

Predmetna obrada i izrada paremiološkog tezaurusa suvremenih hrvatskih poslovice

Subject indexing and creation of the paremiological thesaurus of contemporary Croatian proverbs

Borna Petrović¹

SAŽETAK: Cilj je ovoga rada bio predmetno obraditi korpus suvremenih hrvatskih poslovice terminima u denotativnom (doslovnom) i konotativnom (prenesenom) značenju te na temelju korištenih termina izraditi dvostruki tezaurus suvremenih hrvatskih poslovice koji će biti korišten za pretraživanje na Portalu hrvatskih poslovice. Uzorak istraživanja činio je korpus od 108 suvremenih hrvatskih poslovice preuzet iz upitnika nastalog u sklopu projekta određivanja hrvatskog paremiološkog minimuma i optimuma. Metodologija obrade poslovice podrazumijevala je dva dijela: predmetnu obradu poslovice i izradu dvostrukog tezaurusa. Predmetna se obrada temeljila na Lancasterovoj metodologiji, a odvijala se u sedam faza: denotativna analiza poslovice, traženje sinonima, kvazisinonima i srodnih termina denotativnim terminima uz njihovo popisivanje, konotativna analiza poslovice određivanjem očemnosti korpusnom analizom, dodavanje termina za predmetno označivanje u konotativnom značenju, traženje sinonima, kvazisinonima i srodnih termina u rječniku i jezičnim portalima te završna provjera konotativnog značenja korištenjem umjetne inteligencije uz ručnu provjeru dobivenih izvora. Izrada oba tezaurusa provedena je prema Shearerovoj metodologiji. Rezultati su pokazali da je poslovice dodijeljeno ukupno 1074 termina, od kojih su 586 denotativni i 488 konotativni. Denotativni tezaurus sadrži 29 glavnih faseta i 380 podfaseta, a konotativni 23 fasete i 244 podfaseta. Originalnost ovog istraživanja leži u činjenici da je ovo prvi hrvatski paremiološki tezaurus i ujedno prvi tezaurus koji zasebno obrađuje građu prema denotativnom i konotativnom značenju. Praktična primjena ovih rezultata ogledat će se u implementaciji tezaurusa na Portal hrvatskih poslovice, čime će se, uz omogućavanje jednostavnijeg, preciznijeg i sveobuhvatnijeg pretraživanja, doprinijeti digitalizaciji i popularizaciji hrvatske paremiološke baštine u pedagoškom i znanstvenom smislu.

KLJUČNE RIJEČI: Portal hrvatskih poslovice, poslovice, predmetna obrada, predmetno označivanje, tezaurus

ABSTRACT: The aim of this paper was to subject-index a corpus of contemporary Croatian proverbs using denotative (literal) and connotative (figurative) terms, and to develop a dual thesaurus for searching on the Croatian Proverbs Portal. The research sample consisted of 108 proverbs collected as part of a project defining the Croatian paremiological minimum and optimum. The methodology included two main steps: subject processing of proverbs and creation of a dual thesaurus. Subject processing followed Lancaster's methodology and consisted of seven phases: denotative analysis, identification of synonyms and related terms, connotative analysis by determining aboutness through corpus analysis, addition of connotative indexing terms, further synonym searches in a dictionary and language portals, and final verification of connotative meanings using artificial intelligence and manual review. Both thesauri were constructed according to Shearer's methodology. The results showed that

Ovaj rad nastao je na temelju metodologije i rezultata diplomskog rada autora. Diplomski rad obranjen je 11. srpnja 2025. godine pod mentorstvom izv. prof. dr. sc. Kristine Feldvari i sumentorstvom izv. prof. dr. sc. Melite Alekse Varge, a pohranjen je na institucijskom Repozitoriju Filozofskog fakulteta Sveučilišta Josipa Jurja Strossmayera u Osijeku (DABAR) na sljedećoj poveznici: <https://urn.nsk.hr/urn:nbn:hr:142:488805>.

¹ Borna Petrović, mag. informatol. et mag. educ. philol. angl., Sveučilište Josipa Jurja Strossmayera u Osijeku, Filozofski fakultet, Odsjek za informacijske znanosti, bpetrovic@ffos.hr.

1,074 terms were assigned to the proverbs – 586 denotative and 488 connotative. The denotative thesaurus contains 29 main facets and 380 subfacets, while the connotative thesaurus includes 23 facets and 244 subfacets. The originality of this research lies in its presenting of the first Croatian paremiological thesaurus, also the first thesaurus to separately process denotative and connotative meanings. The practical outcome of this research will be the implementation of the thesaurus on the Croatian Proverbs Portal, enabling more precise and comprehensive searches, and contributing to the digitalization and popularization of Croatian paremiological heritage in educational and scientific contexts.

KEYWORDS: Croatian Proverbs Portal, proverbs, subject processing, subject indexing, thesaurus

1 Uvod

Ideja o izradi hrvatskog paremiološkog tezaurusa proizlazi iz činjenice kako hrvatski jezik, za razliku od drugih jezika kao što su engleski, bugarski, njemački i mađarski (usp. Wander, 1867-1880; Mieder, Kingsbury i Harder, 1996; Litovkina, 2005; Stojkova, 2007; Doyle, Mieder i Shapiro, 2012), nema sveobuhvatnu zbirku poslovice koja bi odgovarala suvremenoj paremiologiji i paremiografiji i kako je većina hrvatskih poslovice dokumentirana u zbirkama iz 19. i 20. stoljeća (Grzybek et al., 1993; Aleksa Varga, 2023). Napori za stvaranjem takve suvremene zbirke, inovativnog digitalnog alata koji bi podupro zaštitu, pristup i pedagošku diseminaciju hrvatskih poslovice, ostvareni su kroz izradu *Portala hrvatskih poslovice* (Portal hrvatskih poslovice, 2025). Poslovice kao jezične cjeline nose dva značenja – denotativno (doslovno) i konotativno (preneseno), a njihova se predmetna obrada na *Portalu* vršila prema oba. Tako je dobivene termine bilo potrebno sustavno organizirati u cjelinu kakvu će korisnici *Portala* moći koristiti za lak pronalazak poslovice ne samo prema željenim terminima/konceptima u oba značenja, već i prema novima koje im možda nisu pale na pamet. Ovakvu vrstu pretraživanja omogućuje upravo dvostruki tezaurus poslovice u denotativnom i konotativnom značenju. Teorijske i metodološke implikacije izrade ovakvog tezaurusa postavile su Aleksa Varga i Feldvari (2022), a ostvarene su predmetnom obradom korpusa suvremenih hrvatskih poslovice kao prethodnim korakom izrade hrvatskog paremiološkog tezaurusa, što su ujedno i ciljevi ovog rada.

2 Predmetna obrada

Predmetna obrada jedan je od temeljnih postupaka organizacije informacija na području knjižnične i informacijske znanosti čiji je primarni cilj što konkretniji opis teme, odnosno predmeta određenog dokumenta. Ovaj proces omogućuje raspoređivanje dokumenata prema njihovom sadržaju u fizičkim i digitalnim okruženjima, što je temelj za učinkovito dohvaćanje informacija u suvremenim informacijskim sustavima (Stipanov, 2007).

Iako je postupak predmetnog označivanja opisan u brojnim pravilnicima, priručnicima, standardima i udžbenicima, kao i od strane brojnih autora, u kontekstu ovog rada držat ćemo se Lancastera (1998), koji ovaj postupak opisuje u dva koraka: konceptualnu/intelektualnu/predmetnu analizu, prilikom koje se određuje predmet dokumenta, odnosno njegova *očemnost* (engl. *aboutness*) i prijevod rezultata dobivenih konceptualnom analizom u konkretne termine za predmetno označivanje. Pojam očemnosti u kontekst je informacijskih znanosti prvi uveo Fairthorne (1969) razlikujući ekstenzijsku očemnost, odnosno sadržaj inherentan dokumentu koji je svojom prirodom neodvojiv od

njega, i intenzijsku očemnost, odnosno ono što se može zaključiti iz sadržaja dokumenta i ovisi o interpretaciji predmetnog stručnjaka i korisničkim potrebama. Prilikom predmetne obrade, svaki se predmetni stručnjak treba voditi trima pitanjima: *Što je dokument?*, *O čemu je dokument?* i *Za koga je dokument?* (Langridge, 1989). Opis sadržaja dokumenta omogućuju jezici za predmetno označivanje, odnosno predmetni jezici – dokumentacijski jezici koji omogućuju dosljedan pristup sadržaju dokumenta u katalogu, bibliografiji ili kazalu i koji se sastoje od kontroliranog rječnika i prikazuju odnose između termina u njemu (semantičke strukture), a mogu sadržavati i sintaktička pravila za slaganje pojmova u nizove (Kolbas, 2000). Prema Rowley i Farrowu (2017), predmetni se jezici dijele na kontrolirane i nekontrolirane, pri čemu kontrolirani jezici koriste termine iz kontroliranog rječnika, dok nekontrolirani preuzimaju termine izravno iz teksta dokumenta. Kontrolirani predmetni jezici se dijele na klasifikacijske (npr. klasifikacijske sheme) i abecedne (npr. tezaursi i popisi predmetnih odrednica).

Izrada predmetnih jezika oslanja se na tri jamstva koja osiguravaju legitimnost temelja samog rječnika: jamstvu predložka, jamstvu uporabe i jamstvu strukture. Jamstvo predložka podrazumijeva empirijski razvoj rječnika na osnovu dokumenata koji se opisuju i osigurava da se terminologija predmetnog jezika razvija na temelju stvarnih potreba korpusa koji se obrađuje, a ne na teoretskim pretpostavkama. Jamstvo uporabe uzima u obzir rječnik korisnika i potrebu prilagodbe predmetnog jezika terminima koje korisnici upotrebljavaju, čemu ide u prilog Cutterova izjava kako je *uporaba vrhovni arbitrar*, pritom se odnoseći na uporabu od strane korisnika, a ne predmetnih stručnjaka. Ovo jamstvo osigurava poboljšanu učinkovitost dohvaćanja informacija u sustavima. Naposljetku, jamstvo strukture opravdava uključivanje termina koji pružaju korisnu strukturalnu funkciju u predmetnom jeziku, omogućujući popunjavanje praznina u hijerarhiji ili okupljanje specifičnijih srodnih termina (Svenonius, 2000). Iako je uporaba jamstva prilikom izrade predmetnih rječnika vrlo korisna, ponekad se u ovom smislu susrećemo i s poteškoćama u definiranju domene rječnika, odnosno termina koje ćemo uključiti, a koje ne u rječnik. Naime, pojedine discipline, u prvom redu u društvenim i humanističkim znanostima, imaju razmjerno male tehničke rječnike te je predmet njihovog istraživanja često leksički neodređen, što za posljedicu ima nemogućnost jednostavnog izražavanja tog predmeta jednom riječju ili frazom. Najbolji je primjer ovog fenomena termin *kultura* koji, ovisno o predmetnom području, može imati brojna značenja, što nameće pitanje svrsishodnosti normativne kontrole ovakvog vokabulara (Svenonius, 2000).

3 Tezaurus

Terminološki rječnik koji sadrži sustavno organizirano stručno nazivlje danog znanstvenog područja okosnica je svake znanstvene discipline. Povijesno gledano, pojam *tezaurus* potječe od latinskog *thesaurus* i grčkog *θησαυρός* (*riznica, blago*), a u suvremenom smislu prvi puta ga je u upotrebu uveo Peter Mark Roget 1852. godine objavom *Tezaurusa engleskih riječi i fraza* koji je postao temelj za taksonomske rječnike, čime se koncept tezaurus počinje širiti kao naziv za leksikografski priručnik organiziran po konceptima i značenjima. U knjižničnoj i informacijskoj znanosti (engl. *Library and Information Science*, LIS) ovaj se termin odnosi na kontrolirani rječnik koji služi terminološkoj kontroli, odnosno prevođenju prirodnog jezika iz dokumenta kojeg se opisuje, jezika predmetnog stručnjaka i korisnika na kontrolirani jezik za predmetno označivanje, i obrnuto (Nikolić-Hoyt, 1996; Reitz, 1996; ISO, 1986).

U kontekstu LIS-a i istraživanja provedenog u ovome radu, razlikujemo dvije vrste tezaurusa: tezaurese opće namjene (engl. *General-Purpose Thesauri*) i strukturirane tezaurese, odnosno tezaurese za dohvaćanje informacija (engl. *Information Retrieval (IR) Thesauri*). Tezaursi opće namjene koriste se prije svega na području jezikoslovlja i uobičajeno se temelje na razini jednog leksema, čime su nepogodni za traženje i pretraživanje informacija, dok su IR tezaursi strukturirani te uspostavljaju hijerarhiju i odnose između termina, čime olakšavaju dohvaćanje informacija (Aitchison et al., 2000). U ovome ćemo se radu baviti opisom i izradom IR tezaurusa.

Tehnička dokumentacija koja prati razvoj i izradu tezaurusa vrlo je opširna, no u kontekstu ovoga rada ističemo standard ISO 2788:1986 – *Smjernice za izradbu i razvoj jednojezičnih tezaurusa* koji navodi kako učinkovitost predmetnog označivanja kao sredstva za identificiranje i pronalazak dokumenta ovisi o dobro složenom jeziku za predmetno označivanje, pri čemu se sastavljač takvog jezika suočava s tri glavne poteškoće: određivanjem predmeta dokumenta, odabirom termina koji primjereno sažimaju taj predmet i označavanjem odnosa između koncepata koje ti termini reprezentiraju (ISO, 1986).

Konkretno korake izrade tezaurusa, koji će ujedno biti i uvršteni u metodologiju ovoga istraživanja, predlaže Shearer (2004), a oni su:

1. prikupljanje prirodnih riječi iz dokumenata koji će biti obuhvaćeni
2. spajanje sinonima i razrješenje homografije
3. stvaranje faseta okupljanjem deskriptora
4. slaganje faseta i podfaseta
5. dodavanje novih deskriptora radi popunjavanja značajnijih praznina
6. notacija
7. identificiranje veza između deskriptora među fasetama
8. stvaranje abecednog kazala i
9. stvaranje abecednog prikaza.

Deskriptori, odnosno preferirani termini, jedni su od osnovnih gradivnih elemenata tezaurusa koji jedinstveno predstavljaju pojmove u kontroliranom rječniku. Odabir deskriptora temelji se na prethodno spomenuta tri jamstva: na jamstvu predložka empirijskim utemeljenjem deskriptora u analiziranoj građi, na jamstvu uporabe usklađivanjem termina korištenih kao deskriptora s rječnikom korisnika te na strukturnom jamstvu popunjavanjem hijerarhijskih praznina radi dosljedne organizacije. Uz to, deskriptori moraju biti jednoznačni, treba izbjegavati korištenje sinonimnih i dvostrukih termina kao deskriptora te, ukoliko je potrebno, mogu se koristiti napomene o uporabi i značenju termina (engl. *scope note*, SN) za razrješenje polisemije. Semantička mreža tezaurusa temelji se na tri vrste veza: hijerarhijski odnosi (općenitiji, širi termini – BT i specifičniji, uža termini – NT), odnosi ekvivalencije (uputnice s nepreferiranih na preferirane termine – USE i obratno – UF) i asocijativni odnosi (semantički/funkcionalno srodni termini izvan stroge hijerarhije – RT), što omogućuje dosljedno predmetno označivanje, a posljedično i intuitivno pretraživanje informacijskih sustava (Aitchison, Gilchrist i Bawden, 2000; Urbanija, 2005; ANSI i NISO, 2010).

Tezaurus se uobičajeno prezentira kroz abecedni i hijerarhijski prikaz, uz moguće grafičke sheme i permutirano kazalo složenih termina. Također je potrebno definirati tipografiju, pravila uporabe velikih i malih slova, format prikaza i opseg metapodataka (definicije,

napomene o upotrebi, frekvencije pojavljivanja) ovisno o kategorijama korisnika (stručnjaci, predmetni stručnjaci, krajnji korisnici). Grafički prikazi pomažu u vizualizaciji strukture, a abecedna kazala olakšavaju brzo traženje termina (Urbanija, 2005; ANSI i NISO, 2010).

4 Metodologija

4.1 Ciljevi istraživanja

Prvi cilj ovoga istraživanja bio je predmetno obraditi 108 suvremenih poslovice i njihovih varijantnih oblika prikupljenih u sklopu istraživanja Alekse Varge (2023), a koje se smatraju najčešće korištenim poslovicama prema *Hrvatskom nacionalnom korpusu*, *Hrvatskoj jezičnoj riznici* i korpusu *hrWaC* označavajući svaku pojedinu poslovicu ključnim riječima u denotativnom (doslovnom) i konotativnim (prenesenim) značenjima.

Drugi, krajnji cilj ovoga istraživanja proizlazi iz prvoga, a to je iz termina koji su dobiveni predmetnom obradom izraditi deskriptore te pomoću njih, prateći metodologiju izrade tezaurusa, izraditi dvostruki tezaurus suvremenih hrvatskih poslovice u denotativnom (doslovnom) i konotativnom (prenesenom) značenju. Važno je naglasiti da cilj tezaurusa poslovice nije potpuna terminološka preciznost kao pri izradi znanstvenog tezaurusa, nego grupiranje poslovice po dominantnim temama i značenjima, kako bi korisnik lako pronašao srodne poslovice. Pritom su važniji iskustveni, emocionalni i narativni odnosi među pojmovima nego stroga semantička hijerarhija. Za razliku od tehničkih ili znanstvenih tezaurusa koji počivaju na preciznim definicijama i formalnim klasifikacijama, tezaurus poslovice temelji se na interpretaciji kulturnog i jezičnog značenja, pri čemu naglasak nije na jednoznačnosti termina, već na njihovoj sposobnosti povezivanja poslovice u zajedničke tematske cjeline.

4.2 Metodologija predmetnog označivanja i izrade tezaurusa

S obzirom na dvojak cilj ovoga istraživanja, metodologija se također može razložiti na dva dijela.

4.2.1 Predmetno označivanje suvremenih poslovice: postupci i metode

Prvi korak predmetnog označivanja suvremenih poslovice bila je predmetna obrada poslovice u denotativnom (doslovnom) značenju, pri čemu se denotativna (doslovna) analiza odvijala tako da su se leksički elementi poslovice interpretirali, odnosno preuzimali *de visu*². Tako su, primjerice, za poslovicu *Tko rano rani, dvije sreće grabi* odabrani termini *ranjenje*, *sreća* i *grabljenje*, pri čemu se, za potrebe pretraživanja *Portala*, slijedilo pravilo prebacivanja svih imenica u nominativ jednine, svih pridjeva u imenice u nominativu jednine i svih glagola u glagolske imenice, također u nominativu jednine. Od ovog se pravila odstupa iznimno, u slučajevima kada bi tako tvorene riječi bile prosječnom korisniku nejasne ili kada bi tvoreni oblik graničio s negramatičnim, odnosno kada bi deskriptor bio nepretraživ i neupotrebljiv.

Drugi korak podrazumijevao je pregledavanje Šarić-Wittscheninog *Rječnika sinonima hrvatskoga jezika* (Šarić i Wittschen, 2008) kako bi se utvrdile istoznačnice i bliskoznačnice dobivenih termina, ali i mrežnih jezičnih baza ukoliko određena riječ nije jasna. U tu su svrhu

² Denotacija je „dogovoreno doslovno, svjesno i relativno stabilno značenje znaka” (Institut za hrvatski jezik i jezikoslovlje, 2025a). Kako denotativno značenje predstavlja osnovno, doslovno značenje riječi neovisno o kontekstu, odnosno ono što većina govornika pod tim izrazom podrazumijeva bez dodatnih tematskih/predmetnih asocijacija, kod denotativnog (doslovnog) značenja poslovice radi se samo doslovna analiza prisutnih ključnih riječi u samoj poslovisi (Olimovna i Ibodullayevna, 2025).

korišteni portali *Kontekst.io* (Kontekst.io, 2025), Hrvatski jezični portal (Hrvatski jezični portal, n. d.), korpusne baze (*Hrvatska jezična riznica*, *Hrvatski nacionalni korpus*, *Hrvatski mrežni korpus hrWaC* te korpusi *MaCoCu* i *ClassSla*) putem sustava *Sketch Engine* (Sketch Engine, 2025), portal *Struna* (Institut za hrvatski jezik i jezikoslovlje, 2025a) te Kolokacijska baza hrvatskoga jezika (Institut za hrvatski jezik i jezikoslovlje, 2025b). Primarni je izvor za određivanje sinonima i kvazisinonima bio Šarić-Wittschenin *Rječnik sinonima hrvatskoga jezika* s obzirom na to da je on prvi hrvatski leksikografski rječnik sinonima. On je trenutno i jedini hrvatski rječnik svoje vrste, upravo iz razloga što u hrvatskom jezikoslovlju nije postignuto suglasje oko toga što se sve podrazumijeva pod pojmom sinonima (Opašić, 2010). Pri odabiru sinonima i kvazisinonima iz *Rječnika* se uzimala u obzir samo prva kategorija sinonima, odnosno ona semantički najbliža terminu koji se obrađuje³. Iako su termini iz prve kategorije semantički bliski terminima koji se predmetno obrađuju, nisu svi sinonimi jednako prikladni u svim kontekstima poslovice. Iako sinonimija, kao što je već navedeno, može podrazumijevati zamjenjivost u svim kontekstima, u praksi je ta zamjenjivost često ograničena. Stoga su odabrani oni sinonimi koji se mogu najprirodnije uklopiti u značenje poslovice koja se predmetno obrađuje. Također, neki termini iz prve kategorije, iako bliski terminu koji se obrađuje, mogu imati dodatna značenja ili asocijacije koje bi mogle dovesti do nejasnoća⁴. Termini iz druge i treće, odnosno ostalih razina razmatrali su se kao potencijalni srodni termini. Uz to, portal *Kontekst.io* također je korišten kao izvor za određivanje srodnih termina jer pomoću statističke sličnosti u kontekstu predlaže riječi koje se često pojavljuju u sličnim kontekstima, no ne nužno sa sličnim značenjem.

Ovakvim uvidom u kontekstualnu sličnost termina, a ne njihovu značenjsku istovjetnost, pogodan je izvor za određivanje srodnih termina. U nastavku se donosi primjer traženja sinonima, kvazisinonima i srodnih termina za termin *sreća* u Šarić-Wittscheninom *Rječniku sinonima hrvatskoga jezika* i na portalu *Kontekst.io*. Tako ćemo, primjerice, pomoću *Rječnika sinonima* povezati termine *zadovoljstvo* i *ushit* kao sinonime i kvazisinonime te *blagostanje* kao srodan termin (Slika 1.), a pomoću portala *Kontekst.io* na termin *sreća* povezati termin *ugoda* kao srodan termin (Slika 2.).

³ U Rječniku pod Napomenama o strukturi rječnika i uputama za upotrebu pod odjeljkom Poredak sinonima u sinonimnim nizovima stoji kako „[s]inonimi u sinonimnim nizovima nisu poredani abecedno. U određivanju poretka nastojala se uzimati u obzir značenjska bliskost s natuknicom: riječi značenjski bliže natuknici načelno su smještene prije riječi značenjski udaljenijih od nje.“ (Šarić i Wittschen, 2008)

⁴ Primjerice, kod poslovice *Tko rano rani, dvije sreće grabi*, za termina *sreća* u denotativnom značenju, prema Šarić-Wittscheninom (2008) *Rječniku* određeni su sljedeći sinonimi prve kategorije: *zadovoljstvo*, *smirenost*, *duševni mir/sklad*, *sklad*, *blaženstvo*, *ushit* i *zanos*. Sinonimi iz prve kategorije poput *smirenost*, *duševni mir/sklad*, *sklad*, *blaženstvo* i *zanos* nose konotacije koje nisu u potpunosti kompatibilne s praktičnim (svakodnevnim) značenjem termina *sreća* u ovoj poslovice. Primjerice, *smirenost* i *duševni mir* više upućuju na pasivno, statično stanje, dok ova poslovice govori o aktivnom postizanju nagrade. *Blaženstvo* i *zanos* često imaju religiozne i filozofske konotacije. Stoga su kao najprirodniji sinonimni/kvazisinonimni termini za termin *sreća* u ovom kontekstu odabrani *zadovoljstvo* i *ushit*.

sreća* f 1 (*osjećaj sreće*) zadovoljstvo, smirenost, duševni mir/sklad, sklad, blaženstvo, ushit, zanos **usp. 2** (*život u sreći*) blagostanje *usp.*, dobrobit, blagoslov **3** (*~ ih prati u poslu*) uspjeh, pobjeda, sretan slučaj, sretne/povoljne okolnosti, sretan splet okolnosti, hazard, (*~ im u životu nije bila uvijek sklona*) sudbina *usp.*, fortuna; nef. orij. talih • nef. narok, berićet

Slika 1: Potencijalni sinonimi, kvazisinonimi i srodni termini za termin „sreća“ u Šarić-Wittscheninom Rječniku sinonima hrvatskoga jezika (izvor: Šarić i Wittschen, 2008)

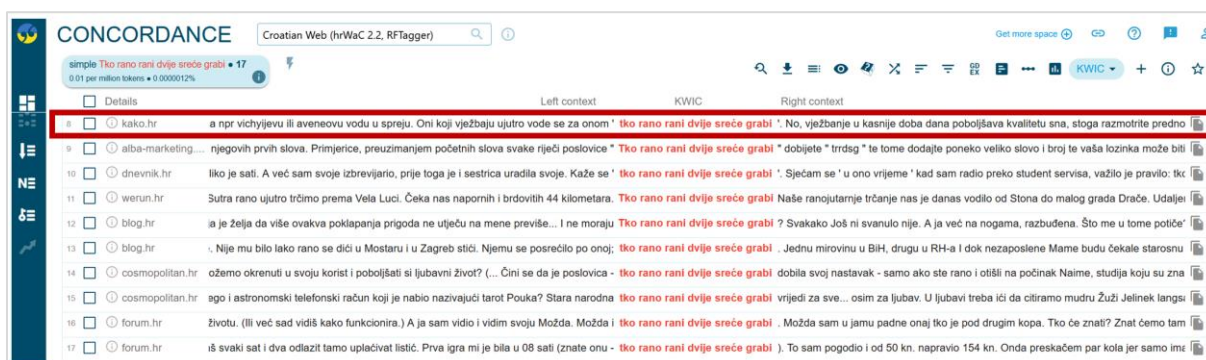
Slični izrazi i sinonimi za sreća			Primjeri iz općenitog korpusa Korpus hrWac
Kliknite za traženje	UČESTALOST	SLIČNOST	hrWac je korpus hrvatskog jezika (Clarín.si)
sreća	3.07	69%	novčanih sredstava. sreće, te na odgajanje i usmjeravanje primjereno svojim razvojnim sposobnostima.  JOSEPH TAFUR KLIC REKE Avtor raziskuje svet amazonskega šamanizma in opisuje, kako duh zdravi telo, kako ljubezen obnavlja dušo in kako iskanje višjega smisla in življenjskega poslanstva odpira vrata zdravju in dolgemu življenju. www.eno.si oglas , meni su iznenađenja romantika. Iako plaćem na sve romantične filmove, nisam romantična kao moj dragi. Imala sam sreće, našla sam dobrog
sreča	0.39	67%	
tuga	8.9	65%	
patnja	6.03	63%	
utjeha	5.38	63%	
dobrota	3.6	57%	
samoća	1.96	57%	
nevolja	5.4	57%	
znatiželja	2.17	56%	
ugoda	0.89	56%	
ljubav	175.2	56%	

Slika 2: Potencijalni sinonimi, kvazisinonimi i srodni termini za termin „sreća“ na portalu Kontekst.io (izvor: Kontekst.io, 2025)

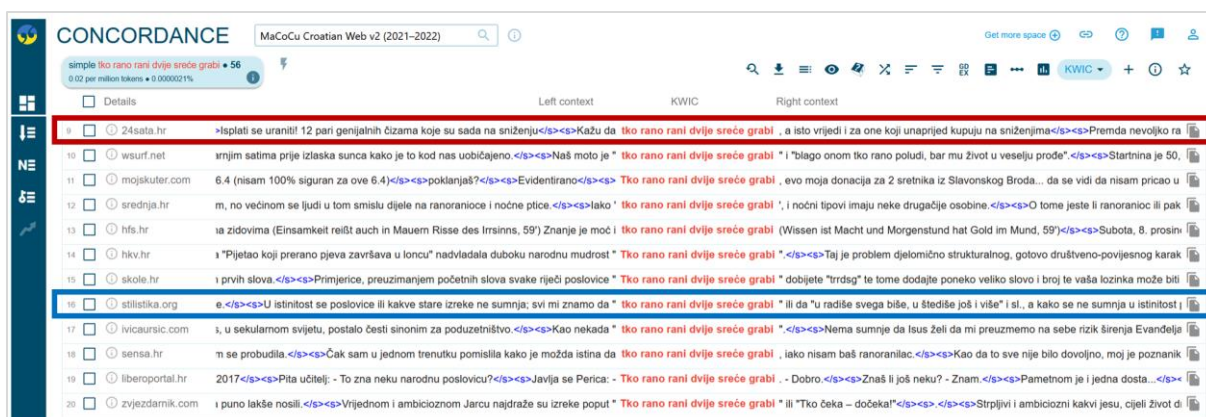
Treći korak podrazumijevao je završnu predmetnu obradu za denotativno (doslovno) značenje, pri čemu su se popisivali svi odabrani termini iz poslovice te svi sinonimi, kvazisinonimi i srodni termini koji su pronađeni u navedenim rječnicima, portalima, bazama i korpusima. Primjerice, za termin *sreća* popis bi izgledao kako slijedi: *sreća* (sinonimi/kvazisinonimi: *zadovoljstvo*, *ushit*; srodni termini: *blagostanje*, *ugoda*).

U četvrtom se koraku kreće s konotativnom analizom poslovice koja se temelji na Lancasterovoj metodi predmetnog označivanja, a podrazumijeva dva koraka: intelektualnu ili konceptualnu analizu, prilikom koje se određuje predmet (očemnost) poslovice i prijevod

rezultata dobivenih konceptualnom analizom u konkretne termine za predmetno označivanje (Lancaster, 1998). U ovom se koraku najprije konzultiraju korpusne baze, gdje se iz konteksta u kojima se određena poslovice koristi zaključuje koje je njezino značenje, odnosno očemnost. U nastavku se donose rezultati korpusne analize na primjeru korpusnih baza *hrWaC* i *MaCoCu* za poslovice *Tko rano rani, dvije sreće grabi*. Tako, primjerice, vidimo da se prema korpusu *hrWaC* na portalu *kako.hr* (8. rezultat) ova poslovice koristi u kontekstu vježbanja, gdje se ističe da oni koji rano ustaju i vježbaju ujutro ostvaruju dvostruku korist – ne samo da su fizički aktivni, već i dan započinju energično i s osjećajem postignuća (Slika 3.). U korpusu *MaCoCu* vidimo kako se ova poslovice koristi na portalu *24sata.hr* u kontekstu kupovine obuće na sniženju (9. rezultat). Poslovice ovdje prenosi poruku kako oni koji djeluju pravovremeno, odnosno koji prvi reagiraju na sniženja i kupuju unaprijed, uspijevaju kupiti željene proizvode po povoljnijoj cijeni. Također, u istom korpusu vidimo da se ova poslovice koristi na portalu *stilistika.org* (16. rezultat), gdje se dodatno uspoređuje s poslovicom *U radiše svega biše, u štediše još i više.*, koja dodatno ističe vrijednost marljivosti i upornog rada kao put do većeg uspjeha i blagostanja. Obje poslovice zajedno šalju poruku kako proaktivnost, trud i disciplina donose višestruke prednosti u svakodnevnom životu (Slika 4.). Tako smo za navedenu poslovice odredili da joj je konotativno (preneseno) značenje vezano uz *važnost ranog ustajanja, odnosno započinjanja nekog posla što je ranije moguće kako bi se postigli optimalni rezultati*.



Slika 3: Korpusna analiza poslovice „Tko rano rani, dvije sreće grabi“ u korpusu *hrWaC* (izvor: Sketch Engine, 2025)



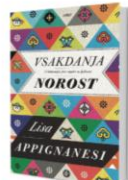
Slika 4: Korpusna analiza poslovice „Tko rano rani, dvije sreće grabi“ u korpusu *MaCoCu* (izvor: Sketch Engine, 2025)

Nakon što se iz korpusnih baza preko konteksta zaključilo koja bi mogla biti očemnost poslovice, kreće se na peti korak, odnosno drugi korak predmetnog označivanja – dodavanje termina za predmetno označivanje u konotativnom (prenesenom) značenju. U ovom bi slučaju za navedenu poslovicu *Tko rano rani, dvije sreće grabi* to bili termini *ustrajnost* i *revnost*.

Šesti korak podrazumijeva pretraživanje odabranih termina u spomenutom *Rječniku sinonima* i mrežnim jezičnim bazama kako bi se dobilo što više potencijalnih sinonima, kvazisinonima i srodnih termina po kojima bi korisnik mogao tražiti određenu poslovicu. Ovdje se vodi jamstvom uporabe i obogaćivanja ulaznog korisničkog rječnika sa svrhom olakšavanja pretraživanja korisnicima prema nedeskriptorima. U nastavku se donosi primjer traženja sinonima, kvazisinonima i srodnih termina za termin *ustrajnost* u Šarić-Wittscheninom *Rječniku sinonima hrvatskoga jezika* i na portalu *Kontekst.io*. Tako ćemo, primjerice, na termin *ustrajnost* pomoću *Rječnika sinonima* povezati termine *izdržljivost* i *postojanost* kao sinonime i kvazisinonime⁵ (Slika 5.), a pomoću portala *Kontekst.io* povezat ćemo srodne termine *strpljivost* i *marljivost*, koji su ujedno jedni od posljednjih termina (odnosno semantički najudaljenijih od termina *ustrajnost*) u *Rječniku sinonima* (Slika 6.). Dakle, za poslovicu *Tko rano rani, dvije sreće grabi* za termine za predmetno označivanje na temelju Rječnika sinonima i portala *Kontekst.io* odabrani su sljedeći termini: *ustrajnost* (sinonimi/kvazisinonimi: *izdržljivost*, *postojanost*; srodni termini: *strpljivost*, *marljivost* i *revnost*.

ustrajnost f izdržljivost, postojanost, stalnost, upornost usp., nepokolebljivost, nepopustljivost, strpljivost, strpljenje; marljivost, revnost; okorjelost; nef. istrajnost

Slika 5: Potencijalni sinonimi, kvazisinonimi i srodni termini za termin „ustrajnost“ u Šarić-Wittscheninom *Rječniku sinonima hrvatskoga jezika* (izvor: Šarić i Wittschen, 2008)

Slični izrazi i sinonimi za ustrajnost			Primjeri iz općenitog korpusa
Kliknite za traženje			Korpus hrWac
UČESTALOST SLIČNOST			hrWac je korpus hrvatskog jezika (CLarin.si)
strpljivost	1.92	82%	s tobom, odgovoriše mu drugi učenici. Kako je to lijepo čuti! Ima još nešto što je Isus nagradio: njihovu marljivost i ustrajnost. Da su čekali prekrštenih ruku što će se dogoditi, njihove bi mreže bile još i danas prazne. Da su lovili sat ili dva, a  LISA APPIGNANESI VSAKDANJA NOROST Iskrena, autobiografska priprema o življenju, smrti, norosti, ljubavi i vnukih, kakor jih vidi pisateljica, ki se ves čas zaveda zgodovinskih in sodobnih muhavosti ženske usode. oglas
upornost	5.07	80%	
odvažnost	0.9	76%	
predanost	4.24	75%	
istrajnost	0.1	74%	
nepokolebljivost	0.11	74%	
marljivost	0.9	73%	
odlučnost	4.9	73%	
dosljednost	2.27	73%	
požrtvovnost	1.24	72%	
razboritost	0.75	71%	

Slika 6: Potencijalni sinonimi, kvazisinonimi i srodni termini za termin „ustrajnost“ na portalu *Kontekst.io* (izvor: Kontekst.io, 2025)

⁵ Pojedine rječničke natuknice nemaju kategorije sinonima, već se samo nudi nenumerirani popis. U takvim su se slučajevima kao sinonimi i kvazisinonimi preferirali termini koji su najizravnije semantički povezani s terminom koji se obrađuje, tako da se slijedio prethodno obrazloženi *Poredak sinonima u sinonimnim nizovima*.

Sedmi i završni korak podrazumijeva korištenje alata umjetne inteligencije, odnosno velikih jezičnih modela kao pomoć pri dodatnoj provjeri konotativnog (prenesenog) značenja poslovice, pri čemu su odabrani modeli GPT-4.5, Sonar i Claude 3.7 (koji su dio Perplexityja)⁶. U nastavku slijedi primjer odgovora Perplexityja na upit *Molim te, reci mi koje je konotativno, odnosno preneseno značenje poslovice Tko rano rani, dvije sreće grabi te navedi sve izvore koje si koristio za pisanje odgovora*. Nakon što Perplexity ponudi odgovor, na temelju pročitaneog teksta potvrđuje se kako se konotativno (preneseno) značenje ove poslovice odnosi na vrijednost marljivosti, discipline i proaktivnosti, naglašavajući kako rano ustajanje i pravovremeno djelovanje povećavaju izgleda za uspjeh, sreću i ostvarenje životnih ciljeva (Slika 7.). Nakon potvrde konotativnog (prenesenog) značenja poslovice, tekst odgovora Perplexityja predmetno se obrađuje tako da se izdvajaju termini u konotativnom (prenesenom) značenju. Potom se do tada izdvojenim terminima dodaju novi sinonimi/kvazisinonimi/srodni termini, ukoliko bi se isti pojavili tijekom predmetne obrade odgovora Perplexityja. Tako ćemo, primjerice, na osnovu dolje navedenog odgovora (Slika 7.) za ovu poslovicu terminima *ustrajnost* (sinonimi/kvazisinonimi: *izdržljivost, postojanost*; srodni termini: *strpljivost, marljivost*) i *revnost* dodati i srodne termine *redovitost, proaktivnost* i *disciplina*. Tako ćemo ukupno za poslovicu *Tko rano rani, dvije sreće grabi* imati 9 termina: *ustrajnost* (sinonimi/kvazisinonimi: *izdržljivost, postojanost*; srodni termini: *disciplina, redovitost, strpljivost, marljivost*) i *revnost* (srodni termin: *proaktivnost*).

⁶ Navedeni su modeli dio *Perplexityja*, alata umjetne inteligencije čija je komercijalna inačica besplatno dostupna studentima za korištenje u osobne i istraživačke svrhe (Perplexity, 2025). Standardizirani *prompt*, *Molim te, reci mi koje je konotativno, odnosno preneseno značenje poslovice [poslovica] te navedi sve izvore koje si koristio za pisanje odgovora*, oblikovan je prema načelima koja su postavili Chen, Zhang, Langrené i Zhu (Chen i sur., 2025) te su u svakom *promptu* izričito traženi izvori iz kojih proizlaze odgovori, a isti su dodatno provjereni.

Molim te, reci mi koje je konotativno, odnosno preneseno značenje poslovice „Tko rano rani, dvije sreće grabi“ te navedi sve izvore koje si koristio za pisanje odgovora

Answer Sources · 10 Tasks

Jutarnji list

KONAČNO OTKRIVENO JE LI IZREKA 'TKO RANO RANI, DVIJE SREĆE GRABI' ISTINITA

Studentski.hr

Deset staromodnih, ali zlata vrijednih savjeta -...

Rečnik manje poznatih ...

Objašnjenje poslovice: Ko rano rani, dve sreće grabi? ...

Zagreb.info

Pogled iz drugog kuta: Narodne izreke, kada ih...

Konotativno, odnosno preneseno značenje poslovice „Tko rano rani, dvije sreće grabi“ jeste da onaj tko je marljiv, discipliniran i redovit, odnosno tko se rano budi i započne raditi, ima veće šanse za uspjehom i boljim rezultatima u životu. Poslovica naglašava vrijednost proaktivnosti, organiziranosti, dosljednosti i pozitivne rutine, sugerirajući da rano ustajanje i početak rada daju više vremena za ostvarenje ciljeva, donose bolje raspoloženje i veću vjerojatnost za uspjeh ^{2 3 5}.

Ova izreka potiče iz narodne mudrosti koja je nastala u ruralnim zajednicama, gdje je rano ustajanje bilo ključno za uspjeh u poljoprivredi i drugim svakodnevnim obvezama, jer je rad od zore do sumraka bio uobičajen. Kroz povijest, poslovica se proširila i na druge aspekte života, podržavajući ideju da redovitost i marljivost vode do više mogućnosti i „sreća“ – odnosno uspjeha ^{3 2 5}.

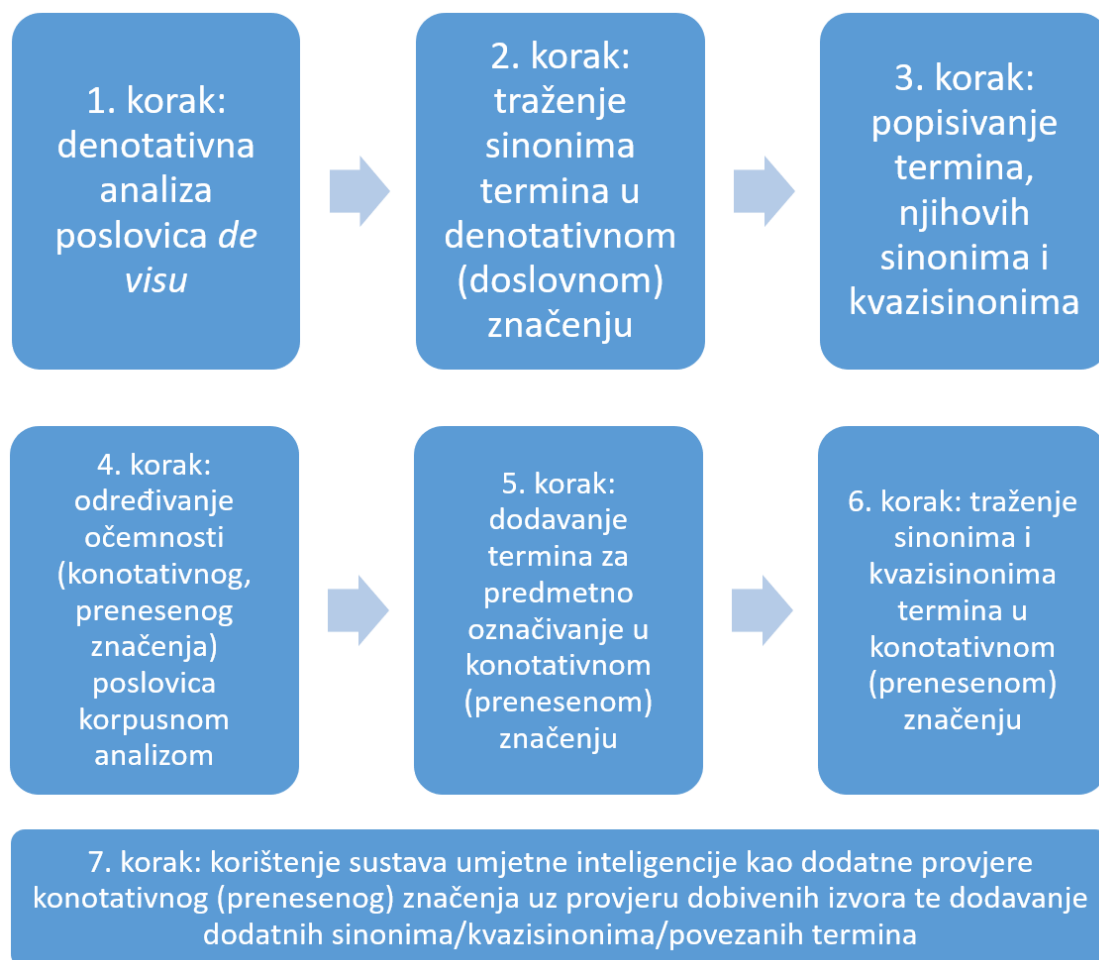
Osim prenesenog značenja, poslovica ponekad se koristi i humoristično ili s ironijom, kako bi se naglasilo da rano ustajanje ne mora nužno donijeti dvostruku sreću, već da je važno prilagoditi se vlastitim potrebama i ritmu ^{4 6}.

Izvori koje sam koristio za pisanje odgovora:

1. **Jutarnji.hr (2017):** KONAČNO OTKRIVENO JE LI IZREKA 'TKO RANO RANI, DVIJE SREĆE GRABI' ISTINITA Jutarnji ili večernji tipovi: Evo tko je depresivniji, a tko uspješniji ¹.
2. **Studentski.hr (2018):** Deset staromodnih, ali zlata vrijednih savjeta ².
3. **Rečnikopedija (2024):** Objašnjenje poslovice: Ko rano rani, dve sreće grabi? – Poreklo, značenje, upotreba i primeri ³.
4. **Zagreb.info (2024):** Pogled iz drugog kuta: Narodne izreke, kada ih 'analizirate' na ovaj način nose još dublji besmisao ⁴.
5. **NEON Televizija (2012):** Ko rano rani, dvije sreće grabi ⁵.
6. **Tportal (2013):** Tko rano rani, dvije sreće ipak ne grabi! ⁶

Slika 7: Primjer konotativne analize poslovice „Tko rano rani, dvije sreće grabi“ u Perplexityju (izvor: Perplexity, 2025)

Grafički bi koraci predmetne obrade suvremenih poslovice izgledali kako slijedi (Slika 8.):



Slika 8: Grafički prikaz toka predmetne obrade suvremenih poslovice (izvor: autor)

Provedbom navedenih koraka kod svih 108 poslovice i njihovih varijantnih oblika dobiven je popis termina (zajedno s njihovim sinonimnim, kvazisinonimnim i srodnim oblicima) za svaku poslovicu u denotativnom i konotativnom značenju.

4.2.2 Izrada dvostrukog tezaurusa suvremenih poslovice: postupci i metode

Nakon stvaranja popisa termina za denotativno i konotativno značenje svake poslovice, pristupilo se izradi dvostrukog tezaurusa suvremenih poslovice. Prvi tezaurus sadržava deskriptore u denotativnom značenju, a drugi sadržava deskriptore u konotativnom značenju.

Izrada tezaurusa uvelike se temeljila na Shearerovoj metodologiji izrade tezaurusa (Shearer, 2004), dok su se prilikom izrade poštivala sva tri jamstva vezana uz domenu rječnika: jamstvo predložka, strukturno jamstvo i korisničko jamstvo. Poštivanje jamstva predložka očituje se kroz odabir termina za predmetno označivanje, a posebice deskriptora u denotativnom (doslovnom) značenju, iz dokumenata koji se obrađuju, u ovom slučaju iz poslovice. Tako se, primjerice, u tezaurusu mogu pronaći termini *granuće*, *grehota* i *povaljivanje* koji se, iako nisu dijelom svakodnevnog hrvatskog leksika, nalaze kao termini u tezaurusu jer su to riječi koje se koriste u poslovice. Nadalje, poštivanje jamstva uporabe očituje se kroz korištenje sinonima, kvazisinonima i srodnih termina koji se nalaze na mrežnom portalu *Kontekst.io*, gdje

je vidljiva i sličnost njihova korištenja u kontekstu i učestalost, čime se osigurava da će korisnici moći pretraživati prema terminima koji su u skladu sa svakodnevnim rječnikom. Tako je, primjerice deskriptoru *odsutnost* dodan sinonim *odsustvo*, koji uz sličnost od 72 % ima indeks učestalosti od 2,25, a terminu *iskorištavanje* dodan je sinonim *korištenje*, koji uz sličnost od 69 % ima indeks učestalosti od čak 76,1. Naposljetku, jamstvo strukture očituje se kroz uključivanje termina koji nužno ne poštuju jamstva predloška ili uporabe, no imaju svrhu popunjavanja praznina u tezaursu ili okupljanja specifičnih srodnih termina. Tako su, primjerice, oblikovane glavne fasete tezaurusa, pri čemu su nakon okupljanja deskriptora stvoreni novi termini koji su označavali čitavu fasetu, primjerice, faseta *Promjena stanja i kretanje* ili faseta *Sukob, borba i nametanje* i slično.

Metodologija korištena pri izradi oba tezaurusa podrazumijevala je:

1. prikupljanje prirodnih riječi koje označavaju denotativno i konotativno značenje poslovice
2. spajanje sinonima, kvazisinonima i srodnih termina
3. stvaranje fasete okupljanjem deskriptora
4. slaganje fasete i podfasete
5. dodavanje novih deskriptora radi popunjavanja značajnijih praznina
6. notacija
7. identificiranje veza između deskriptora među fasetama
8. stvaranje abecednog kazala i
9. stvaranje abecednog prikaza (Shearer, 2004).

Prvi korak, prikupljanje prirodnih riječi koje označavaju denotativno i konotativno značenje poslovice već je obavljen predmetnim označivanjem poslovice. U drugom su koraku svi sinonimi i kvazisinonimi odabranih termina povezani tako da se odredio jedan deskriptor, a svi njegovi drugi oblici postali su nepreferirani termini (*nedeskriptori*). S obzirom da je u literaturi jedan od koraka i razrješavanje homografije, a da na našem uzorku nisu zamijećeni slučajevi homografije, istu nije bilo potrebno razrješavati. U trećem koraku okupljeni su deskriptori koji označavaju pojmove iste vrste (npr. *grana*, *grm*, *stablo*, *lan*, *jabuka* i sl.) tako da se skupovima od više takvih pojmova dodjeljuju nazivi koji čine grupu, odnosno fasetu (npr. *biljni svijet*). Za potrebe izrade ovog tezaurusa, izrađivala se samo jedna razina podfasete. U četvrtom su se koraku izrađene fasete i podfasete složile kako bi tezaursi dobili strukturu (Slika 9.):

BILJNI SVIJET (sinonim: flora)
jabuka
grana (sinonim: grančica)
žito (sinonim: klas)
lan
jagoda
trešnja
korov (sinonim: pirika)
drvo (sinonim: stablo, srodni termin: sjekira, nošenje)
grozd (sinonim: grožđe)
pupoljak (sinonim: izdanak)
grm (sinonim: žbun)
kopriva
vrba
zrno (sinonimi: kučina, kukolj, ljulj)

Slika 9: Primjer izrađenih fasete i podfasete u tezaursu (izvor: autor)

U petom koraku su se, po uočavanju praznina, a u skladu sa strukturnim jamstvom, u dobivenom popisu deskriptora stvarali novi, pri čemu se držalo pravila kako oni trebaju biti jednoznačni te se izbjegavalo korištenje kvazisinonima ili polisema kao deskriptora. Tako je u sljedećem primjeru stvoren novi deskriptor, nadređeni termin *ekonomija*, kako bi se u jednu fasetu mogli okupiti deskriptori *bogatstvo*, *boljitak*, *dobit*, *financije*, *korist*, *održivost* i *profit* (Slika 10.):

EKONOMIJA dobit (sinonim: zarada) bogatstvo (sinonim: imetak) korist (srodni termin: iskorištavanje) boljitak održivost financije (sinonim: sredstva) profit

Slika 10: Primjer izrade novog deskriptora *ekonomija* u službi nadređenog termina (izvor: autor)

Šesti korak podrazumijevao je olakšavanje pronalaska deskriptora korisnicima preinakom dobivenog poretka deskriptora u standardnu notaciju (širi termin, engl. *Broader Term*, BT; uži termin, engl. *Narrower Term*, NT; srodni termin engl. *Relater Term*, RT; USE (eng. *use*) – uporabi: izabrani deskriptor za sinonim ili kvazisinonim; uputa za sinonim-sinonim; UF (eng. *use for*) – uporabi za: sinonim, kvazisinonim; uputa za deskriptor). Primjerice, za prethodno navedeni primjer preinaka poretka izgledala bi kako slijedi (Slika 11.):

EKONOMIJA NT bogatstvo UF imetak NT boljitak NT dobit UF zarada NT financije UF sredstva NT korist RT iskorištavanje NT održivost NT profit

Slika 11: Preinačeni poredak deskriptora iz prethodnog primjera u standardnu notaciju (izvor: autor)

Sedmi korak podrazumijevao je identificiranje veza između deskriptora među fasetama i izrađivanje unakrsnih uputnica između srodnih deskriptora u različitim fasetama. U nastavku slijedi primjer takvih uputnica za termine *kruh* (faseta *hrana i prehrana*) i *želudac* (faseta *ljudsko tijelo i zdravlje*) u tezaurusu termina u denotativnom značenju (Slika 12.):

HRANA I PREHRANA	LJUDSKO TIJELO I ZDRAVLJE
NT glad UF gladovanje	NT glava
NT jedenje UF hranjenje UF prehranjivanje	RT pamet
NT juha UF čorba UF supa	NT grlo
RT lonac	NT krv
NT kruh UF hleb	RT srce
RT želudac	NT noga
NT luk	RT hodanje
NT med	NT peta UF taban
NT pečenje	NT ruka
NT pogača UF pogačica	NT tijelo
NT sitost UF zasićenost	NT usta
	NT zdravlje
	RT liječenje
	NT želudac UF stomak UF trbuh
	RT kruh

Slika 12: Primjer unakrsnih uputnica između termina kruh i želudac (izvor: autor)

Osmi korak podrazumijevao je stvaranje abecednog kazala kojim će korisnik moći olakšano pretraživati termine iz tezaurusa u abecednom poretku koristeći prethodno spomenute USE uputnice, primjerice, ako korisnik traži termin *crnina*, kraj istoga će stajati uputnica uporabi (USE), odnosno da se u tezaurusu kao preferirani, odobreni termin koristi termin *crno*; ili ako korisnik traži termin *cura*, kraj istoga će također stajati uputnica uporabi (USE), odnosno da se u tezaurusu kao preferirani, odobreni termin koristi termin *djevojka*. Primjerice, za slovo C u abecednom kazalu tezaurusa termina u denotativnom značenju, navedeno bi izgledalo kako slijedi (Slika 13.):

C
carevanje
cigla
cipela
crnina USE crno
crno
cura USE djevojka

Slika 13: Primjer slova C u abecednom kazalu (izvor: autor)

Deveti i posljednji korak podrazumijevao je stvaranje abecednog prikaza samog tezaurusa. Primjerice, za fasetu *ekonomija* u tezaurusu termina u konotativnom značenju to bi izgledalo kako slijedi (Slika 14.):

EKONOMIJA
NT bogatstvo UF imetak
NT boljitak
NT dobit UF zarada
NT financije UF sredstva
NT korist
RT iskorištavanje
NT održivost
NT profit

Slika 14: Primjer fasete ekonomija u abecednom prikazu (izvor: autor)

Predmetna obrada 108 suvremenih poslovice i izrada tezaurusa odvijala se od siječnja do svibnja 2025. godine

4.3 Uzorak

Postupak odabira poslovice uključenih u ovo istraživanje započeo je projektom određivanja hrvatskog paremiološkog minimuma i optimuma⁷ (2014. – 2018.), u sklopu kojega su provedena dva istraživanja 2014. i 2018. godine sa standardiziranim upitnikom poznavanja hrvatskih poslovice. S obzirom na to da je, kao što je već navedeno, većina hrvatskih poslovice dokumentirana u zbirkama iz 19. i 20. stoljeća (odnosno od 1846. do 1987. godine, a za razliku od nekih drugih jezika u kojima postoje već paremiološke zbirke ili rječnici poslovice), u istraživanju se nije moglo osloniti na metodološki opravdan izvor u vidu rječnika poslovice, već na vlastiti korpus autorice Alekse Varge koji je nastao prikupljanjem poslovice u razdoblju trajanja projekta i potvrđivanjem frekvencije njihovog korištenja u *Hrvatskom nacionalnom korpusu*, *Hrvatskoj jezičnoj riznici* i *hrWaC* korpusu. Tako je najprije sastavljen popis od 239 poslovice, koji je sužen na 207 oblika koji se pojavljuju u navedenim korpusima. Kako u trenutku istraživanja nije bio određen hrvatski paremiološki minimum i optimum, odnosno frekvencije korištenja i poznavanja poslovice među općom populacijom, pri odabiru poslovice oslanjalo se na prethodno spomenutu korpusnu analizu i frekvencije pojavljivanja poslovice u hrvatskim korpusima i mrežnim stranicama. Podvrgavanjem ovog popisa pilot-istraživanju na 21 lingvistu, popis je smanjen na 105⁸ najčešćih hrvatskih poslovice poredanih po frekvenciji pojavljivanja, pri čemu su poslovice podijeljene u tri kategorije zastupljenosti: od najzastupljenijih (1-35), preko manje zastupljenih (36-70) do najmanje zastupljenih (71-105) (Aleksa Varga, 2023).

5 Rezultati⁹

5.1 Suvremene poslovice: rezultati predmetne obrade

Predmetnom obradom navedenim poslovicama dodijeljeno je ukupno 1723 termina, od toga 859 u denotativnom (doslovnom) značenju, od kojih su 188 sinonimi i kvazisinonimi, a 864 u konotativnom (prenesenom) značenju, od kojih su 231 sinonimi i kvazisinonim (za popis termina po poslovi i način njihove statističke obrade vidi Prilog 1 u Repozitoriju). Iz Tablice 1 vidljivo je kako su najčešće u denotativnom (doslovnom) značenju pojavljuju termini *bivanje* i *čovjek* (pojavljuju se 8 puta) te *kuća* i *dva* (pojavljuju se 6 puta), odnosno radi se o vrlo niskoj

⁷ Paremiološki je minimum skup najvažnijih, najčešće korištenih i najpoznatijih poslovice određenog jezika, a koje su potrebne za osnovno razumijevanje i sudjelovanje u usmenim i pisanim komunikacijskim situacijama. U praksi, to je popis poslovice koje bi svaki korisnik jezika trebao poznavati kako bi mogao učinkovito komunicirati i razumjeti kulturološke reference u korištenom jeziku. Paremiološki je optimum širi pojam i predstavlja skup različitih paremioloških minimuma, odnosno najpoznatijih i najkorištenijih poslovice potrebnih za poučavanje jezika kao materinjeg ili stranog te za kulturološku pismenost prosječno obrazovane osobe (Aleksa Varga i Hrisztova-Gotthard, 2019).

⁸ Na popisu korištenom za potrebe ovog diplomskog rada navedeno je 108 poslovičkih oblika s pripadajućim varijantnim oblicima. Autorica istraživanja i voditeljica projekta navodi kako se na proširenom popisu, uz poslovice, nalaze također dvije izreke i jedna ponovljena poslovice.

⁹ U rezultatima će se deskriptivno i na odabranim primjerima prikazati termini korišteni za predmetnu obradu suvremenih poslovice i fasete tezaurusa. Tablica s cjelovitim popisom dodijeljenih termina i klasifikacijskih oznaka po poslovi, oba tezaurusa i pripadajućim abecednim kazalima dostupni su na institucijskom Repozitoriju Filozofskog fakulteta Sveučilišta Josipa Jurja Strossmayera u Osijeku (DABAR) na sljedećoj poveznici: <https://urn.nsk.hr/urn:nbn:hr:142:488805>.

frekvenciji riječi. Ove su riječi uglavnom konkretne i vezane uz opipljive pojmove (*novčić, ovan, škuda, hleb, kozlič* i slično). Iz Tablice 2 vidljivo je kako termini u konotativnom (prenesenom) značenju pokazuju višu maksimalnu frekvenciju (termin *ishod* se pojavljuje 12 puta; *oprez, pobjeda, postignuće* i *postizanje* pojavljuju se 9 puta), ali i znatno veći broj pojmova s nižim frekvencijama ($n = 4$; $n = 3$) u odnosu na denotativne termine. Ove su riječi apstraktnije te pokrivaju širi raspon značenja (*efikasnost, poticaj, pouzdanje, pouzdanost, rezultat* i slično).

Tablica 1: Frekvencije dodijeljenih ključnih riječi u denotativnom (doslovnom) značenju (izvor: autor)¹⁰

Termin (denotativno značenje)	frekvencija (n)
bivanje, čovjek	8
kuća, dva	6
dan, dobronamjernost, dobrota, dobroćudnost, imanje	5
dolazak, duša, miš, novčić, ovan, pad, rađanje, stizanje, vođenje, škuda	4
djelovanje, dom, duljina, dužina, gladovanje, govorenje, hleb, hranjenje, htjenje, jarić, kamen, kozlič, kraj, kruh, kučina, mišljenje, muka, novac, nošenje, ovca, padanje, pamet, para, patnja, pravljenje, rad, radost, rođenje, ruka, san, tri, voda, zlo, ćud,	3

Tablica 2: Frekvencije dodijeljenih termina u konotativnom (prenesenom) značenju (izvor: autor)¹¹

Termin (konotativno značenje)	frekvencija (n)
ishod	12
oprez, pobjeda, postignuće, postizanje	9
uspjeh	8
korist, ponašanje, sigurnost	7
efikasnost, poticaj, pouzdanje, pouzdanost, rezultat, strpljenje	6
inicijativa, ljubav, odgovornost, poduzimanje, popustljivost, rad, savjesnost, snošljivost, svojta, tolerancija, uzvraćanje	5
akcija, aspiracija, cilj, djelovanje, dobrobit, izbjegavanje, izmjena, karakter, karakteristike, konflikt, konkurencija, mir, namjera, neprijateljstvo, održivost, prevencija, promjena, reciprocitet, roditelj, samodostatnost, svađa, svrha, utjecaj, vrijeme	4
domaćinstvo, harmonija, iskustvo, izgled, jednakost, kauzalnost, mudrost, napor, nepouzdanost, nepromišljenost, nerealnost, obitelj, odluka, opasnost, opstojnost, optimizam, osobitost, podrška, pokvarenost, posljedica, potpora, potvrda, pragmatičnost, praksa, problem, prolaznost, promišljenost, razboritost, realizacija, rizik, rutina, samostalnost, sličnost, sposobnost, sreća, sukob, suradnja, uzročnost, vanjština, vrijednost	3

5.2 Suvremene poslovice: tezaursi u denotativnom (doslovnom) i konotativnom (prenesenom) značenju

Što se tiče rezultata vezanih za drugi dio istraživanja, izrađeni tezaurus suvremenih poslovice u denotativnom (doslovnom) značenju sadrži 29 glavnih faseta i 380 podfaseta. Ukupno je evidentirao 409 deskriptora i 177 sinonima i kvazisinonima koji su označeni odnosom

¹⁰ U tablici su prikazane sve frekvencije veće od 2.

¹¹ Idem.

ekvivalencije i uputnicom UF s nedeskriptora na deskriptor. Budući da je tezaurus izrađen s jednom razinom faseta, 29 glavnih faseta predstavljaju i 29 nadređenih termina (BT¹²), dok je podređenih termina (NT) 380. Srodnih termina ima ukupno 31 par, odnosno 62 RT oznake.

U nastavku se donosi primjer nekoliko faseta (faseta *osjećaji* i faseta *hrana i prehrana*) tezaurusa u denotativnom značenju, dok se cijeli tezaurus denotativnih značenja nalazi u Prilogu 2.1. na Repozitoriju (Slika 15.):

OSJEĆAJI UF emocije	HRANA I PREHRANA
NT drago UF milo	NT glad UF gladovanje
NT ljubav UF voljenje	NT jedenje UF hranjenje UF prehranjivanje
RT Bog	NT juha UF čorba UF supa
RT strast	RT lonac
NT milost UF milosrđe	NT kruh UF hleb
NT muka UF patnja	RT želudac
NT samoća UF osama UF osamljenost UF usamljenost	NT luk
NT sramota UF sram	NT med
NT sreća UF ushit UF zadovoljstvo	NT pečenje
RT blagostanje	NT pogača UF pogačica
RT ugoda	NT sitost UF zasićenost
NT strah	
RT prijetnja	
NT veselje UF radovanje	

Slika 15: Fasete *osjećaji* i *hrana i prehrana* tezaurusa u denotativnom značenju (izvor: autor)

Izrađeni tezaurus suvremenih poslovice u konotativnom (prenesenom) značenju sadrži 23 glavne fasete i 244 podfasete, s ukupno 267 deskriptora i 216 sinonima i kvazisinonima (uputnica UF s nedeskriptora na deskriptor). Također je izrađena samo jedna razina faseta, stoga 23 glavne fasete predstavljaju 23 nadređena termina (BT), dok podređenih termina (NT) ima 244. Srodnih termina ima 27 parova, odnosno 54 RT oznake.

U nastavku se donosi primjer nekoliko faseta (faseta *rizik i posljedice*, faseta *komunikacija* i faseta *mudrost*) tezaurusa u konotativnom značenju, dok se cjeloviti tezaurus konotativnih značenja nalazi u Prilogu 3.1. na Repozitoriju (Slika 16.):

RIZIK I POSLJEDICE	KOMUNIKACIJA
NT izazov UF prepreka	NT govor
RT hrabrost	NT informacija
NT iznenađenje UF iznenadnost UF šok	NT odgovaranje UF odgovor
NT katastrofa UF havarija UF tragedija	NT pomoć UF asistencija UF podrška
RT nesreća	NT vjerodostojnost UF autentičnost
NT nepredvidljivost UF neizvjesnost	UF pouzdanost
NT posljedica UF ishod UF rezultat	
RT uzročnost	
NT prijetnja	
RT nesigurnost	MUDROST
RT strah	NT iskustvo UF praksa
NT sklonište UF zaklon	NT odgoj
RT zaštita	NT pouka UF lekcija
NT zamka UF klopka	NT učenje UF obrazovanje

Slika 16: Fasete *rizik i posljedice*, *komunikacija* i *mudrost* tezaurusa u konotativnom značenju (izvor: autor)

¹² Notacija za nadređene termine (BT) nije stavljena u tezaurus jer su glavne kategorije podebljane (engl. *bold*).

Naposljetku, nakon predmetne obrade poslovice i izrade tezaurusa suvremenih poslovice u denotativnom i konotativnom značenju, donosimo primjer njihove buduće funkcionalnosti na *Portal hrvatskih poslovice*. Tako će, primjerice, korisnik koji u tražilicu upiše termin *oprez* (jer želi dobiti sve poslovice koje govore, odnosno čija je tema oprez) biti upućen (UF) na to da pretraži termin *sigurnost* jer je on preferirani termin. Tada će mu se prikazati faseta *sigurnost* (Slika 16.):

SIGURNOST UF oprez
NT nesigurnost UF neizvjesnost
RT prijetnja
RT strah
NT opstojnost
RT prilagodba
NT povoljnost
NT zaštita UF prevencija
RT sklonište

Slika 16: Faseta *sigurnost* koja će korisniku biti prikazana u mrežnom sučelju nakon pretrage po terminu *oprez* i uputnici *sigurnost* (izvor: autor)

gdje će vidjeti kako su za termin *sigurnost* uži termini (NT) *nesigurnost*, *opstojnost*, *povoljnost* i *zaštita*. Uz to što i podređeni termini imaju uputnice s neusvojenih oblika (UF), neki od njih imaju i srodne termine (RT) koji se nalaze u drugim fasetama. Na taj način će moći proširiti svoje pretraživanje dodatnim terminima koji mu možda prije uvida u tezaurus ne bi pali na pamet.

Ukoliko korisnika, primjerice, budu zanimale sve poslovice koje u konotativnom (prenesenom) značenju govore o sigurnosti, odabirom tog termina prikazat će mu se sve poslovice (i njihovi varijantni oblici, ovdje navedeni unutar zagrada) koje su označene deskriptorom *sigurnost* (Slika 17.):

SIGURNOST UF oprez
> <i>Nije zlato sve što sja.</i> (Nije zlato sve što sja.; Nije zlato sve što blista.; Nije zlato sve što sjaji.)
> <i>Uzdaj se u se i u svoje kljuse.</i> (Uzdaj se u se i na se.; Uzda se u se i u svoje ljude.)
> <i>Bolje vrabac u ruci, nego golub na grani.</i> (Bolje vrabac u ruci, nego ptica na grani.; Bolje vrabac u ruci nego golub na krovu.; Bolje vrabac u ruci nego golub na ramenu.; Bolje vrabac na ruci nego vrabac na grani.; Bolje vrabac u ruci, nego dva u grmlju.; Bolje vrabac u ruci, nego lasta na grani.; Bolje vrabac u ruci, nego 2 na grani.; Bolje vrabac u ruci nego kos na grani.; Bolje vrabac u ruci nego na grani.; Bolje vrabac u ruci, već ptica na grani.)
> <i>Neće grom u koprive.</i>
> <i>Tko radi, ne boji se gladi.</i> (Tko radi, ne zna za glad.; Tko radi neće biti gladan.; Tko radi nije gladan.; Tko radi, ne gladije.; Tko radi, zlo ne misli.)
> <i>U radu je spas.</i>
> <i>Hvali more, drž se kraja.</i> (Hvali more, drž se kopna.)
> <i>Svugdje je lijepo, ali kod kuće je najbolje.</i> (Svugdje je lijepo, ali kod kuće je najljepše.; Svugdje je lijepo, ali doma je najljepše.; Svugdje je lijepo, ali kući je najljepše.; Svugdje je lijepo, doma je najljepše.; Svugdje je lijepo, doma najljepše.; Svugdje je lijepo, ali najljepše je doma.; Svugdje je lijepo, ali je kod kuće najljepše.; Svugdje je lijepo, kod kuće je najljepše.; Svugdje je lijepo, ali doma je doma.; Svugdje je lijepo, al je doma najljepše.; Svugdje je lijepo, a doma najljepše.; Svugdje je lijepo, doma je najbolje.; Svugdje je lijepo, a kući najljepše.; Svugdje je lijepo, ali je kući najljepše.; Svugdje je lijepo, a kod kuće najljepše.; Svugdje je lijepo, doma je nalipše.; Svugdje je lijepo, kući je najljepše.; Svugdje je lijepo, ali je najljepše kući.; Svugdje je lijepo, ali najljepše je kod kuće.; Svugdje je lijepo, ali nigdje kao kući.; Svugdje je lijepo, a kod nas najljepše.; Svugdje je lijepo, a kući je ljepše.; Svugdje je lijepo, samo je kod kuće najljepše.; Svugdje je lijepo, ali ne kao kod kuće.; Svugdje je lijepo, ali nigdje kao kod svoje kuće.; Svugdje je lijepo, ali kod kuće najljepše.; Svugdje je lijepo, ali doma najljepše.)

Slika 17: Prikaz poslovice i njihovih varijantnih oblika označenih deskriptorom *sigurnost* u mrežnom sučelju (izvor: autor)

Ovakav način pretraživanja omogućuje korisnicima da brzo i učinkovito pretražuju i pronađu predmetne/tematske skupove poslovice koje odgovaraju njihovoj informacijskoj potrebi, uključujući i njihove varijantne oblike.

6 Zaključna rasprava

Rezultati ovog istraživanja potvrđuju važnost teorijskih pretpostavki organizacije informacija, predmetne obrade i izrade tezaurusa te ih nadopunjuju konkretnom primjenom na korpusu suvremenih hrvatskih poslovice. Istraživanje se oslanja na temeljna načela organizacije informacija, pri čemu je predmetna obrada poslovice modificiranim Lancasterovim (1998) postupkom predmetnog označivanja pokazala kako i minimalne tekstualne jedinice, kao što su poslovice, mogu biti sustavno i višeslojno analizirane, što je u skladu s pristupima iz literature koji naglašavaju važnost konceptualne analize i rječničke kontrole.

Korištenje prilagođenog Lancasterovog (1998) modela predmetnog označivanja, uz konzultaciju rječnika, korpusnih baza, mrežnih jezikoslovnih portala i umjetne inteligencije kao sredstva dodatne provjere potvrdilo je važnost uporabe kontroliranih rječnika za predmetno označivanje, odnosno rječničke kontrole, u skladu s međunarodnim standardom Z39.19 (ANSI i NISO, 2010), čime je postignuta dosljednost i preciznost u dodjeli i organizaciji termina. Međutim, važno je naglasiti da cilj izrade tezaurusa poslovice nije postizanje terminološke preciznosti kakva je karakteristična za znanstvene tezauruse. Umjesto toga, tezaurus poslovice usmjeren je na grupiranje poslovice prema dominantnim temama/predmetima i značenjima kako bi korisnik mogao što lakše pronaći srodne poslovice. Stoga, iako je u ovome radu postignuta određena dosljednost i preciznost u dodjeli termina korištenjem kontroliranih rječnika, naglasak i dalje ostaje na interpretaciji kulturnog i jezičnog značenja te na sposobnosti termina da povezuju poslovice u zajedničke tematske cjeline, što je neophodno za razumijevanje apstraktne i kulturno specifične građe kao što su poslovice.

Primjena Shearerove (2004) metodologije izrade tezaurusa pri izradi dvostrukog tezaurusa ishodovala je strukturiranim i preglednim kontroliranim rječnicima u denotativnom (doslovnom) i konotativnom (prenesenom) značenju, koji će korisnicima *Portala hrvatskih poslovice* omogućiti pretraživanje prema različitim varijacijama apstrakcije i značenja, što je u skladu s teorijskim implikacijama koje navodi Svenonius (2000). Dodatnu vrijednost ovakvog tezaurusa čini upravo uspostavljanje odnosa srodnosti među deskriptorima koje opisuje Urbanija (2005), koji su od velikog značaja ne samo za predmetno označivanje, već i za samu obradu informacijskih zahtjeva korisnika. Takvi odnosi vode predmetnog stručnjaka i/ili korisnika kroz tezaurus upozoravajući ga na sve pojmove bliske traženom terminu i otvarajući mogućnost otkrivanja termina koje možda nije planirao koristiti, ali koji mogu biti relevantni za predmetnu obradu ili pretraživanje.

U izradi tezaurusa na području LIS-a posebno je važno prepoznati da pojedini termini, poput *kultura*, mogu imati višestruka značenja ovisno o kontekstu primjene. Za razliku od znanstvenih tezaurusa, gdje se, kako navodi Svenonius (2000), termini nastoje definirati što univerzalnije i kontekstualno neovisno, ovaj tezaurus namjerno uključuje kontekstualno ovisne termine kako bi bolje odgovarao stvarnim potrebama korisnika i specifičnostima građe. Na primjer, termin *kultura* može u svakodnevnom označavati *sveukupnost idejne, djelatne i materijalne organizacije načina života koja služi kao sredstvo prilagodbe okolini i osiguranja opstanka pojedinca i zajednice* (Institut za hrvatski jezik i jezikoslovlje, 2025a), dok će u biologiji

značiti *proces uzgoja stanica, tkiva ili mikroorganizama u kontroliranim laboratorijskim uvjetima, izvan njihovog prirodnog okoliša* (Biology Online Dictionary, 2021). Ovakav pristup omogućuje preciznije pretraživanje, ali i razumijevanje informacija, pri čemu ističe i razlike normativne kontrole vokabulara u humanističkim i društvenim znanostima, gdje je značenje često određeno širim društvenim i kulturnim kontekstom (Svenonius, 2000).

Poštivanjem jamstava predloška, uporabe i strukture pri izradi tezaurusa omogućeno je uključivanje relevantnih i strukturno nužnih termina, ali i onih svakodnevnih, korisnicima važnih, čime je u tezaurusima naglašen element praktične i svakodnevne upotrebljivosti te se potvrđuje važnost smjernica o nužnosti prilagodbe predmetnih jezika korisničkim potrebama i analiziranoj građi koje navodi Svenonius (2000).

Međutim, unatoč brojnim prednostima ovakvog paremiološkog tezaurusa kao znanstvenog, istraživačkog i heurističkog alata, treba istaknuti ograničenja pri njegovoj izradi. Prvo ograničenje leži u činjenici kako je prilikom odabira sinonima i kvazisinonima za termine u denotativnom i konotativnom značenju bilo moguće na metodološki opravdan način osloniti se samo na Šarić-Wittschenin Rječnik sinonima hrvatskoga jezika s obzirom na to da je u trenutku provođenja ovog istraživanja to bio jedini rječnik svoje vrste na hrvatskom govornom području (Opašić, 2010). Ovakvo ograničenje izvora sinonima i kvazisinonima može bitno utjecati na razinu raznolikosti i sveobuhvatnosti paremiološkog tezaurusa. Budući da je leksička raznolikost hrvatskog jezika bogata i često regionalno uvjetovana, oslanjanjem na jedan rječnik potencijalno sužavamo opseg izraza koji bi mogli biti relevantni za paremiološku analizu, odnosno ograničavamo mogućnost uočavanja nijansi u značenju i uporabi pojedinih termina, pogotovo u konotativnom značenju gdje su semantičke razlike često suptilne i kontekstualno ovisne. Slijedom navedenog, nije moguće uputiti da se buduća istraživanja provode s više leksikografskih izvora. Međutim, daljnji razvoj hrvatske leksikografije, posebice u vidu izrade novih i specijaliziranih rječnika sinonima, bio bi od iznimne važnosti za daljnja istraživanja i projekte ove vrste. Drugo ograničenje svakako je subjektivnost predmetnog označavanja, prije svega prilikom određivanja očemnosti. Iako je za potrebe istraživanja korištena standardizirana Lancasterova metodologija predmetnog označavanja, svaki je njezin korak neminovno obilježen autorovom subjektivnom interpretacijom značenja. U kontekstu ovog istraživanja, ovaj je postupak tim više subjektivan jer se provodila predmetna obrada poslovice, koje su kulturno specifična i kontekstualno ovisna građa. Zbog toga je interpretacija značenja poslovice zahtijevala korištenje kombinacije različitih korpusnih baza i umjetne inteligencije, pri čemu je svaki alat otvorio mogućnost novih, različitih tumačenja, osobito u slučajevima gdje su se u poslovicama nalazili višeznačni izrazi, što dodatno potvrđuje neizbježnost subjektivnosti pri analizi ovakve vrste građe.

S obzirom na to da se praktična dimenzija ovog rada ogleda u integraciji tezaurusa na Portal hrvatskih poslovice, koji je namijenjen istraživačima na području jezikoslovlja, ali i onima u informacijskim znanostima te u obrazovne svrhe na svim razinama obrazovanja, u obzir je trebalo uzimati i rječnike korisnika iz svih navedenih skupina te na tragu toga oblikovati termine kontroliranog rječnika – kako bi što veći broj korisnika mogao pronaći poslovice prema vlastitom rječniku.

Ovakav metodološki pristup predmetnoj obradi i izradi tezaurusa kakvoga su postavile Aleksa Varga i Feldvari (2022) uvelike doprinosi pedagoškoj i znanstvenoj, ali i općedruštvenoj popularizaciji hrvatske paremiološke baštine. Izrađeni dvostruki tezaurus predstavlja značajan

doprinos digitalizaciji i popularizaciji hrvatske paremiološke baštine, pritom omogućujući preciznije pretraživanje, bolje razumijevanje i širu dostupnost poslovice u digitalnom okruženju. Uz to, ovakav pristup obradi paremiološke građe nudi i implikacije za daljnja istraživanja i projekte na području informacijskih znanosti, kulturne baštine i jezikoslovlja, primjerice, predmetne obrade i izrade tezaurusa hrvatskih poslovice od 12. do 19. stoljeća.

Reference

- Aitchison, J., Gilchrist, A. i Bawden, D., 2000. *Thesaurus construction and use: a practical manual*. 4. izd. London: Aslib.
- Aleksa Varga, M. i Hrisztova-Gotthardt, H., 2019. Towards a Croatian paremiological minimum/optimum: a work in progress. U: Jukić, S., Macan, Ž., Visinko, K., Vodopija-Krstanović, I. i Zubčić, S., ur. *Frazeologija, učenje i poučavanje: zbornik radova s Međunarodne znanstvene konferencije održane od 19. do 21. travnja 2018. godine u Rijeci*. Rijeka: Sveučilište u Rijeci, Filozofski fakultet. Str. 11-23.
- Aleksa Varga, M. i Feldvari, K., 2022. Constructing the digital proverbial thesaurus: theoretical and methodological implications. U: Corpas Pastor, G. i Mitkov, R., ur. *Computational and Corpus-Based Phraseology: 4th International Conference, Europhras 2022: Malaga, Spain, 28-30 September, 2022: Proceedings*. Cham: Springer. Str. 1-15.
- Aleksa Varga, M., 2023. Inačice hrvatskih poslovice. U: Mikić Čolić, A. i Glušac, M., ur. *Standardni i nestandardni idiomi: zbornik radova s međunarodnoga znanstvenog skupa Hrvatskog društva za primijenjenu lingvistiku održanoga od 9. do 11. lipnja 2022. u Osijeku*. Zagreb: Hrvatsko društvo za primijenjenu lingvistiku. Str. 171-185.
- American National Standards Institute (ANSI) i National Information Standards Organization (NISO), 2010. *Guidelines for the construction, format, and management of monolingual controlled vocabularies*. (ANSI/NISO standard Z39.19-2005 (R2010)). Baltimore, MD: National Information Standards Organization. Dostupno na: <https://doi.org/10.3789/ansi.niso.z39.19-2005R2010> [25. 2. 2025.]
- Biology Online Dictionary, 2021. *BiologyOnline*. Dostupno na: <https://biologyonline.com/dictionary/> [10. 6. 2025.]
- Chen, B., Zhang, Z., Langrené, N. i Zhu, S., 2025. Unleashing the potential of prompt engineering for large language models. *Patterns* 6(6), 101260. Dostupno na: <https://doi.org/10.1016/j.patter.2025.101260> [1. 6. 2025.]
- Doyle, C. C., Mieder, W. i Shapiro, F. R. (ur.), 2012. *The Dictionary of modern proverbs*. New Haven, CT: Yale University Press.
- Fairthorne, R., 1969. Content analysis, specification, and control. *Annual Review of Information Science and Technology*, 4, str. 73-109.
- Grzybek, P., Škara, D. and Heyken, Z. 1993. Der Weisheit der Gasse auf der Spur. – Eine empirische Pilotstudie zur Bekanntheit kroatischer Sprichwörter. *Zeitschrift für Balkanologie*, 29(2), str. 85–98.

Hrvatski jezični portal, n. d. *Hrvatski jezični portal*. Dostupno na: <https://hjp.znanje.hr/> [25. 2. 2025.]

Institut za hrvatski jezik i jezikoslovlje, 2025a. *Struna: Hrvatsko strukovno nazivlje*. Dostupno na: <http://struna.ihjj.hr/> [25. 2. 2025.]

Institut za hrvatski jezik i jezikoslovlje, 2025b. *Kolokacijska baza hrvatskoga jezika*. Dostupno na: <http://ihjj.hr/kolokacije/> [25. 2. 2025.]

International Organisation for Standardisation (ISO), 1985. *Methods for examining documents, determining their subjects, and selecting indexing terms* (ISO standard br. 5963:1985). Geneva: International Organization for Standardization.

International Organisation for Standardisation (ISO), 1986. *Guidelines for the establishment and development of monolingual thesauri*. 2. izd. (ISO standard br. 2788:1986). Geneva: International Organization for Standardization.

Kolbas, I., 2000. Načela izrade jezika za predmetno označivanje. *Suvremena lingvistika* 25(1-2), str. 153-167. Dostupno na: <https://hrcak.srce.hr/22633> [25. 2. 2025.]

Kontekst.io, 2025. *Kontekst.io*. Dostupno na: <https://kontekst.io/hrvatski> [25. 2. 2025.]

Lancaster, F. W., 1998. *Indexing and abstracting in theory and practice*. Illinois: University of Illinois.

Langridge, D. W., 1989. *Subject analysis: principles and procedures*. London: Bowker-Saur.

Litovkina, A., 2005. *Magyar közmondástár*. Budapest: Tinta Kiadó.

Mieder, W., Kingsbury, S. A. i Harder, K. B. (ur.), 1996. *A dictionary of american proverbs*. New York, NY: Oxford University Press.

Nikolić-Hoyt, A., 1996. U potrazi za boljom riječi. *Suvremena lingvistika* 22(41/42), str. 483-488. Dostupno na: <https://hrcak.srce.hr/file/38499> [25. 2. 2025.]

Olimovna, N. F. i Ibodullayevna, H. K., 2025. The role of denotation and connotation in communication. *Journal of Scientific Research, Modern Views and Innovations* 1(3), str. 210-211. Dostupno na: <https://spaceknowledge.com/index.php/JOISRMVI/article/view/943> [13. 6. 2025.]

Opašić, M., 2010. Hrabar, odvažan, izazovan... pothvat Ljiljane Šarić i Wiebke Wittschen. *Fluminensia* 22(2), str. 175-180. Dostupno na: <https://hrcak.srce.hr/file/98039> [25. 2. 2025.]

Perplexity, 2025. *Perplexity*. Dostupno na: <https://www.perplexity.ai/> [25. 2. 2025.]

Portal hrvatskih poslovice, 2025. *Portal hrvatskih poslovice*. Dostupno na: <https://poslovice.ffos.hr/> [10. 6. 2025.]

Reitz, J.M., 1996. *Thesaurus*, Online dictionary for library and information science (ODLIS). Dostupno na: https://odlis.abc-clio.com/odlis_t.html [25. 2. 2025.]

Rowley, J. i Farrow, J., 2017. *Organizing knowledge: an introduction to managing access to information*. 3. izd. London: New York: Routledge.

Shearer, J. R., 2004. A practical exercise in building a thesaurus. U: Roe, S. K. i Thomas, A. R., ur. *The thesaurus: review, renaissance, and revision*. Binghamton: Haworth International Press.

Sketch Engine, 2025. *Sketch engine*. Dostupno na: <https://www.sketchengine.eu/> [25. 2. 2025.]

Stipanov, J., 1998. Hrvatski knjižnični sustav i predmetna obrada. U: Lasić-Lazić, J., Leščić, J. i Petric, J., ur. *Predmetna obradba – ishodišta i smjernice: zbornik radova*. Zagreb: Hrvatsko knjižničarsko društvo. Str. 33-42.

Stojkova, S., 2007. *Balgarski poslovice i pogovorki*. Sofija: Izdatelstvo Kolibri.

Svenonius, E., 2009. *The intellectual foundation of information organization*. Cambridge, MA: MIT Press.

Šarić, Lj., i Wittschen, W., 2008. *Rječnik sinonima hrvatskoga jezika*. Zagreb: Naklada Jesenski i Turk.

Urbanija, J. 2005. *Metodologija izrade tezaurusa*. Zagreb: Naklada Nediljko Dominović.

Wander, K. F. W., 1867-1880. *Deutsches Sprichwörter-Lexikon: Ein Hausschatz für das deutsche Volk* (5 sv.). Leipzig: Brockhaus.