



UDK 811.163.6'373.7

Matej Meterc

ENOTE SLOVENSKEGA PAREMIOLOŠKEGA MINIMUMA V GOVORNEM KORPUSU GOS

Prispevek predstavlja prvi poskus raziskave slovenske paremiologije v korpusu govorne slovenščine GOS, ki obsega 120 ur posnetega govora. S korpusno raziskavo, v katero smo vključili 300 najbolj poznanih paremioloških enot iz slovenskega paremiološkega minimuma, smo našli 37 enot. Opisujemo razpon variant, prenovitev in transformacij pregovorov ter prikazujemo primere njihove neprototipne rabe. Po analizi uvajalnih sredstev pregovorov pokažemo, kako ta sredstva lahko uporabimo za odkrivanje novega frazeološkega gradiva v korpusih. Analiziramo tudi zastopanost enot v različnih tipih diskurzov in razmišljamo o mestu paremiologije med preostalimi stavčnimi frazemi.

Ključne besede: korpus govorne slovenščine GOS, frazeologija, paremiološki minimum, paremiološka enota, pregovor

The paper represents the first attempt to examine Slovene paremiology as it is shown in the corpus of spoken Slovene (GOS) which includes 120 hours of recorded speech. Thirty-seven units were identified with corpus analysis that included 300 most well-known paremiological units from the Slovene paremiological minimum. We describe the range of proverb variants, their actualisations as well as transformations, and provide examples of non-prototypical usage of proverbs. After analyzing the text introducers of the proverbs, we show how these introducers can be used to detect new phraseological material in corpora. We also analyze the distribution of units in different types of discourse and discuss the place of paremiology among other sentential phrasemes.

Key words: corpus of spoken Slovene GOS, phraseology, paremiological minimum, paremiological unit, proverb

0 Terminološko in empirično izhodišče ter glavna vprašanja

Paremiološke enote (v nadaljevanju PE) razumemo kot tisti del stavčne frazeologije, ki se prekriva s kratkimi folklornimi žanri (Permjakov 1988: 82), med katerimi močno prevladujejo pregovori in reki. Razmejevanju pregovora in reka se je posvetil že Mlacek (1983: 131), ki ugotavlja, da je meja prehodna in za frazeološke raziskave ne predstavlja bistvene teme. Zaenkrat je bilo v slovenskem prostoru med stavčnimi frazemi največ pozornosti posvečeno prav tema osrednjima paremiološkima žanroma, širšemu spektru stavčnih frazemov pa skoraj nič. Temu spektru ustreza Toporišičev termin *rečenica*,¹ ki označuje enote, ki so »povedne že v našem pomnilniku«

¹ Raba tega izraza je začela pešati skupaj s terminom *reklo* za besednozvezne frazeme in se skorajda ne uporablja več. Za njegovo rabo je bilo še v devetdesetih letih značilno, da so ga raziskovalci omejevali le na obravnavo pregovorov, ali pa ga z njimi celo enačili. Za njim je ostalo prosto strukturno mesto, ki ga samo pregovori in reki ne zapolnijo.

(1985: 33). Upoštevati moramo indice o dokaj nizki številčnosti pregovorov v (še dokaj neraziskani) širši množici stavčnih frazemov (Mlacek 1983, Čermák 2009).² Obenem v svetovni in slovenski frazeologiji še vedno prevladuje predstava o središčnem položaju pregovorov v tem delu frazeologije. Čeprav nas v prispevku zanima predvsem govorna raba slovenske paremiologije, se bomo k temi širšega spektra stavčne frazeologije še vrnili.

Vprašanje frazeologije v govorni komunikaciji je pomembno, ker obsega v vsakdanji komunikaciji znatno večji delež kot pisna, jezikoslovje pa je govorne vire dolgo časa zanemarjalo (Čermák 2005: 19–21). Čermák je prepričan, da bodo to vlogo v prihodnosti prevzeli govorni korpusi, ki pa so zaenkrat še dokaj majhni. Najprej si moramo zastaviti vprašanje smotrnosti raziskovanja PE v govornem korpusu v velikosti, ki nam je na voljo. V korpusu GOS, ki obsega 120 ur posnetkov, lahko pričakujemo relativno malo zgledov PE. Zato je smiselno, da se v prvi tovrstni raziskavi posvetimo dovolj obsežni množici PE, ki v sodobni slovenščini predstavljajo aktivno paremiološko jedro. Če bi se lotili iskanja poljubnih PE z opiranjem na jezikovno intuicijo, bi se morda takemu jedru približali, vendar bi šlo za subjektivno oceno enega govorca. Ocenjujemo, da tako jedro predstavlja slovenski paremiološki minimum, ki smo ga delno že predstavili (Đurčo, Meterc 2013). Minimum 300 najbolj poznanih slovenskih PE, ki izvira iz Permjakovega koncepta (1988: 143–44), smo ustvarili s sociolingvistično raziskavo s spletnim vprašalnikom po Đurčevi (2002) metodologiji na naslovu vprašalnik.tisina.net (Meterc 2015). V izhodiščnem korpusu za raziskavo je 918 PE, izpisanih iz SSKJ-ja in *Frazeološkega slovarja v petih jezikih* (Pavlica 1960), med katerimi prevladujejo pregovori in reki, nekaj pa je tudi drugih paremioloških žanrov (kritič in wellerizmov). V prispevku uporabljamo minimum 1.0, ki je nastal oktobra 2012 po tem, ko je vprašalnik v celoti izpolnilo 316 anketirancev iz vse Slovenije, rojenih med letom 1929 in 2000.³ V okviru doktorskega študija (Meterc 2014a) smo minimum po Đurčevem konceptu (2006: 3) nadgradili v paremiološki optimum, v katerem vrstni red PE določata tako poznanost PE kakor njihova pogostnost v korpusu pisnih besedil FidaPLUS. O enotah paremiološkega minimuma si lahko zastavimo naslednja vprašanja:

1. Koliko jih lahko najdemo v govornem korpusu velikosti GOS, kako pogosto se pojavljajo posamezne enote in kakšen je delež ti. neprototipnih pojavitev?
2. Kaj lahko po količini najdenih PE iz minimuma sklepamo o njihovi rabi na splošno?
3. Ali je vzorec zgledov dovolj velik za prikaz razpona oblik, v katerih obstajajo in v katere prehajajo?
4. Kakšna je zastopanost uvajalnih sredstev?
5. Kakšna je distribucija PE po tipih diskurza in govornih kanalih in v kakšnem odnosu je ta delež z zastopanostjo teh tipov diskurza in kanalov v samem korpusu?

² V Čermákovem slovarju češke stavčne frazeologije (2009), ki velja za vrhunski in morda tudi v svetovnem merilu najbolj izpopolnjen pristop k temu delu frazeologije, je pregovorov komaj desetina.

³ Raziskava je še vedno v teku, saj želimo podatkovno zbirko povečevati tako v smislu demografskega vzorca, kakor tudi vpisov variant in odgovorov na dodatna vprašanja. Septembra 2014 je število v celoti izpolnjenih vprašalnikov preseglo 400.



6. Kako védenje o govorni rabi PE iz minimuma umestiti v širšo sliko stavčne frazeologije?

Podatkom o regiji snemanja in regiji, iz katere prihaja govorec, se ne bomo posebej posvetili. V prihodnosti bi jih bilo treba proučiti v navezavi na variantnost PE.

1 Iskanje PE v korpusu GOS

Ena izmed teženj sodobnega razvoja korpusnega jezikoslovja je izpopolnjevanje taktik iskanja stavčnih in večstavčnih enot v korpusih (Čermák, 2005a: 158), ki se ga lotevajo nekateri frazeologi (prim. Gantar 2004, Đurčo 2006, Čermák 2007). PE iz minimuma smo večinoma iskali z enostavnimi iskalnimi pogoji po standardiziranem zapisu.⁴ Enote smo skušali razdeliti na čim več sestavnih delov, ki bi jih lahko še uporabili za iskalne pogoje s prikazom obvladljivih seznamov konkordanc, saj smo želeli zajeti tudi variantne in prenovljene oblike. Poleg tega smo sestavinske variante iskali z različnimi sestavinami, npr. »imeti hudič« za enoto *Kadar ima vrag mlade, nima nikoli samo enega*. Pogosto je zadoščala že kombinacija polnopomenske in nepolnopomenske sestavine, npr. samostalnika in predloga: »vrabec v«, »golob na«, »na strehi« in »v roki«. Ker je korpus manjši po obsegu, smo nekatere enote lahko iskali tudi z iskalnimi pogoji iz ene same leme, npr. »pameten« za enoto *Pametnejši odneha* (52. mesto), kar v večjih pisnih korpusih ni možno. V primerih PE z arhaizmom, nekrotizmom ali kakšno drugo posebnostjo smo iskali z enim samim leksemom po pogovornem zapisu, npr. »gliha« pri enoti *Gliha vkup štriha* (239. mesto), »Trsti« pri enoti *Le povrsti, kakor so hiše v Trsti* (216. mesto) in »komod«/»komot«/»kamot« pri enoti *Čez komod ga ni* (271. mesto).

Iz izkušenj s korpusoma FidaPLUS in Gigafida ter primerjave s podobnimi korpusnimi raziskavami (Gantar 2004), lahko pričakujemo, da je delež stavčnih frazemov in med njimi PE tudi v govornem korpusu dokaj nižji od deleža besednozveznih frazemov. Gantar (2004: 95) npr. ugotavlja, da je 10 do 200 zgledov za besednozvezne frazeme nizka, 200 do 500 srednja in 500 do 1000 visoka pogostnost, v primerjavi s tem pa smo ugotovili, da v FidaPLUS predstavlja povprečje že 101 zgledov PE (Meterc 2014).⁵ Precej besednozveznih frazemov smo našli že z iskalnimi pogoji za PE iz minimuma:

1. z iskalnim pogojem »na zob« (*Podarjenemu konju se ne gleda na zobe*): »je bla kr žleht a veš k ga je mal na zob vrgla«;
2. z iskalnim pogojem »je zlato« (*Čas je zlato*): »je zlat«, »un je zlata vreden če mene vprašaš« in »tist je zlatga dnarja vreden« in
3. z iskalnim pogojem »od srca« (*Daleč od oči, daleč od srca*): »se od srca nasmeje« in »se ti zvali kamen od srca«.

Najdene zglede navajamo v Razpredelnici 1 na koncu tega poglavja v podobi,

⁴ V večjem korpusu bi za izdelavo zahtevnejšega iskalnega pogoja uporabili konstrukcijski vzorec PE in ga zaradi možnosti zamenjave, izpusta in dodajanja sestavin v variantah in prenovitvah razčlenili na niz iskalnih pogojev, npr. »#2zc*_za_*_*_za_#2zc*«, »#2zc*_za_*_*_*_za_#2zc*« in »#2zc*_za_*_*_*_*_*_za_#2zc*« za PE *Vsi za enega, eden za vse* v korpusu FidaPLUS.

⁵ Povprečje (101,54 zgledov na PE) smo izračunali tako, da smo delili število vseh zgledov paremij (30462) s številom PE iz minimuma (300).



v kakršni so registrirani v korpusu oz. brez popravkov in vstavljanja ločil. Izmed neprototipnih pojavitev navajamo le po eno na PE. Govor različnih govorcev je ločen s poševnico. Navajamo tudi stopnjo poznanosti PE v odstotkih, njeno zaporedno številko v paremiološkem minimumu ter pogostnost v korpusih GOS in FidaPLUS. Zgledi so opremljeni tudi s podatki o tipu diskurza, v katerem se pojavlja, o kanalu in regiji snemanja.

Legenda:

Regija snemanja:

- KR (Kranj)
- LJ (Ljubljana)
- MB (Maribor)
- NM (Novo mesto)
- SLO (celotna Slovenija)
- JZ (jugozahodna Slovenija)
- JV (jugovzhodna Slovenija)
- SZ (severozahodna Slovenija)
- SV (severovzhodna Slovenija)

Kanal:

- TV (televizija)
- RA (radio)
- OS (osebni stik)
- TEL (telefonski pogovor)

Tip diskurza:

- JR (javni razvedrilni)
- JIN-IZ (javni informativno-izobraževalni)
- NN (nejavni nezasebni)
- NZ (nejavni zasebni)

Razpredelnica 1:

pozna- nost	mini- mum	paremiološka enota in zgledi iz GOS (tip diskurza, kanal, regija snemanja)	GOS	FP
99,4 %	7.	<i>Kjer se prepirata dva, tretji dobiček ima:</i> »aha tm kjer se prepirata dva/ in mi je rekla« (JR, TV, SLO)	1	114
99 %	18.	<i>Kdor čaka, dočaka:</i> »čakam že nekej časa /eee kdor čaka dočaka pravijo« (JR, RA, SLO)	1	207
99 %	19.	<i>Kdor drugim jamo koplje, sam vanjo pade:</i> »ma tud en nauk ne/ mhm/ kdor drugemu jamo koplje sam svoje sreče kovač [smeh]« (JR, TV, SLO)	1	113
99 %	22.	<i>Rana ura – zlata ura:</i> »nadaljujemo z našim bojem proti jutranji nenagrajenosti/ rana ura zlata ura« (JR, RA, SV)	6*	81
99 %	24.	<i>Kar danes lahko storiš, ne odlašaj na jutri:</i> »sej praim ne odlagaj nikol kaj lahko dans narediš ti jutri ni treba to si daj toti moto si daj« (NZ, OS, MB)	1	48



98,4 %	44.	<i>Vsak je svoje sreče kovač:</i> »ma tud en nauk ne/ mhm/ kdor drugemu jamo koplje sam svoje sreče kovač [smeh]« (JR, TV, SLO)	1	196
98,4 %	47.	<i>Vse ob svojem času:</i> »vse bomo obdelali/ vse ob svojem času« (JR, RA, JZ) »leta dva tisoč vse ob praven času ja« (JR, RA, SLO)	2	116
98,4 %	49.	<i>Hiti počasi:</i> »kaj menite ostali?!/ bom pa tako vprašala/ hiti počasi« (JIN-IZ, OS, MB)	1	99
98,4 %	52.	<i>Pametnejši odneha:</i> »a jez sem pa mislu de je miza rekla pametnejši odneha« (NZ, OS, LJ)	1	73
98,4 %	53.	<i>Brez dela ni jela:</i> »na primer ne vem eee brez dela Kikiriki ni jela« (JR, RA, JZ)	2*	68
98,1 %	61.	<i>Obljuba dela dolg:</i> »ampak ker obljuba dela dolg smo eee počakali« (JR, RA, JZ)	1	251
98 %	67.	<i>Dober glas seže v deveto vas:</i> »samo dober glas se širi v deveto vas ne« (NN, OS, MB) »boste slišali še njen glas kaj pravi ampak naj vas opozorim da seže daleč v deveto vas« (JR, RA, JZ)	2	397
97,8 %	79.	<i>Dobro blago se samo hvali:</i> »dobro blago se samo hvali slabega moram pa jaz« (JR, TV, SLO)	1	84
97,8 %	80.	<i>Pomagaj si kakor veš in znaš:</i> »pomaga si vsakdo kokr si more in ve« (JIN-IZ, TV, SLO)	1	30
97,5 %	87.	<i>Obleka dela človeka:</i> »ja ja sej sej sej pravim ne obleka naredi človeka zdaj če bi se jaz sleko bi tud po mojem mlajši zgleo« (JR, TV, SLO)	1	298
97,5 %	90.	<i>Mladost – norost:</i> »dragi moji gledalci a vidte kak je to mladost norost ne« (JR, TV, SLO)	1	115
97,5 %	91.	<i>Kar na dan z besedo:</i> »jov povej/ kr besedo na dan« (JR, TV, SLO)	1	101
96,8 %	108.	<i>Napad je najboljša obramba:</i> »jaz bi reku napad je najboljša obramba« (JIN-IZ, TV, SLO)	1	121
96,8 %	110.	<i>V vinu je resnica:</i> »v vinu se skriva ogromno resnice/ vino veritas« (JR, RA, JZ)	1	60
96,5 %	114.	<i>Kar je preveč, je preveč:</i> »ker kar je pa preveč je pa preveč no« (JIN-IZ, OS, KR) »ne preveč kar je preveč je preveč« (NZ, OS, MB)	2	242
96,5 %	115.	<i>Previdnost je mati modrosti:</i> »previdnost je mati modrosti« (JR, RA, JV) »previdnost je mati modrosti saj veste« (JR, RA, JV)	2	136
96,2 %	119.	<i>Skomnost je lepa čednost:</i> »no skomnost je pa tudi lepa čednost a ne?« (JIN-IZ, RA, SLO)	1	103
95,9 %	126.	<i>Vsaka šola nekaj stane:</i> »draga šola draga šola mpak vsaka šola neki stane ja« (JIN-IZ, TV, SLO)	1	106



95,6 %	136.	<i>Hočeš, nočeš, moraš:</i> »eh moreš to ni zaj hočeš nočeš moreš« (NN, OS, MB)	1	22
95,2 %	144.	<i>Bo, kar bo:</i> »ne vemo še napovedi ampak/ mmm bo kar bo« (JR, RA, SV) »sm reku kure kar bo bo« (NZ, TEL, NM)	2	416
93,9 %	160.	<i>Ni dima brez ognja:</i> »običajno je f Holivudu tak da kjer je dim je tud ogenj« (JR, RA, SV)	1	33
91,4 %	188.	<i>Enkrat je treba začeti:</i> »enkrat treba začet« (NN, OS, MB)	1	20
91,1 %	191.	<i>Kdor jezika špara, kruha strada:</i> »kdor jezika špara vina strada ne/ na zdravje živeli« (JR, TV, SLO)	1	16
90,5 %	192.	<i>Boljše nekaj kakor nič:</i> »a dve ure so pa zadost/ ja boljš k nč« (NZ, OS, LJ)	1	50
90,2 %	197.	<i>Gora ni nora, tisti je nor, ki gre gor:</i> »torej gora ni nora kikiriki je tisti ki gre gor« (JR, RA, JZ)	4*	32
87,7 %	218.	<i>Še precej vode bo steklo, preden se bo to uresničilo:</i> »da bo še veliko veliko postopkov in vode preteklo« (JIN-IZ, TV, SLO)	1	145
86,1 %	236.	<i>Najbolj neumen kmet ima najdebelejši krompir:</i> »zdej na vjem kaku pravjo to katjer ma ta debjel krompir kk je tist pregovor« (JR, TV, LJ) »veš neumnimu kmet zmeri najbl debel kmpir zraste ne« (NZ, OS, LJ) »pregovor pravi da neumnemu kmetu zraste najdebelejši krompir in to na asvaltu ampak kaj pa če ne zraste?« (JR, TV, SLO)	3	65
85,7 %	241.	<i>Pri denarju se vse neha:</i> »mhm gospa ministrca se bo vse na koncu končalo pri denarju?« (JIN-IZ, TV, SLO)	1	168
85,4 %	245.	<i>S silo nič ne dosežeš:</i> »seveda ne moreš biti hiter na silo se ne da s silo prej« (JR, TV, SLO)	1	35
84,8 %	252.	<i>Kar je, to je:</i> »okej kar je je prteklosti se ne da spremenit« (JIN-IZ, RA, SV) »o ja kar je je« (NN, OS, LJ) »zdej je ker je« (NZ, OS, LJ)	3	459
83,5 %	268.	<i>Naj se zgodi, kar se mora:</i> »naj se zgodi kar se mora« (JR, RA, JZ)	1	72
82,3 %	275.	<i>Lahko je biti poveljnik po bitki:</i> »ampak saj veste kako je eee po bitki smo generali vsi« (JIN-IZ, , SV)	1	114

2 Korpusni zgledi PE iz paremiološkega minimuma

Skupaj smo našli 37 PE iz paremiološkega minimuma in 55 zgledov zanje,⁶ kar znaša v povprečju 1,49 zгледа na PE. Število zgledov moramo dodatno zamejiti na 46 (oziroma 1,2 na PE), saj smo med zgledi našli 12 ti. *neprototipnih pojavitvev*, ki jih bomo predstavili v podpoglavju, upoštevali pa smo le po en neprototipen zgleđ na PE. Zaradi razlike med govorno in pisno komunikacijo ter tudi strukture korpusov smo pričakovali, da v korpusu govornjenih besedil ne bomo našli le enot, ki so med

⁶ Število zgledov bi lahko zmanjšali na 54, če križanje dveh enot, ki ga bomo omenili v nadaljevanju, ne bi šteli dvakrat.



najbolj pogostimi v pisnem korpusu, temveč tudi take, ki so med manj pogostimi. To se je potrdilo npr. pri enotah:

1. *Kdor jezika špara, kruha strada* (16 zgledov v FidaPLUS);
2. *Enkrat je treba začeti* (20) in
3. *Hočeš, nočeš, moraš* (22).

Obenem v GOS nismo zasledili nekaterih najbolj poznanih ter tudi najbolj pogostih enot v korpusu FidaPLUS, npr. prvih treh pregovorov v paremiološkem optimumu:

1. *Denar je sveta vladar* (419. zgledov v FidaPLUS);
2. *Jabolko ne pade daleč od drevesa* (313) in
3. *Boljši je vrabec v roki kakor golob na strehi* (323).⁷

S podatkom o številu besed v GOS (1035101) lahko izračunamo, da se PE iz minimuma pojavi na okoli 22502 besed. Zanimivo je, da je ta ocena zelo blizu oceni, na koliko besed se PE iz minimuma pojavi v korpusu pisnih besedil; v doktorski raziskavi smo našli skupno 30462 zgledov v šestomilijonskem korpusu FidaPLUS, kar zneso okoli 19696 besed na eno zgled PE iz minimuma. Čermák (2007: 570) je v korpusni raziskavi iz leta 1997, ko je češki korpus pisnih besedil obsegal okoli 23 milijonov besed, na podlagi 284 zgledov za 99 PE iz Schindlerjevega češkega minimuma izračunal, da bi se ena PE iz minimuma morala v korpusu pojaviti na okoli 80000 besed. Njegov rezultat lahko v luči obeh slovenskih raziskav pripišemo ožjemu naboru enot v češkem minimumu, do nastanka katerega je bil kritičen tudi sam Čermák (2003: 15–16).

Če upoštevamo vseh 55 najdenih zgledov in obseg korpusa (7200 minut), lahko izračunamo, da bi se enota slovenskega paremiološkega minimuma ob podobni strukturi govora, kakršna je v korpusu, morala v povprečju pojaviti približno na dve uri (na 130,91 minut).⁸ Skupina raziskovalcev (Mehl, Vazire, Ramírez-Esparza, Slat-cher, Pennebaker 2007) je v nekajdnevni poskus s snemalnimi napravami vključila 396 govorcev in ugotovila, da tako moški kot ženske izrečejo okoli 16000 besed na dan, medtem ko pri dveh tretjinah govorcev število besed niha med 11000 in 25000. Kljub temu izračunu in podatku o številu besed v korpusu, je težko trditi, da povprečen govorec na dan izgovori eno PE iz minimuma, saj struktura korpusa ni osnovana na oceni distribucije tipov diskurzov v vsakdanji komunikaciji enega govorca. Lahko pa si zamislimo opazovalca, ki bi vsak dan osem ur poslušal govor v različnih situacijah in bi v povprečju slišal štiri enote iz minimuma.

2.1 Nprototipske pojavitve med korpusnimi zgledi

Nprototipne pojavitve so pojavitve PE izven njihove osnovne funkcije (2007:

⁷ V korpusu GOS nismo našli prvih desetih PE slovenskega optimuma. Našli smo šele 11. enoto *Kdor čaka, dočaka*. Če ne upoštevamo mere poznanosti enot oziroma namesto prvih desetih enot iz optimuma za primerjavo vzamemo prvih deset enot po pogostnosti v FidaPLUS, ugotovimo, da smo med njimi našli le tri enote: *Kar je, to je* (459 pojavitve, 5. mesto po pogostnosti), *Bo, kar bo* (416, 7.) in *Dober glas seže v deveto vas* (397, 8.).

⁸ Če upoštevamo le število, ki vključuje prototipne zgledje in po en nprototipni zgled na PE (46), se ta časovni interval podaljša na približno dve uri in pol (156,52 minut).



572–73). Zato v razpredelnici število pojavitev ob njih označujemo z *. Čermák izpostavlja primer rabe pregovorov in rekov v kvizih, pogosto pa so povezane tudi s popularizatorji (npr. naslovi del), ki dodatno krepijo njihovo rabo. Tudi v našem vzorcu se dve PE pojavita v (istem) radijskem kvizu, v katerem so eno izmed sestavin nadomestili z medmetom *kikiriki*: »torej gora ni nora kikiriki je tisti ki gre gor« (4 zgledi) in »ne vem eee brez dela kikiriki ni jela«. Poleg tega gre pri vseh šestih zgle-
dih PE *Rana ura – zlata ura* za naslov jutranje rubrike na nekem drugem radiu. Če ne upoštevamo neprototipnih pojavitev, se najpogosteje pojavita PE *Najbolj neumen kmet ima najdebelejši krompir* in *Kar je, to je s tremi zgledi*.

2.2 Razpon pojavnih oblik PE med korpusnimi zgledi

Glede vprašanja variant ter prenovitev PE upoštevamo klasifikacije, ki jih je v slovenskem prostoru predstavila Erika Kržišnik (1990: 403, 1998: 111). Med zgledi PE iz minimuma v korpusu GOS najdemo vrsto variantnih oblik, npr. sestavinsko varianto *Obleka naredi človeka (Obleka dela človeka)* ali pa skladijsko varianto *Kar bo, bo (Bo, kar bo)*. Razumljivo je, da se v govornem korpusu zvrstnost slovensčine kaže v večji meri kot v pisnih korpusih. Našli smo nekaj pogovornih in narečnih variant, npr.

1. »eee skratka pomaga si vsakdo kokr si more in ve«;
2. »katjer ma ta debjel krompir« in
3. »veš neumnimu kmet zmeri najbl debel kmpir zraste ne«.

Očitno razhajanje variante, ki prevladuje v rabi, z uslovarjeno podobo (ti. *ničto varianto*) smo morali upoštevati pri naslednjih enotah:

1. *Čas je zlato (SSKJ) – Čas je denar* (113. mesto v minimumu);
2. *Nasprotja se privlačujejo (SSKJ) – Nasprotja se privlačijo* (112. mesto) in
3. *Govoriti je srebro, molčati pa zlato (SSKJ) – Molk je zlato* (294. mesto).

Enoto *Lahko je biti poveljnik po bitki* smo iskali tudi s sestavino *general*. Našli smo le en zgled in sicer v varianti *Po bitki smo generali vsi*, ki se je po naši oceni razširila zaradi istoimenske pesmi skupine Pepel in kri v vlogi popularizatorja. O tem priča tudi 114 zgledov v korpusu FidaPLUS, med katerimi ni variant, ki bi bile blizu ničti varianti, ter spletna raziskava, saj so med 41 variantami, ki so jih dopisali anketiranci, le PE s sestavino *general*. Našli smo še primer latinske izvirne PE, ki se pojavlja hkrati s slovensko ustreznico: »v vinu se skriva ogromno resnice/ vino veritas«.

Med zgledi PE smo priča tudi bogatemu naboru prenovitvenih postopkov:

1. z vrinjenimi, zamenjanimi ali podvojenimi sestavinami (»da bo še veliko veliko postopkov in vode preteklo«, »Kdor jezika špara, vina strada«);
2. s podaljšavo PE (»dobro blago se samo hvali slabega moram pa jaz«), razgradnjo (»sedaj boste slišali še njen glas kaj pravi ampak naj vas opozorim da seže daleč v deveto vas«) in
3. z aluzijo na pregovor (»zdej na vjem kaku pravjo to katjer ma ta debjel krompir kk je tist pregovor«).

Razpredelnica 2:

tip	delež v GOS	delež (zgledi)	Podtip	delež v GOS	delež (zgledi)	kanal	delež v GOS	delež (zgledi)
javni	56,39 %	76,09 % (35)	informativni izobraže- valni	34,12 %	21,74 % (10)	televizija	10,5 %	10,87 % (5)
						radio	9,19 %	6,52 % (3)
						osebni stik	14,88 %	4,35 % (2)
			razvedrilni	22,7 %	54,35 % (25)	televizija	10,14 %	23,91 % (11)
						radio	12,13 %	30,43 % (14)
nejavni	43,61 %	23,91 % (11)	nezasebni	15,06 %	10,87 % (5)	telefon	3,27 %	0 %
						osebni stik	11,79 %	10,87 % (5)
			zasebni	28,55 %	13,04 % (6)	telefon	6,67 %	2,17 % (1)
						osebni stik	21,88 %	10,87 % (5)

Pogost je zamolk druge polovice PE, kakršne na primeru zgledov PE v češkem pisnem korpusu podrobneje opisuje Blatná (1998): »aha tm kjer se prepirata dva«. PE najdemo tudi v vprašalnem naklonu: »mhm gospa ministrica se bo vse na koncu končalo pri denarju?«. Že na majhnem vzorcu 46 zgledov lahko opazujemo (neustaljene) skladijske posebnosti PE v govorni rabi:

1. »dragi moji gledalci a vidte kak je to mladost norost ne« (*Mladost je norost*);
2. »enkrat treba začeti« (*Enkrat je treba začeti*) in
3. »Kr besedo na dan« (*Kar na dan z besedo!*).

Našli smo tudi enoto, o kateri lahko zaradi sobesedila sklepamo, da je nastala s križanjem enote *Potrpežljivost je božja mast* in *Skromnost je lepa čednost*: »potrpežljivost je božja mast ne/ ja... potrpežljivost je lepa čednost.«⁹ Primer je zanimiv, ker tudi v korpusu FidaPLUS z iskalnim pogojem »#lčednost/#llep !skromnost« najdemo veliko število prenovitev vzorca *X je lepa čednost*, med katerimi so najpogostejše prenovitve s sestavinami *točnost* (22), *vztrajnost* (19) in *hvaležnost* (7). Ob iskanju pregovora *Kdor drugim jamo koplje, sam vanjo pade* z iskalnim pogojem »kdor drug« smo našli antipregovor, ki je nastal s križanjem dveh PE: »ki pa ma tud en nauk ne/ mhm/ kdor drugemu jamo koplje sam svoje sreče kovač«. Teme antipregovorov kot posebnega paremiološkega žanra smo se lotili v ločenih prispevkih (Meterc 2012, 2014). Prav tako smo obravnavali križanje v ustaljenem antipregovoru *Rana ura – slovenskih fantov grob* (Meterc 2014: 114).

Naleteli smo na primere medravninskih frazeoloških transformacij – tudi iz PE, za katere nismo našli zgleda: »ona je tiha voda kot se temu reče«. Medravninskim

⁹ Enota se je s 70,3-odstotno poznanostjo sicer znašla izven minimuma na 388. mestu.

transformacijam smo se posvetili na primerih iz slovenščine in slovaščine (Meterc 2012). Ob iskanju enote *Hočeš, nočeš, moraš* smo bili pozorni na transformacije (adverbializacije) te PE v besednozvezni frazem *hočeš nočeš*: »pa vendarle hočeš nočeš naletite v medijih na tak in drugačen jezik«. Funkcijo prislova opravljajo tudi transformacije enote *Vse ali pa nič*, ki smo jo našli v treh zgledih in je nismo prišteli k PE:

1. »gremo na vse ali nič«;
2. »na vse ali nič ajde gremo strinjam se« in
3. »ja ne sej smo šli na vse ali nič tko da«.

Frazem *na vse ali nič* je v *Slovarju slovenskih frazemov* pojasnjen z razlago »tako da se največ dobi ali pa izgubi« (Keber 2012: 1075). V korpusu FidaPLUS najdemo PE *Vse ali nič*:

1. »Ali vse ali pa nič. Obljubil si mi!«;
2. »so dobili javno žaljiv odgovor. Vse ali nič. NDR naj z eno potezo popolnoma zgine, ali tega ne prizna.« in
3. »da sistemi, ki delujejo po logiki „vse ali nič!“, ki bi torej „raje“ nič, če ne morejo dobiti vsega«.

Iz sobesedila je tu razvidna funkcija stavčnega frazema, ki je v korpusu GOS nismo zasledili. Dva zgleda iz GOS kažeta ponovitev istih stavčnih frazemov, kar lahko razumemo kot frazeološki idiolekt določenega govorca (v tem primeru radijskega spikerja): »tak da okej brez panike ampak vseeno/ previdnost je mati modrosti saj veste« in »tako da eee/ ni takšne panike ampak vsak v vsakem primeru previdnost je mati modrosti«.

2.3 Uvajalna sredstva PE v govornem korpusu

Sredstvo, ki ima funkcijo uvajanja drugih jezikovnih sredstev (zelo pogosto frazemov) imenuje Čermák (2004: 71–72) *uvajalno sredstvo* (po češko *introduktor*). Na podlagi korpusnih raziskav ugotavlja, da so v okviru frazeologije najbolj pogosta prav v paremiologiji, v njihovi vlogi pa nastopajo samostalniki, glagoli, prislovi, vezniki ter besedne zveze (Čermák 2004: 75). V našem vzorcu korpusnih zgledov smo taka sredstva našli v desetih primerih. Med njimi najdemo tipične samostalniške žanrske oznake, ki jih našteva tudi Čermák: »nauk«, »moto«, »po logiki« in »pregovor pravi«. Zadnja oblika sodi tudi v nabor večbesednih uvajalnih sredstev, ki se s pogosto rabo ustaljujejo. Tudi preostala najdena uvajalna sredstva so večbesedna in se sklicujejo na kolektivno vedenje, ki ga posredujejo PE.

Umeščena so:

1. pred PE (»sej praim«, »sej pravim ne«, »a vidte kak je to«, »saj veste kako je«, »veš kak pravijo nek«) ali
2. neposredno za njimi (»saj veste«, »pravijo«).

V enem zgledu sta prisotni dve uvajalni sredstvi; večbesedno uvajalno sredstvo pred PE in žanrska oznaka »moto« po zaključku PE: »Sej praim ne odlagaj nikol kaj lahko dans narediš ti jutri ni treba to si daj toti moto si daj.« Posebnost je vprašalna oblika uvajalnega sredstva: »veš kak pravijo ne/ kak?/ dobro blago se samo hvali«.

Z uvajalnimi sredstvi kot iskalnimi pogoji lahko najdemo tudi PE, ki niso vključene v minimum. Oglejmo si nekaj primerov, ki smo jih v GOS našli z uvajalnimi

sredstvi *pregovor, rek, včasih so rekli* in *v stilu*:

1. »no torej eem/ dopolni pregovor [glas] na dan prežene doktorja fstran«;
2. »tkole pravi str čebelji pregovor ki provzaprov sploh ni čebelji ampak je tadžiški ne boj se človeka ki močno kriči boj se tistega ki tiho sedi«;
3. »je povedal da rek nisi pravi Slovenec če ne osvojiš Triglava ne drži«;
4. »eee sej včasih so rekli dajte ljudem igre pa pa bodo pozabli na to« in
5. »v stilu prve ne pozabiš nikoli«.

Ker se temi zbiranja frazeološkega gradiva s tem pristopom na tem mestu ne moremo podrobneje posvetiti, se bomo k njemu morali vrniti v prihodnosti.

2.4 Distribucija PE po tipih diskurza in govornih kanalih

Delež posameznih tipov in podtipov diskurzov ter govornih kanalov med zgledi PE iz slovenskega paremiološkega minimuma sopostavimo s strukturo celotnega govornega korpusa, objavljeno na spletni strani GOS.¹⁰ Zastopanost PE iz slovenskega paremiološkega minimuma je v javnem tipu diskurza občutno večja kot v nejavnem. Delež PE iz nejavnega diskurza (23,91 %) je skoraj za polovico manjši od zastopanosti besedil nejavnega diskurza v korpusu (43,61 %). Ta razlika se krepi predvsem na račun polovičnega deleža nejavnega zasebnega diskurza med najdenimi PE, za tretjino pa je zmanjšan tudi delež iz nejavnega nezasebnega diskurza.

Med PE iz javnega diskurza ima podtip razvedrilnega diskurza (54,35 %) občutno prednost pred podtipom informativno-izobraževalnega diskurza (21,74 %). Treba je sicer upoštevati, da so se snovalci korpusa odločili uvrstiti prosti moderirani radijski program v razvedrilni diskurz, ker ta »skuša podajati informativne vsebine na razvedrilen način« (Verdonik, Zwitter Vitez 2011: 49). Največ je PE, ki se v okviru razvedrilnega diskurza pojavljajo v radijskih besedilih. Predstavljajo približno tretjino vseh najdenih PE (30,43 %), delež radijskih besedil tega podtipa pa v korpusu znaša le 12,13 %. Opazno povečan je tudi delež PE iz televizijskega javnega razvedrilnega diskurza (23,91 %), če upoštevamo, da televizijska besedila tega tipa predstavljajo 10,14 % vseh posnetkov v korpusu. Tako delež PE iz televizijskih kakor tudi radijskih posnetkov iz javnega informativno-izobraževalnega diskurza sta primerljiva z zastopanostjo besedil v korpusu. Občutno nizek je delež PE v osebnem stiku javnega izobraževalno-informativnega diskurza (4,35 %), če ga primerjamo z deležem besedil tega tipa v korpusu (14,88 %). Korpus GOS vsebuje tudi podsklop šolskega diskurza (javni izobraževalni diskurz preko kanala osebnega stika), ki obsega 10 %. Našli smo le dve PE, kar sicer priča o prisotnosti paremiologije v šolskem pouku, je pa za kakršnekoli druge zaključke premalo. Delež PE iz osebnega stika v okviru nejavnega nezasebnega podtipa diskurza (10,87 %) je primerljiv z deležem besedil tega podtipa v korpusu (11,97 %). Izrazita razlika se kaže v okviru osebnega stika nejavnega zasebnega diskurza, saj je delež PE prepolovljen (10,87 %) v primerjavi z deležem takih besedil v korpusu (21,88 %). Zelo majhen je tudi delež enot iz telefonskih pogovorov nejavnega nezasebnega diskurza, saj smo na tem kanalu našli le eno PE iz minimuma (2,17 %). V korpusu pa je delež telefonskih pogovorov bistveno večji (6,67 %).

¹⁰ Razpredelnico navajamo neposredno pod razpredelnico PE in najdenih zgledov na koncu razprave.

Če primerjamo deleže najdenih zgledov z deleži v celotnem korpusu le na podlagi kanala, ugotovimo, da so PE najbolj zastopane v radijskem etru. Delež takih zgledov je za polovico višji (36,96 %) od deleža radijskega govora v korpusu (21,49 %). Z upoštevanjem neprototipnih pojavitev, ki so vse z radia, bi delež presegel polovico vseh najdenih zgledov (56,52 %). Podobno razmerje se kaže pri televiziji (34,78 % med zgledi in 20,64 % v strukturi korpusa). Delež zgledov PE iz minimuma na kanalu osebnega stika, ki dosega skoraj polovico (48,55 %) vseh posnetkov v govornem korpusu, je občutno znižan (26,09 %). Zastopanost PE v posnetkih telefonskih pogovorov (2,17 %) je zelo nizka v primerjavi z zastopanostjo teh posnetkov v korpusu (9,94 %).

3 Zaključki

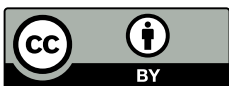
Potrdilo se je pričakovanje, da v korpusu govornih besedil ne najdemo le PE, ki so med najbolj pogostimi v pisnem korpusu, temveč tudi take, ki so med najmanj pogostimi. V velikem govornem korpusu bi verjetno našli vse PE iz paremiološkega minimuma in optimuma. Predstavili smo povprečno število (prototipnih) pojavitev PE iz minimuma v korpusu GOS (1,2 zgleda). Ocenili smo tudi, da se PE iz slovenskega paremiološkega minimuma v povprečju v korpusu govornih besedil pojavi na okoli 20000 besed, kar je povsem primerljivo s podatkom iz korpusa pisnih besedil. Ugotovili smo, da že 46 zgledov PE zadošča za pester prikaz variantnosti, prenovitev in nekaterih drugih fenomenov (antipregovorov, transformacij ter celo frazeoloških specifik določenega idiolekt). Med zgledi prevladujejo PE iz javnega razvedrilnega diskurza. S pomočjo korpusa smo izdelali tudi približno oceno pogostnosti pregovorov in rekov iz minimuma v omejenem časovnem intervalu, ki bi jo bilo zanimivo primerjati z izsledki prav tako časovno omejenega terenskega zapisovanja kot dopolnilne raziskave.

Ker gre v raziskavi za sondo, ki predstavlja presečišče omejene množice govornih besedil z omejeno množico najbolj poznanih PE, ne smemo zanemariti vprašanj glede območij, ki so ostala izven njega. Raziskavo bi lahko razširili še na druge PE, ki so v anketi presegle vsaj prag 50-odstotne poznanosti (takih je bilo 546). Z iskalnimi pogoji smo povsem po naključju našli nekaj PE, ki niso bile uslovarjene v virih za izhodiščni korpus in zato niso bile vključene v spletno raziskavo za določitev paremiološkega minimuma, npr.:

1. »ampak mene so še tako učili tak daj no nekej imeti v rokah da te vrag brez dela ne najde ko pride po tebe«;
2. »hjec mora biti pri nas na kmetiji«;
3. »sej je bil strošek ampak ziher je ziher« in
4. »veš kk bom rekla vse se da če se hoče«.¹¹

Odrpoto ostaja vprašanje neparemioloških stavčnih frazemov. Sliko tega dela frazeologije bi v prihodnosti lahko dopolnili z uporabo pisnih in govornih korpusov. V kor-

¹¹ Anketiranci so po zaključku ankete odgovarjali na vprašanje, ali se spominjo še kakšnega pregovora, reka ali podobne enote, ki je v jedru vprašalnika niso zasledili. Mnoge PE se v navedkih ponavljajo, saj si ne morejo zapomniti vseh 918 enot iz jedra. O tem, da je množica PE v anketi dovolj reprezentativna, pričča dejstvo, da le en pregovor (*Po jutru se dan pozna*) izmed 24 PE, ki se med odgovori pojavljajo več kot petkrat, ni bil vključen vanjo. Enote v GOS sicer nismo našli, v FidaPLUS pa se pojavi 32-krat.



pusu GOS smo sicer našli tudi naslednje stavčne frazeme, ki še niso bili uslovarjeni:

1. »pa še tak na ših tu k je blu zdele pa čao Miki«;
2. »je adijo pamet«;
3. »ooo to mi deli« in
4. »ne morš verjet jebi ga lajf saks«.

Zadnji zgled odpira tudi vprašanje citatne frazeologije, ki bi ji bilo prav tako treba posvetiti pozornost. Stavčni frazem *Ne moreš verjeti* je v primerjavi s PE iz minimuma zelo številčen (35), zgledi pa kažejo tudi številne različne pogovorne in narečne prvine, npr.:

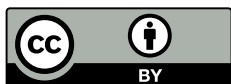
1. »nea morš verjet«;
2. »ne rem vervati«;
3. »na morš vert« in
4. »človik ne more vrjet«.

Domnevamo, da je za pretežen (neparemiološki) del stavčnih frazemov značilna pogostnost, ki je večja od pogostnosti PE, a vseeno nižja od pogostnosti besednozveznih frazemov. Pri nalogi sistematičnega zbiranja doslej še neuslovarjene frazeologije ne smemo zanemariti tudi možnosti klasičnega terenskega zapisovanja enot.

Tako analizo variant in prenovitev kakor tudi analizo distribucije PE po tipih diskurza in govornih kanalih bi bilo zanimivo v prihodnosti, ko bomo verjetno z večanjem korpusa imeli tudi več gradiva, navezati še na vprašanje socialne zvrstnosti frazeologije (prim. Kržišnik 1998). Idealno bi bilo primerjati enako številčna korpusa govorjenih in pisnih besedil. Šele takrat bomo lahko primerjali tako pogostnosti širokega razpona stavčnih frazemov ter besednozveznih frazemov v obeh korpusih, kakor tudi razmerji med frazemi teh dveh ravni v obeh tipih korpusov. Končno bomo lahko nazorno prikazali prevlado določenih PE (in frazemov na splošno) v govorni rabi in določenih v pisni rabi. Zaenkrat lahko določimo le nabor nadpovprečno aktivno rabljenih enot, ki se v korpusu pojavljajo podpovprečno, če podatke o pogostnosti v pisnem korpusu sopostavimo s podatki sociolingvistične raziskave poznavanja PE. Na ta način je na suprasemantični ravni možno določiti tudi diamedialno razliko med enotami, ki se v enem jeziku pogosteje pojavljajo v pisni, v drugem pa v govorni rabi (Đurčo, Meterc 2013: 30, Meterc 2014a: 154–57). Zanimivo bi bilo tudi primerjati PE iz minimumov drugih jezikov v njihovih govornih korpusih.

Glede metodološke nadgradnje paremiološkega minimuma v optimum (Đurčo 2006: 17) si lahko v okviru empirične paremiologije zastavimo nalogo »govornega« optimuma, ki bo mogoč v prihodnosti z upoštevanjem podatkov o pogostnosti PE v govorjeni slovenščini. Govorni in pisni optimum bomo lahko primerjali med seboj. Poleg tega bi lahko izdelali skupni pisno-govorni optimum, v katerem bi združili podatke o pogostnosti enot iz primerljivo velikega pisnega in govornega korpusa.

Zaključimo lahko, da nam ob primernem okviru, ki nam ga ponuja minimum kot zamejitev sodobne aktivne paremiologije, tudi enostavne poizvedbe v majhnem govornem korpusu prinesejo zanimive ugotovitve in del slike o aktualni govorni rabi stavčnih frazemov v slovenščini.



VIRI IN LITERATURA

- Renata BLATNÁ, 1998: Czech proverbs vanishing in the black hole. Elaborated on the material of the Czech National Corpus. *EUROPHRAS ,97 phraseology and paremiology International symposium*. Ur. P. Ďurčo. Bratislava: Akadémia PZ. 24–29.
- František ČERMÁK, 2003: Paremiological Minimum of Czech: The Corpus Evidence. *Flut von Texten – Vielvalt der Kulturen. Ascona 2001 zur Methodologie und Kulturspezifik der Phraseologie*. Ur. H. Burger, A. Häcki Buhofer in G. Greciano. Hohengehren: Schneider Verlag. 15–31.
- , 2004: Besedilna uvajalna sredstva rečenic in drugih stalnih besednih zvez. *Slavistična revija* 52/1. 71–88.
- , 2005: Korpus, informace a lingvistika. *Přednášky z 48. běhu Letní školy slovan-ských studií*. Ur. J. Hasil. Praga: Filozofická fakulta Univerzity Karlovy. 15–24.
- , 2005a: Jezikovni korpus: sredstvo in vir spoznanj. Študije o korpusnem jeziko-slovju. Ur. V. Gorjanc in S. Krek, Ljubljana: Krtina. 137–71.
- , 2007: *Frazeologie a idiomatika česká a obecná*. Praga: Univerzita Karlova v Pra-ze, Karolinum.
- , 2009: *Slovník české idiomatiky a frazeologie 4. Výrazy větné*. Praha: Leda.
- Peter ĎURČO, Matej METERC, 2013: Empirične paremiološke raziskave tipov ekviva-lentnosti in suprasemantičnih razlik v slovenščini in slovaščini. *Slavia centralis* 6/2. 20–37.
- Peter ĎURČO, 2002: K výskumu súčasnej živej slovenskej paremiológie. *Přednášky XXXVIII. letnej školy slovenského jazyka a kultúry. Studia Academica Slovaca* 31. Ur. J. Mlacek. Bratislava: Stimul. 51–60.
- , 2006: Methoden der Sprichwortanalysen oder Auf dem Weg zum Sprichwörterop-timum. *Phraseology in Motion. Methoden und Kritik. Akten der Internationalen Tagung zur Phraseologie (Basel, 2004)*. Ur. A. Häcki Buhofer in H. Burger. Balt-mannsweiler: Schneider Verlag Hohengehren. 3–20.
- Polona GANTAR, 2004: *Frazem in njegovo besedilno okolje. Doktorska disertacija*. Ljubljana: FF.
- Peter GRZYBEK, Christoph CHLOSTA, 2008: Some Essentials on the Popularity of (American) Proverbs. *Festschrift on the Occasion of Wolfgang Mieder's 65th Birthday*. Ur. K. McKenna. Burlington: University of Vermont. 95–110.
- Janez KEBER, 2011: *Slovar slovenskih frazemov*. Ljubljana: Založba ZRC, ZRC SAZU.
- Erika KRŽIŠNIK, 1990: Tipologija frazeoloških prenovitev v Cankarjevih prozih be-sedilih, *Slavistična revija* 38/4. 399–421.
- , 1998: Socialna zvrstnost in frazeologija. *34. seminar slovenskega jezika, literatu-re in kulture. Zbornik predavanj*. Ur. E. Kržišnik. Ljubljana, 53–69.



Matthias R. MEHL, Simine VAZIRE, Nairán RAMÍREZ-ESPARZA, Richard B. SLATCHER, James W. PENNEBAKER, 2007: Are women really more talkative than man? *Science magazine* 317. Na spletu.

Matej METERC, 2012: Transformacije frazeoloških enot in prepoznavanje transformacij v drugem slovanskem jeziku med slovenskimi in slovaškimi govorcji. *Slavica Slovenica. Individualna in kolektivna dvojezičnost*. Ur. P. Stankovska, M. Wtorowska in J. Pallay. Ljubljana: FF. 323–30.

--, 2014: Je prihodnost slovenskih antipregovorov (le) pregovorna? *Prihodnost v slovenskem jeziku, literaturi in kulturi*. 50. Seminar slovenskega jezika, literature in kulture. Ur. H. Tivadar. Ljubljana: FF. Center za slovenščino kot drugi/ tuji jezik. 113–16.

--, 2014a: *Primerjava paremiologije v slovenskem in slovaškem jeziku na osnovi paremiološkega optimuma. Doktorska disertacija*. Ljubljana: FF.

--, 2015: Online questionnaire providing information on most well-known and well-understood proverbs in Slovene language. *Phraseologie im Wörterbuch = Phraseology in Dictionaries, (Mednarodna knjižna zbirka Zora, 97)*. Ur. V. Jesenšek in P. Grzybek. Maribor: Mednarodna založba Oddelka za slovanske jezike in književnosti, Filozofska fakulteta. 205–15.

Jozef MLACEK, 1983: Problémy komplexného rozboru prísloví a porekadiel. *Slovenská reč* 48/2. 129–40.

Josip PAVLICA, 1960: *Frazeološki slovar v petih jezikih*. Ljubljana: DZS.

Grigorij PERMJAKOV, 1988: *Osnovy strukturnoj paremiologii*. Moskva: Nauka.

Darinka VERDONIK, Ana ZWITTER VITEZ 2011: *Slovenski govorni korpus GOS*. Ljubljana: Trojina, zavod za uporabno slovenistiko.

Korpus govornjene slovenščine GOS. Na spletu.

Korpus slovenskega jezika FidaPLUS. Na spletu.

Slovar slovenskega knjižnega jezika. Ljubljana: ZRC SAZU, DZS, 2008. Na spletu.

SUMMARY

Phraseology in spoken language is worth examining firstly because it presents the basic form of everyday communication and secondly, because this aspect, as Čermák points out, has so far been neglected. This paper shows how a paremiological minimum can be used as the starting platform for the research of paremiological units in the spoken corpora. Although the corpus of spoken Slovene (GOS) is quite small, we assumed we would be able to get quite an informative picture of Slovene paremiology by checking which units from the list of the 300 most well-known paremiological units (mostly proverbs and sayings), appear in it. In the survey, we used the experimental corpus that consisted of 918 units from two lexicographical sources. The first version of the paremiological minimum is based on the answers of 316 respondents from all over Slovenia as well as abroad. The units were presented to them in the full