko so vse druge zveze ostajale zunaj ožje frazeološke posledično pa tudi slovarske problematike.

### 1.1 Idiomaticnost in frazeološki pomen

Ruska frazeološka teorija, kot sta jo v 50. in 60. letih zasnovala V. V. Vinogradov in N. N. Amosova, pa tudi večina vzhodnoevropskih frazeoloških šol, ki so zrasle na teh temeljih, je poskušala oblikovati sistem kategorij, po katerih bi bilo mogoče ločevati področje frazeologije od področja, na katerem delujejo splošno veljavna pravila besednega povezovanja. Temeljna značilnost tako pojmovane FE izhaja iz pojmov *idiomaticnost* in *frazeološki pomen*.

Idiomaticnost, ki se nanaša na razmerje med vstopno in izstopno semantiko sestavin FE glede na celoto, je v splošnem mogoče razumeti kot univerzalni jezikovni pojav, pri čemer se razlikovalne lastnosti med morfemi, besedami in besednimi zvezami v posameznih jezikih formalno in vsebinsko razlikujejo. Če pustimo ob strani možnost idiomaticnega kombiniranja že na ravni morfemov in zanemarimo obstoj enobesednih idiomov tipa *duhamoren*, *nepriprav*, *dolgočasiti se* ipd., je znotraj ožje frazeološke problematike idiomaticnost vezana na prepoznavanje stopnje pomenske neodvisnosti celotne FE glede na pomen posameznih sestavin. In to kljub dejstvu, da je vedno znova prihajalo do različnega razumevanja stopnje pomenske motiviranosti konkretne FE, hkrati pa ni bilo mogoče postaviti ostrih mej med posameznimi stopnjami tako pojmovane FE, saj je uvrstitev v določen tip na tej podlagi v veliki meri odvisna od jezikovne in kulturne izkušnje vsakega posameznega govorca (Cowie 1998: 215).

Na podlagi pojmovanja frazeološkega pomena, tj. pomena FE kot celote, ki ni vsota njenih sestavnih elementov, kot tiste lastnosti, ki je za pojav FE bistvena, se je v frazeološki teoriji oblikovala delitev FE v dve temeljni skupini: na FE, ki jih je mogoče pomensko analizirati, kar pomeni, da je njihov pomen v različni meri razpršen med posamezne sestavine, in na take, ki tega razmerja ne izkazujejo in so v celoti pomensko nemotivirane (Erbach 1992: 12; Nunberg et al. 1994: 496–497). Obenem se je nezmožnost dobesedne prevedljivosti frazeološkega pomena postavljala kot eden temeljnih pogojev za prepoznavnost FE, čeprav so najnovejše na besedilih utemeljene raziskave pokazale, da je pojav medjezikovne idiomaticnosti relativen in dinamičen, saj je lahko določena besedna zveza v nekem jeziku idiomaticna, njen tujejezični ustreznik pa ne nujno (Mlacek et. al 1995: 64). Omenjena izhodišča so bila poglavitni vzrok za to, da so bile FE obravnavane kot specializirani segment leksikalnega fonda in s tem tudi vsebina specializiranih slovarjev, medtem ko so bile v splošnih slovarjih predstavljene precej neučinkovito in nesistematično.

**1.2 Slovenistični pristopi v preučevanju FE** že od prve teoretične razprave<sup>2</sup> o frazeologiji (Toporišič 1973/74) sledi poskusom umestitve FE na posamezne ravnine strukturno pojmovane jezikovne zgradbe. Ob tem so se oblikovala merila za določitev FE, ki so upoštevala večbesednost, stalnost in možnost avtomatične reprodukcije. Z vključitvijo besednozveznih terminov med FE<sup>3</sup> so bila sprva na podlagi ruskih teoretičnih pristopov postavljena izhodišča za pojmovanje frazeologije v širšem smislu, medtem ko se je v drugi polovici 80. let z zahtevo po »neslovarskem pomenu« vsaj ene sestavine (Kržišnik-Kolšek 1986: 435) in z izločitvijo terminoloških besednih zvez iz ožje frazeološke problematike oblikovala delitev na stalne besedne zveze (SBZ) na eni in FE na drugi strani. Z vpeljavo pojma kolokabilnost oz. povezovalnost (Kržišnik-Kolšek 1988: 51–54) se je pojmovanje FE omejilo na enopovezovalne (monokolokabilne) zveze tipa *priti/spraviti na kant*, medtem ko se t. i. omejena povezovalnost, npr. *kriv + obtožba, ovadba; pričevanje, izpoved, prisega; nauk, vera* (Kržišnik 1994: 33) v razmerju do FE ni posebej opredeljevala, čeprav so na ta način ostajale zunaj leksikoloških in frazeoloških raziskav zveze, kjer se besede v katerem od svojih pomenov povezujejo z relativno omejenim naborom drugih besed, prim. tudi *star + mama, mati, oče starši; star + celina, kontinent, svet*. Pomembno merilo, ki je preusmerilo slovenistične frazeološke raziskave na skrajno omejen segment večbesednih poimenovalnih enot (t. i. frazeologija v ožjem smislu), je izpostavitve zgolj tistih zvez, ki izkazujejo pomemben delež konotativnih pomenskih sestavin oz. imajo opazno pragmatično vlogo (Kržišnik 1990: 400), kar je poleg že omenjenih terminov iz frazeološke problematike izločalo tudi frazne glagole kot: *držati s kom, pristati na kaj* ipd. ter predložne kolokacije tipa: (*razlikovati, sortirati*) *po barvi*; *v barvi (kože, lesa)*; (*igrati*) *na mestu (branilca)*, in zveze s t. i. slovničnim pomenom kot: *ne glede na*; *za razliko od*; *v primerjavi z/s* ipd. V zvezi z določitvijo zgornje meje oz. največje možne še frazeološke in ne morda že besedilne entitete, se je skušalo rešiti vprašanje stavčne zgradbe SBZ, pri čemer so bile zlasti sporne nekatere glagolsko-samostalniške zveze – t. i. nepravilni glagolski frazemi (Kržišnik 1994: 83), kot npr. *luna trka koga, srce pade v hlače komu* ipd. Pri preverjanju tega tipa FE je zlasti značilna vloga intuicije (npr. prek anketnega prepoznavanja strukture in pomena zveze, njihovih pretvorbenih možnosti ipd.) ob odsotnosti obsežnejših korpusov (do leta 1997) pa tudi omejitev zgolj na določen tip besedil (časopisni članki, umetnostna besedila, gradivo za Slovar slovenskega knjižnega jezika).

<sup>2</sup> O frazeologiji se je v slovenističnem jezikoslovju razmišljalo že vsaj od konca 50. dalje, ko so se oblikovala izhodišča za Pavličev Frazeološki slovar v petih jezikih (1960) in ko so se oblikovala izhodišča za predstavitev frazeologije v Slovarju slovenskega knjižnega jezika, katerega prva knjiga je izšla leta 1970.

<sup>3</sup> Pomemben prispevek v tem smislu je opredelitev skupine terminoloških SBZ z vrstnim pridevnikom, tip. *mehki, trdi les; črni bor* ipd., z vidika formalne prepoznavnosti stopnje pomenske zlitosti, kot se odraža v besednozvezni oz. morfemski zgradbi terminov (Vidovič Muha 1988).



še vedno odraža prav v oblikovanju zaključkov na podlagi majhnega obsega podatkov (Čermák 2001: 5). Področje frazeologije je v tem smislu posebno tudi zato, ker se, kot že omenjeno, od drugod znane kategorije znotraj stalnih zvez zabrisujejo in jih je nemogoče razbrati iz manjšega števila vzorcev.

Pri uporabi korpusa za leksikografske namene se tako ponuja možnost identificiranja besednih sopojavitev, kadar so te za jezik tipične. Preučevanje dobljenih vzorcev pa daje glede na dejanski besedilni kontekst tudi možnost vpogleda v tipične pomenke in sporočanje vloge, ki se v zvezi z njimi uveljavljajo. Dejstvo je, da izraba korpusa v gradivnem in metodološkem smislu leksikografu omogoča več fleksibilnosti pri oblikovanju zasnove slovarja, zlasti tudi v razmerju do potencialnega uporabnika. Verjetno zato tudi ni naključje, da sta se gradnja korpusov tako glede zajemanja čim večje jezikovne variantnosti kot tudi razvoj orodij za delo s korpusi sooblikovala in izpopolnjevala tudi skozi slovarske projekte. Zlasti tiste, ki so želeli zagotoviti čim večjo realnost jezikovnih podatkov, v prvi vrsti z izbiro geselskih besed, zvez in njihovih oblik, ki so opazno zastopane v dejanski rabi. Prav na podlagi korpusov pridobljenih podatkov so bili leksikografi prisiljeni premisliti upravičenost vključevanja oblik, ki so zgolj odraz tvorbenih možnosti jezika in nimajo potrditve v dejanskih besedilih. Slovarji, ki nastajajo na podlagi korpusa, lahko z večjo zanesljivostjo zajamejo pomenke vrednosti leksikalnih elementov in ugotavljajo njihovo dejansko pogostnost. Povsem nove razsežnosti so s korpusi dobili tudi zgledi rabe. Če namreč sprejmemo zahteve po koherenci in sporočanje učinkovitosti besedila, izhajamo iz dejstva, da se besedilo oblikuje na več zelo prefinjenih načinov, ki se jim z umetnim zgledom sicer lahko bolj ali manj približamo, vendar pa občutljivosti in pomembnosti sobesedila ne moremo nadomestiti. V tem smislu je tudi za nabor bistvenih lastnosti večbesednih enot na podlagi korpusa mogoče določiti izmerljiva in s tem razmeroma objektivna merila.

### 2.1 Lastnosti FE – izhodišča in določitev meril

Lastnosti besedne povezovalnosti – predvsem z namenom identificirati različne tipe večbesednih enot kot potencialnih slovarskih entitet – je v korpusnem okolju mogoče ugotavljati glede na pogostnostno, funkcijsko in pomensko izhodišče. Izhodišča so dodatno določena s temeljnimi postopki korpusne analize (Teubert 1999): identifikacijo jezikovnih podatkov, kjer izhajamo iz besede oz. njene oblike; korelacijo jezikovnih podatkov s pomočjo statističnih metod, kjer opazujemo vidike besedne povezovalnosti in jezikovnega vzorčenja, ter interpretacijo dobljenih rezultatov. V ta namen smo uporabili Korpus slovenskega jezika FIDA in konkordančnik, ki je uporabnikom na voljo ob dostopu.

**Pogostnostni vidik**, ki ga postavljamo v izhodišče analize, se nanaša na prepoznavanje očitnih besednih sopojavitev in ugotavljanje tipičnih sopojavnic – t. i. kolo-katorjev preučevane besede v izhodišču konkordančnega niza. Tu so nam, kot kaže zgled 1, v korpusu FIDA na voljo vsaj tri možnosti, pri čemer daje za ugotavljanje večbesednih enot kot potencialnih leksikalnih elementov najboljše rezultate vrednost  $MI^3$ , zlasti v kombinaciji s podatkom o absolutni pogostnosti korpusnega elementa, ki ga preučujemo v izhodišču konkordančnega niza (Gorjanc in Krek 2001).

**Zgled 2:** pogostnostno izhodišče; izhodišče konkordančnega niza: šala.

(a) 10 najpogostejših kolokatorjev levo od izhodišča		(b) 10 vzajemno povezanih kolokatorjev levo od izhodišča – vrednost MI		(c) 10 (so)vzajemno povezanih kolokatorjev levo od izhodišča – vrednost MI <sup>3</sup>	
1 za	544	1 = privihnil	14.987226	1 zbijati	24.799923
2 v	348	2 = jelušičeve	14.987226	2 = prvoaprilska	24.164425
3 biti	188	3 = yorkshirskim	14.987226	3 neslan	23.591403
4 in	59	4 = gattinonijevi	14.987226	4 za	21.953701
5 dober	56	5 = vsplahutale	14.987226	5 = prvoaprilsko	20.973203
6 kota	55	6 = yorkshirskimi	14.987226	6 zbijanje	19.883292
7 kot	55	7 = carmichaelovega	14.987226	7 neslano	19.783085
8 zbijati	55	8 = vzorčan	14.987226	8 = prvoaprilske	19.708851
9 se	49	9 = vidrim	14.987226	9 = severin	19.211598
10 pripovedovati	46	10 = pinčevi	14.987226	10 v	18.915057

**Funkcijski vidik** temelji na prepoznavanju tipičnih skladenjskih vzorcev, v katerih se pojavlja beseda v jedru konkordančnega niza, in na vzpostavljanju pomenskih povezav med njimi. Za take vzorce je značilno, da so rezultat slovničnih in pomenskih lastnosti jezika in zato nekakšni slovnično-leksikalni konglomerati. Kot taki postajajo v korpusnem okolju izhodišče za različne slovnične, leksikalne in skladenjske analize in preusmerjajo pozornost z ravni jezikovnosistemskih možnosti v preučevanje vzorcev besedilnih realizacij, tipološke zakonitosti, ki se na tej podlagi ustvarjajo, upoštevajo tudi vse »kršitve«, ki predstavljajo osnovo za aktualiziran jezikovni opis. Tako denimo glagol *veljati* – poleg tipičnih kolokacij, v katerih so posamezni kolokatorji pomensko razločevalni, npr. (*začeti, prenehati*) *veljati*; *velja* (*poudariti, omeniti*); *korajža velja* – s predlogom *za* tvori skladenjske vzorce, ki so na posameznih stopnjah leksikalizacijskega procesa (prim. a1 in a2, ki sta v nasprotju z b frazna glagola) tudi slovarsko zanimivi.

**Zgled 3:** funkcijsko izhodišče; izhodišče konkordančnega niza: veljati za.

skladenjski vzorec	besedilni vzorec	pomen	besedilni zgled
(a1) veljati za <i>kakšnega</i>	veljati za <i>pomembnega/ nedolžnega/spoštovanega/ uglednega ...</i>	'biti tak'	V nasprotju z Nizozemci, ki <b>veljajo za</b> skromnejše turiste, Belgijci na počitnicah ne stiskajo pri izdatkih.
(a2) veljati za <i>koga/kaj utemeljitelja/ začetnika/</i>	veljati za <i>favorita/ predhodnika ...</i>	'biti kdo/kaj'	Te freske, ki žal niso v celoti ohranjene, <b>veljajo za</b> glavno delo slikarske šole v Ferari.
(b) veljati za <i>koga/kaj</i>	veljati za <i>vse/večino/oba...</i>	'nanašati se na koga kaj;	Prišli boste do spoznanja, da je življenje borba, in če to <b>velja za</b> vas, zakaj ne bi <b>veljalo za</b> druge.

**Pomenski vidik** temelji na prepoznavanju sorodnih pomenskih realizacij na podlagi obsežnih konkordančnih nizov in omogoča prenos tipičnih pomenskih lastnosti v slovarsko razlago. Pomenski vidik je pri identifikaciji stalnih zvez ključnega pomena in ga ni mogoče ločiti od pogostnostnega in skladenjskega, pri čemer je prepoznavanje leksikalne vloge pogostnostno izstopajoče besedne zveze izrazito kompleksen pojav in v veliki meri povezan s prepoznavnimi elementi besedilnega okolja, tipom besedila in drugimi jezikovnimi in zunajjezikovnimi pojavi. Kot rojeni govorci se pri interpretaciji pomenske vsebine tudi v primeru večbesednih enot nikoli resnično ne odrečemo intuiciji, kljub temu pa je na podlagi številnih besedilnih realizacij, kot nam jih odkrivajo konkordančni nizi in različne možnosti urejanja besedilnega gradiva v korpusnem okolju, mogoče z veliko mero zanesljivosti prepoznati sorodne pomenske realizacije in jih abstrahirati v smislu tipiziranih slovarskih razlag.

Predvidevamo torej, da je vsaka pogosta (so)vzajemno povezana večbesedna enota potencialna leksikalna enota ali tipični skladenjski vzorec preučevanega jezika, kar širi ožjo frazeološko problematiko na celoto SBZ, pri čemer je treba njihovo dejansko leksikalno vlogo šele ugotoviti. Tudi na slovenskem gradivu se je namreč pokazalo, da tipične sopojavitve nimajo nujno tudi opazne leksikalne vrednosti. Zlasti pogoste so sopojavitve s slovničnopomenskimi elementi (predlogi in vezniki), ki lahko postanejo slovarsko zanimive kot del širših skladenjskih struktur, npr. *kar tako za šalo; bolj za šalo kot zares; gre za šalo; kot za šalo* (*premagati, opraviti z/s kom/čim, pomesti s kom*); *malo/malce za šalo* (*in malo/malce zares; napol za šalo napol zares; za šalo* (*povprašati, reči ...*); *vzeti, jemati za šalo; imeti smisel za šalo* itd.

## 2.2 Tipologija SBZ na podlagi korpusnih podatkov

V raziskavi predvidevamo tri relativno samostojne tipe SBZ, ki jih je mogoče določiti na podlagi prepoznavnih sintagmatskih in paradigmatskih razmerij, kot jih odkrivajo korpusni podatki, pridobljeni glede na zgoraj omenjena izhodišča. Sintagmatske lastnosti se odražajo v kolokabilni moči besed na horizontalni osi, paradigmatske lastnosti pa se odkrivajo v možnosti kopičenja besed na vertikalni osi znotraj posameznega pomenskega polja. Predvideni tipi SBZ hkrati tudi napovedujejo potencialne slovarske iztočnice oziroma hierarhično razmerje znotraj slovarskega članka, kot bomo prikazali v nadaljevanju.

**Zgled 4:** Razvrstitev kolokatorjev levo od izhodišča glede na sintagmatska in paradigmska razmerja; izhodišče konkordančnega niza: jajce

p a r a d i g m a t i k a	s i n t a g m a t i k a						
		stepeno	kokošje	valiti	zlato	kolumbovo	kukavičje
	kuhano	ptičje	leči				
	umešano	kurje	odlagati				
	pečeno	nojevo					
	ocvrto						
	surovo						

Relativnost razmerij med eno in večbesednimi leksikalnimi enotami ter med pomensko transparentnimi in pomensko netransparentnimi SBZ se v korpusu zabrisuje, kar seveda ne pomeni, da je na podlagi korpusnih podatkov nemogoče določiti temeljni predmet leksikalne problematike, pač pa predvsem preseči skrajnostno delitev na besede in besedne zveze na eni strani ter na stalne in proste zveze na drugi. S tem ko je korpus omogočil prepoznavno sopojavljanje posameznih besed, je hkrati tudi prevrednotil pojme kot so leksikalna in skladijska enota, kar je v prenosu na leksikografsko prakso z upoštevanjem dejstva, da je težišče leksikografije na pripisovanju pomena jezikovnemu znaku, ki se manifestira vedno in samo v besedilu, pomenilo enakovredno obravnavo besedne in besednozvezne problematike ter vključitev SBZ in očitnih skladijskih vzorcev tudi v splošne in ne zgolj specializirane slovarje. S tega vidika pojmovanje *kolokacije* kot formalno-verjetnostnega in hkrati semantičnega pojava, kot se odraža v medsebojni povezanosti leksikalnih elementov, izpostavlja v grobi delitvi dve vrsti tipičnih besednih sopojavitvev. Take, ki indicirajo posamezne pomene večpomenskega leksema, npr. (*rdeča, modra, svetla, temna* ipd.) *barva*; (*tiskarska, oljna*) *barva*; – 'snov za barvanje'; (*politična, klubska*) *barva* – 'odraz pripadnosti čemu', in take, ki se pojavljajo zgolj pri izbranem pomenu besede in tako tvorijo bolj ali manj omejene kolokatne paradigme<sup>4</sup> (Čermák 1985: 171, Kržišnik 1988: 51), npr. *barva kože* – 'rasna pripadnost'; *osnovne barve* 'temeljne barve barvnega spektra' ipd. Prvi tip kolokacij, ki je v spodnji preglednici razvrščen pod I., vzpostavlja nepo-

<sup>4</sup> Pojem kolokatne paradigme izhaja iz različnih vidikov besednega povezovanja povsod tam, kjer lahko upoštevamo (jezikovni) pomen. Pri tem neomejenost na eni in preštevnost prvin na drugi strani, ki se znotraj kolokatne paradigme oblikovno in pomensko povezujejo v smiselno celoto, ponuja dve temeljni množici besednih zvez, med katerima se določa osnovna enota frazeologije: veliko kolokatno paradigmo, za katero velja, da je odprta, število njenih prvin pa je neomejeno in omejeno kolokatno paradigmo, ki je praviloma zaprta, število njenih prvin pa je prešteveno.

sredno povezavo z enobesednimi elementi in predstavlja tipično besedilno umestitev preučevane besede. Drugi tip kolokacij (pod II.) obsega tiste stalne zveze in skladenjske vzorce, ki so na prehodu med pomensko transparentnimi, tip *osnovne barve*, (*cvetje* ipd.) *vseh barv*, (*razlikovati*, *razvrstiti*, *ločiti*) *po barvi*; *obrniti (kaj) na/v šalo* ipd. in pomensko netransparentnimi, kjer je pomenska povezanost kolocirajočih elementov vzajemna, npr. *spreminjati barve* 'izražati jezo, prizadetost' itd.; 'přiti s pravo barvo na dan' 'izraziti svoje dejanske, skrite namere ali značaj'. V frazeološki literaturi predstavlja omenjeni tip skupaj s pomensko netransparentnimi zvezami osrednji del frazeološkega fonda, ki ga najpogosteje opredeljuje izraz *frazem*. Na slovarski ravni gre za zveze, ki zahtevajo razlago, hkrati pa se bolj ali manj navezujejo na katerega od pomenov besede, ki tako zvezo tvori, kar postavlja možna izhodišča za znotrajslovarsko hierarhijo. Skrajno točko v pojmovanju pomenske neizraženosti oz. predvidljivosti pomena iz sestavin besedne zveze predstavljajo *idiomi*, zato se znotraj slovarja navadno obravnavajo kot relativno pomensko samostojne enote.

Ob tem je treba povedati, da nikakor ni mogoče potegniti ostre ločnice med pomensko transparentnimi in netransparentnimi SBZ, kot so npr. *ubiti dve muhi na en mah*; *obrniti komu hrbet*, *imeti zvezane roke* itd. in t. i. »pravimi« *idiomi*, kamor štejejo zveze brez očitne pomenske povezave s katerokoli od sestavin, ki jo tvorijo, npr. *iti se gnilo jajce*, *železna zavesa*, *na vrat na nos* ipd. Ker pa slovar glede na svojo priročniško in uporabniku prijazno naravo zahteva konsistentno strukturo, je smiselno že v izhodišču slovarske zasnove premisliti razmerja med eno- in večbesednimi leksikalnimi enotami kot potencialnimi iztočnicami oz. njenimi elementi. Eno izmed možnih načelnih rešitev predstavlja spodnja preglednica.

## STOPNJA POMENSKE TRDNOSTI

LEKSIKALNA ENOTA	SBZ			
	kolokacija		frazem	idiom
	I.	II.		
enobesedna	<i>barva</i> ( <i>rdeča, modra; svetla, temna ...</i> ) <i>barva</i> ( <i>tiskarska, oljna</i> ) <i>barva</i> ( <i>politična; klubska</i> ) <i>šala</i> ( <i>posrečena, robata, opolzka ...</i> )			
večbesedna		<i>osnovne barve</i> <i>barva kože</i> <i>(cvetje) vseh barv</i> <i>v barvi (kože, lesa)</i> <i>prvoaprilska šala</i> <i>neslana/neokusna šala</i> <i>v šali</i> ( <i>dejati, praviti</i> ) <i>(kot) za šalo</i> ( <i>premagovati ...</i> ) <i>ni šale</i> ( <i>s kom/čim</i> ) <i>(razumeti, vzeti, jemati) kot šalo</i> <i>obrniti (kaj) na/v šalo</i> <i>za šalo</i> ( <i>vprašati, poskusiti ...</i> ) <i>(ne) biti za šalo</i> <i>iti za šalo</i>	<i>spreminjati barve</i> <i>přiti s (pravo) barvo na dan</i> <i>zbijati/stresati šale</i> <i>šalo na stran</i> <i>malo za šalo malo zares</i>	<i>hudič/vrag je odnesel šalo</i>

Omenjeni tipi SBZ nam predstavljajo potencialne slovarske enote, pri tem pa se, vsaj z vidika uporabne slovarske informacije, ne zdi pomembna stopnja pomenske trdnosti potencialne slovarske enote, zlasti ne v smislu izključitve ne dovolj leksikaliziranih besednih zvez iz slovarja, pač pa se za samo ureditev slovarja zdi pomembno na različnih nivojih geselskega članka predstaviti posamezne možnosti besednih kombinacij, ki so v jeziku očitne in hkrati opravljajo opazno leksikalno vlogo.

### **2.3 Razmerje med eno- in večbesednimi leksikalnimi enotami – slovarski primer**

Dejstvo, da so večbesedne leksikalne enote glede na svojo leksikalno vlogo enakovredne enobesednim, v slovarju sicer ne upravičuje znotrajiztočniškega podrejanja, zlasti ne v smislu odsotnosti predvidenih slovarskih informacij (pomen, izgovor, besednovrstna pripadnost, zgledi rabe itd.). Vendar pa je umestitev večbesednih enot v slovarski sistem v marsičem bolj zapletena, kot se zdi na prvi pogled, vsaj iz dveh razlogov. Prvi je že omenjeno dejstvo, da so večbesedne enote sestavljene iz elementov »pravnega« jezika, kar pomeni, da največkrat obstajajo tudi zunaj konkretne idiomatične kombinacije, in drugič, da hkrati v različni meri ohranjajo tudi v idiomatični kombinaciji svoje zunajidiomatične slovnične in pomenske lastnosti. Prav ugotavljanje pomenske asociativnosti med sestavino SBZ in besedo, ki obstaja tudi neodvisno od SBZ, izpostavlja možnost slovarskega razvrščanja SBZ glede na predhodno dane pomenske in slovnične podatke sestavin v vlogi nadrejene (enobesedne) iztočnice. Ker je torej besedilno okolje besede, ki jo preučujemo v izhodišču konkordančnega niza, v središču naše pozornosti, lahko med posameznimi tipičnimi sopojavitvami posebej izločimo tiste povezovalne možnosti, ki jih ustvarja denimo metaforični potencial večpomenske besede, ki postane sestavina SBZ. Pri tem lahko stopnja pomenske trdnosti predstavlja oprijemljivo izhodišče za uvrstitev SBZ pod posamezni pomen besede v vlogi nadrejene enobesedne iztočnice.<sup>5</sup> Poglejmo si nekatere možnosti.

---

<sup>5</sup> Način razvrščanja SBZ znotraj geselskega članka podpira hierarhizirana struktura leksikalne podatkovne zbirke, kot je bila izdelana na podlagi Korpusa slovenskega jezika FIDA v okviru ciljnega raziskovalnega projekta »Zasnova na korpusu temeljčih slovarskih in slovničnih opisov slovenskega jezika« (Gorjanc, Krek, Gantar 2005: 14–16).

<b>kisel</b> prid.		<b>lisica</b> sam.	
<b>1.</b> 'okus'		<b>1.</b> 'žival'	
<b>kolokacije</b>	kislo (grozdje, sadež, solata) kislega (okusa)	<b>kolokacije</b>	(stekla, povožena) lisica (pokončati, upleniti) lisico (cepljenje, vakcinacija) lisic lov na lisico
<b>SBZ</b>	kisla smetana kislo zelje kisla repa kislo mleko kisla juha kisla kumarica	<b>SBZ</b>	šakalska lisica leteča lisica morska lisica polarna/arktična/bela/srebrna lisica puščavska lisica
<b>2.</b> 'nevesel, neprijeten'		<b>2.</b> 'lisičjea krzno'	
<b>kolokacije</b>	kisel (nasmeh, obraz, vreme)	<b>kolokacije</b>	krzlena lisica
<b>SBZ</b>	ugrizniti v kislo jabolko	<b>SBZ</b>	(mariborska, pohorska) zlata lisica
<b>3.</b> 'z visoko stopnjo ph vrednosti'		<b>3.</b> 'prebrisan človek'	
<b>kolokacije</b>	kisla (tla, prst, zemlja; okolje)	<b>kolokacije</b>	(prebrisana) lisica (zvit) kot lisica
<b>SBZ</b>	kisli dež/padavina kisla voda		
<b>idiom</b>	čas, sezona kislih kumaric	<b>lisice</b> sam.	
		<b>1.</b> 'priprava za vklepanje aretirancev'	
		<b>kolokacije</b>	(nataknuti, nadeti, sneti) lisice, vklepani v lisice
		<b>SBZ</b>	policijske lisice
		<b>2.</b> 'priprava za zaklepanje nepravilno parkiranih avtomobilov'	
		<b>kolokacije</b>	(odklepanje, priklenitev) lisic odpraviti lisice uvedba lisic

### 3 Sklep

Poskusi zamejitve področja frazeologije na podlagi ugotavljanja stopenj idiomatičnosti niso zagotovili enotnega pojmovanja FE, zlasti ne v razmerju do prostih in pomensko transparentnih SBZ. Pojmovanje idiomatičnega tj. od sestavin neodvisnega celostnega pomena zveze, oži področje frazeologije na pomensko netransparentne SBZ, ki so bile navadno predmet specializiranih frazeoloških slovarjev, medtem ko so bile vse druge oblike jezikovnega vzorčenja iz njih izključene, v splošnih slovarjih pa predstavljene neučinkovito in nesistematično. Z izpostavitvijo vidikov besedne povezovalnosti ne glede na stopnjo medsebojne gramatične in pomenske zlitosti ter z možnostjo avtomatičnega urejanja konkordančnih nizov in merjenja besedne povezovalnosti v obliki statističnih izračunov v korpusnem okolju je mogoče oblikovati objektivna izhodišča za prepoznavanje tipičnih besednih kombinacij, ki so v različnih stopnjah leksikalizacijskega procesa tudi slovarsko zanimive. Čeprav se razmerja med eno- in večbesednimi leksikalnimi enotami ter med pomensko transparentnimi in pomensko netransparentnimi SBZ v korpusu zabrisujejo, je s kontekstualno obravnavo leksikalnega elementa v leksikografski praksi mogoče povsem suvereno prikazati tako strukturno kot pomensko širitev besede na raven besedne zveze ali obsežnejšega skladiškega vzorca.

#### LITERATURA

- COWIE, Antony, 1998: Phraseological Dictionaries: Some East-West Comparisons. *Phraseology*. A. P. Cowie (ur.). *Theory, Analysis, and Applications*. Oxford University Press. 210–228.
- COWIE, Anthony, 1999: Phraseology and corpora: some implications for dictionary making. *International Journal of Lexicography* 12 (4). Oxford: Oxford University Press. 307–323.

- ČERMÁK, František, 1985: Frazeologie a idiomatika. František Čermák, Josef Filipec: *Česká lexikologie*. Praha: Academia. 166–248.
- – František, 2001: Substance of idioms: perennial problems, lack of data or theory? *International Journal of Lexicography* 14 (1). 1–20
- ERBACH, Gregor, 1992: Head-Driven Lexical Representation of Idioms in HPSG. M. Everaert, idr. (ur.). *International Conference on Idioms Tilburg, NL: Proceedings of Idioms I–II*.
- GANTAR, Polona, 2004: Frazem in njegovo besedilno okolje. Doktorska disertacija. Univerza v Ljubljani, Filozofska fakulteta.
- GORJANC, Vojko, KREK, Simon, 2001: A Corpus-Based Dictionary Database as the Source for Compiling Slovene-X Dictionary. *COMPLEX 2001. 6<sup>th</sup> Conference on Computational Lexicography and Corpus Research: »Computational Lexicography and New EU Languages«*. Birmingham: The University of Birmingham. 41–47.
- GORJANC, Vojko, KREK, Simon, GANTAR, Polona, 2005: Slovenska leksikalna podatkovna zbirka. *Jezik in slovstvo* 50/2. 3–19.
- HUNSTON, Susan in FRANCIS, Gill, 2000: Pattern Grammar. A corpus-driven approach to the lexical grammar of English. John Benjamins Publishing: Amsterdam, Philadelphia.
- Korpus slovenskega jezika FIDA: <http://www.fida.net>
- KREK, Simon, 1999: Računalniški korpusi v slovaropisju. *Razgledi* 13. 14–16.
- KRŽIŠNIK-KOLŠEK, Erika, 1986: Poskus razvrstitve stalnih besednih zvez v Trubarjevi Cerkovni ordningi. Obdobja 6, 16. stoletje v slovenskem jeziku, književnosti in kulturi. Ljubljana: Filozofska fakulteta. 435–445.
- – 1988: *Frazeologija v moderni: Magistrsko delo*. Mentorica B. Pogorelec. Ljubljana: Filozofska fakulteta.
- KRŽIŠNIK, Erika, 1990: Tipologija frazeoloških prenovitev v Cankarjevih prozih besedilih. *Slavistična revija* 4/38. 399–420.
- – 1994: Slovenski glagolski frazemi (ob primeru frazemov govorenja). Doktorska disertacija. Ljubljana: Filozofska fakulteta.
- MLACEK, Jozef, idr., 1995: *Frazeologická terminológia*. Bratislava: STIMUL – Centrum informatiky a vzdelávania FF UK.
- MOON, Rosamund, 1998: *Fixed Expressions and Idioms in English: A Corpus-Based Approach*. Oxford: Clarendon Press (Oxford Studies in Lexicography and Lexicology).
- NUNBERG, Geoffrey, idr., 1994: Idioms. *Language* 3. 491–538.
- TEUBERT, Wolfgang, 1999: Korpuslinguistik und Lexikographie. *Deutsche Sprache* 4/99. John Benjamins. V: Študije o korpusnem jezikoslovju. V. Gorjanc in S. Krek (ur.). Ljubljana: Krtina. 2005. 103–136.
- TOPORIŠIČ, Jože, 1973/74: K izrazju in tipologiji slovenske frazeologije. *Jezik in slovstvo* 8. 273–279.
- VIDOVIČ MUHA, Ada, 1988: Nekatere jezikovnosistemske lastnosti strokovnih besednih zvez. XXIV. seminar slovenskega jezika, literature in kulture. Ljubljana: Filozofska fakulteta. 83–91.

## SUMMARY

Attempts to delimit the field of phraseology based on determining the degree of idiomaticity have not resulted in a single concept of the PU, especially when compared to free and semantically transparent fixed expressions. The concept of an idiomatic meaning of a phrase, i.e. a meaning which does not depend on its constituent parts, narrows the field of phraseology to semantically opaque fixed expressions which used to be the subject-matter of specialised phraseological dictionaries, while all other forms of language patterning were excluded from them

and presented ineffectively and unsystematically in general dictionaries. The presentation of aspects of word combining regardless of the degree of their grammatical and semantic fusion, and the possibility to automatically sort concordance strings and measure word collocability in the form of statistical calculations in a corpus environment make it possible to form objective bases for recognising typical word combinations which may be of dictionary interest in various stages of the process of lexicalisation. Even though the relations between single word and multiword lexical units and the relations between semantically transparent and semantically opaque fixed expressions are blurred in a corpus, it is possible to quite objectively show both the structural and the semantic extension of the word to the level of the phrase or a longer syntactic pattern by considering the lexical element in its context.