

DOI 10.57589/srl.v70i3.4019

UDK 398.9(=163.6):004.65

*Saša Babič*

ZRC SAZU, Inštitut za slovensko narodopisje

sasa.babic@zrc-sazu.si

## RAZISKOVANJE PAREMIOLOŠKIH ENOT IN NASTAJANJE ZNANSTVENE ZBIRKE PREGOVOROV V OKVIRU INŠTITUTA ZA SLOVENSKO NARODOPISJE ZRC SAZU<sup>1</sup>

Pregovori veljajo za kulturno dediščino, za modrost prednikov, estetsko lepši govor in del jezika, ki vključuje izkušnje in podobe naroda. Kot taki so bili tudi v slovenskem prostoru pogosto zapisovani v razne zvezke, zbirke ali predstavljeni v drugih oblikah od konca 18. stoletja dalje, ko je t. i. »ljudsko« pridobilo na vrednosti in v skladu z romantiko postalo želeno gradivo jezikoslovcev in folkloristov. Inštitut za slovensko narodopisje ZRC SAZU ima v svojem arhivu zbirko slovenskih pregovorov, ki temelji na zbiranju in delu več folkloristov in navdušencev. Članek je pregledni prikaz nastanka in razvoja te zbirke, temelječ na virih, ki so navedeni v zbirki.

**Ključne besede:** pregovor, arhiv, zbirka, zbiranje, folkloristika

Proverbs are considered cultural heritage, ancestors' wisdom, aesthetically more beautiful language and pieces of language containing the experiences and images of a nation. As such, they have often been presented in various volumes, collections, or in other forms in Slovenia since the end of the 18th century, when that which was "folk" gained value and, in line with Romanticism, became sought-after material by linguists and folklorists. The Institute of Slovenian Ethnography ZRC SAZU has in its archives a collection of Slovenian proverbs based on the collection and work of several folklorists and enthusiasts. The article gives an overview of the origin and development of this collection, based on the sources listed in the collection.

**Keywords:** proverb, archive, collection, collecting, folkloristics

Zbirka paremioloških enot Inštituta za slovensko narodopisje ZRC SAZU je zbir (trenutno) 36.394 paremioloških enot: pregovorov v širšem smislu, velerizmov, anti-pregovorov, vremenskih paremij itd. Ta je nastajala intenzivneje od druge polovice 1980. let dalje, ko se je urejanja in dopolnjevanja gradiva lotila Marija Stanonik, takrat na novo zaposlena na omenjenem Inštitutu. Zbirka je nastajala že prej, deloma na terenu, kjer je gradivo zbirala folkloristična sekcija, deloma s prepisovanjem starejših gradiv, obe dejavnosti pa je Stanonik še pospešila in razširila tudi znotraj svojega pedagoškega dela na Filozofski fakulteti v Ljubljani in Mariboru. Za vpogled v historiat nastanka zbirke, ki je bila v tokratnem projektu vključena v sistem Clarin.si in s tem širše dostopna, sem se oprla na seznam virov omenjene zbirke – teh je kar 2515, Stanonik pa jih je razdelila na 16 kategorij: Starejši knjižni viri (33 enot), Slovnice in slovarji (79 enot), Periodika – pratike

---

<sup>1</sup> Raziskava je nastala v okviru projekta Tradicionalne paremiološke enote v dialogu s sodobno rabo (ARRS: J6-2579) in raziskovalnega programa Etnološke, antropološke in folkloristične raziskave vsakdanjika (ARRS: P6-0088).

in koledarji (57 enot), Periodika – letopisi (10 enot), Periodika – časopisi, revije, glasila (1119 enot), Leposlovje in literarjenje (478 enot), Strokovni viri (92 enot), Objavljena slovstvena folklor (24 enot), Zbirke pregovorov (10 enot), Rokopisne zapuščine (52 enot), Muzejske zbirke (51 enot), Seminarske in diplomske naloge (10 enot), Priložnostni viri – zasebne zbirke (27 enot), Priložnostni viri – zbiralne akcije (79 enot), Priložnostni viri – osnovnošolsko gradivo (17 enot), Ustni viri (343 enot), Naključni viri (23 enot), Radio in televizija (11 enot).

## 1 Vpogled v starejši tisk

Priljubljenost paremioloških enot oz. pregovorov je kot retorično sredstvo ali ponazoritev vidna že v prvih tiskanih knjigah. Prve zapise slovenskih pregovorov najdemo kot del pridižne retorike; pri tej se je raba pregovorov razcvetela v baročnem obdobju (Babič 2011: 219), saj so njihov ritem, mestoma rima in metaforika ustrezali načelu, da »govorniška literatura obdeluje surovo gradivo glasu« (Gilbert, Kuhn 1967: 125). Raba paremioloških enot je ohranjala svojo priljubljenost tudi v poznejšem času. S folklornimi obrazci, h katerim prištevamo pregovore, so želeli govorniki doseči bolj oz. višje estetsko izražanje, z njimi so ponazarjali ali potrjevali prej povedano, poleg tega pa je bila pomembna tudi moralno-didaktična funkcija pregovorov (Babič 2011: 219). Pregovore tako najdemo v delih Jurija Juričiča, Jurija Dalmatina, Adama Bohoriča, Primoža Trubarja, Matija Kastelca, Janeza Svetokriškega, Rogerija Ljubljanskega, Jerneja Basarja, Filipa Jakoba Repeža, Franca M. Paglovca, Marka Pohlina, Urbana Jarnika, Matija Vertovca, Antona M. Slomška, Petra P. Glavarja itd. Zgodnejša dela vzporejajo latinske in slovenske paremiološke enote (npr. Svetokriški), zato pri vseh enotah ni povsem jasno, ali gre za tudi v slovenščini rabljene enote ali le za prevod latinskih paremioloških enot, medtem ko poznejša dela uporabijo le slovenske enote: slovar Matije Kastelca *Dictionarium Latino-Carniolicum* (1680–1685), rokopisni *Dictionarium Germanico-Slavonicum* (1760) Bernarda Marburgensisa, *Dictionarium Slavo-Carniolicum* (1787–1798) Blaža Kumerdeja, *Deutsch-windisches Wörterbuch* (1789) Ožbalta Gutsmana, *Slovenisches Wörterbuch / Slovensek Besednjak* (1806–1817) Valentina Vodnika, *Nemshko-Slovenske branja* (1813) Janeza Nepomuka, *Collectio vocabulorum* (1818) Frana Bilca, *Lehrbuch der Windischen Sprache* (1824) Petra Dajnka, *Lehrgebäude der Slovenischen Sprache* (1825) Frana Metelka, slovensko-nemški in nemško-slovenski slovar (1833) Antona J. Murka, *Popolni ročni slovar slovenskega in nemškega jezika* (1851) Antona Janežiča, *Deutsch-Slovenisches Wörterbuch* (1860) Mateja Cigaleta in *Slovensko-nemški slovar* (1894–1894) Maksa Pleteršnika. Navedeno kaže, da so bile paremiološke enote še vedno del jezika, ki se je uvrščal v slovarje oz. »največje sestavljene enote, ki jih še prenesejo jezikovni slovarji« (Stanonik 2015: 172). Hkrati pa iz tega lahko razbiramo tudi njihovo ponazarjalno funkcijo, ko so služili kot zgled, tj. kot jezikovna enota, ki je shranjena v jezikovnem mentalnem slovarju za namen ponazoritve pravila, s čimer je pravilo naredila lažje zapomnljivo.

Poleg starejših tiskanih knjig ter slovnice in slovarjev pa so bile polne paremioloških enot tudi periodične tiskovine. Nabori pregovorov so bili tako v pratikah in koledarjih kot tudi v periodičnem časopisju. To kaže primer v *Novi Krainški pratiki* iz 1741,

kjer je bilo natisnjenih nekaj čez 20 vremenskih pregovorov, ali pa v reviji *Vertec*, kjer je bilo v posebnem sklopu vsake številke natisnjenih do 10 pregovorov. Navajanje pregovorov ima dvojni namen: prvi je pokazati nabor pregovorov v slovenščini (pri mladinski periodiki je temu dodan namen naučiti mlade pregovorov), drugi pa poudariti folklorno modrost, izkušnje in sposobnost ubesedovanja po zahtevnejših estetskih načelih izražanja.

Tudi sicer je bila v 19. stoletju precej pestra zbirateljska dejavnost folklornega gradiva, kamor se uvrščajo paremiološke enote. Pohlin naj bi za zbiranje pregovorov navdušil Janeza Miheliča z Dobrave pri Kropi (ni jasno, ali je tam živel ali le služboval) – ta naj bi sestavil prvo samostojno zbirko pregovorov, vendar se je za njo izgubila vsaka sled. Za njegovo zbirko sprva izvemo iz Miheličeve hvale mentorja, tj. Marka Pohlina, v pesmi »Mila pesem, katiro je k hvali teh pregovorov pejl« v 1. zvezku *Pisanic 1779* (po Stanonik 2009: 114). Nato pa je zadnji ohranjeni podatek o njej v Pohlinovi komunikaciji z radeškim župnikom v času svojega ponovnega bivanja v Ljubljani med leti 1781 in 1784. Pohlinovo poročilo, ki se nanaša na dobo Miheličevega župnikovanja v Radečah (po Stanonik 2009: 114):

Mihelič »ima pri sebi (v Radečah) krajnske pregovore (Adagia Carniolica). Rokopis [...] je precej obsežen. Ako bi jih v urejeni in sistematski izdaji objavil, bi izkazal kranjski literaturi nemajhno dobroto.«

Miheličeva zbirka naj bi po njegovih besedah obsegala 1779 pregovorov v »krajnskem gvantu«, in sicer iz »gorenjske, dolenske, iz ptuje dežele« (Kidrič 2013). Nad pregovori je bil navdušen tudi Anton Martin Slomšek, ki jih je poimenoval življenjska vodila, stoletne ljudske izkušnje in ljudska modrost, zlati orehi za življenje itd., poudarjal pa je njihov vzgojni učinek (Stanonik 2009: 134). Poleg zbiranja je marsikatero misel sam oblikoval po vzoru pregovorov, tako kot je po folklornem vzoru denimo pisal basni, prilike in uganke, kar mestoma otežkoča razbiranje, kaj je folklorno gradivo in kaj je njegovo avtorsko delo. Vsekakor pa je v tem času izrednega pomena paremiološko zbirateljsko delo Stanka Vraza, Matija Vertovca, Matija Majarja-Ziljskega, Matevža Ravnikarja-Požencana, Antona Murka in drugih pomembnih zbirateljev folklornega gradiva.

Kot pomembnejši del zbiranja folklornega gradiva in s tem tudi paremioloških enot v 19. stoletju vsekakor velja tudi Štrekljev poziv *Prošnja za narodno blago*, objavljen v Ljubljanskem Zvonu leta 1887. V pozivu je razložil načrt za zbiranje in izdajo folklornega gradiva z napotki zapisovalcem. V njem je poudaril, da je treba zbirati ne le ljudske pesmi in pripovedi, pač pa tudi pregovore, reke, zagovore, šege, običaje, ljudsko pravo, medicino, igre in celo kletvice. Štrekelj sam je uspel izdati le ljudske pesmi, ostalo gradivo, ki so ga informatorji pošiljali iz slovenskega etničnega ozemlja, pa se je ohranilo v arhivu, ki ga hrani Inštitut za slovensko narodopisje ZRC SAZU.<sup>2</sup>

<sup>2</sup> Arhiv ISN: Štrekljeva zapuščina. Gradivo je tudi skenirano in dostopno na spletišču dLib pod naslovom *Rokopisna zapuščina Karla Štreklja*. [Na spletu](#).

Kot prva natisnjena samostojna zbirka slovenskih pregovorov in s tem prvi zaokrožen paremiografski rezultat zbiranja velja delo *Pregovori, prilike in reki*, ki ga je Fran Kocbek objavil leta 1887. Ta velja za mejnik kot temeljno delo slovenske paremiografije (Grzybek 2008: 24), saj je bila izdaja leta 1943 razširjena (Kocbek, Šašelj 1943), poznejše večje zbirke pa temeljijo tudi na njej, npr. Prek (1970) in Bojc (1974).

Pomemben mejnik in pionirsko raziskovalno delo v slovenski paremiologiji je nedvomno prva raziskovalna obravnava pregovorov, tj. članek Viljema Urbasa (etnologa in naravoslovca) *O pregovorih in prilikah, posebno slovenskih* (1870).

V tem obdobju se je naklonjenost pregovorom ter njihovemu zbiranju in objavljanju predvsem v časopisih nadaljevala. Lahko celo trdimo, da so bili pregovori izmed vseh folklornih obrazcev, tudi ugank, načeloma deležni največ pozornosti in zanimanja (gl. Stanonik 2009: 243–55), saj so veljali za neprecenljive stoletne izkušnje, ljudske modrosti in življenjska vodila.

## 2 Zbiranje gradiva v okviru Komisije in Inštituta za slovensko narodopisje

V letu 1947 je bila pri Slovenski akademiji znanosti ustanovljena Komisija za slovensko narodopisje, kjer sta bila prva zaposlena raziskovalca Milko Matičetov in Niko Kuret, leta 1951 pa je bil na podlagi te Komisije ustanovljen Inštitut za slovensko narodopisje. Že od začetka delovanja Komisije in pozneje Inštituta je potekalo izpisovanje folklornih enot iz starejših (zgoraj omenjenih) tiskovin ter zbiranje gradiva na terenu. Folklorni obrazci, predvsem pregovori, so bili načeloma »stranski zbir«, pa vendar zelo pomemben in vedno zapisan. V *Letopisih SAZU* (in pozneje ZRC SAZU) je mogoče tako rekoč vsako leto zaslediti omembo izpisovanja folklornih enot iz tiskovin skupaj z navedbami imen izpisovalcev: Albina Šturbej, Alojzij Bolhar, Niko Kuret, Milko Matičetov, Marija Stanonik. Neposredno omembo pregovorov v *Letopisu SAZU* pa najdemo leta 1979, ko je zabeležen prepis pribl. 15 pregovorov.

Intenzivnejše prepisovanje in zbiranje pregovorov je sprožila Marija Stanonik, ki se je kmalu po prihodu na Inštitut intenzivneje posvetila temu delu.<sup>3</sup> Prek zbiralne Akcije Kluba Nedeljskega 1983–1985 (pobudnica Marija Stanonik) in Zbiralne akcije pri tedniku Družina 2000 (pobudnika Bogdan Dolenc in Marija Stanonik), sodelovanja z osnovnošolskimi učitelji ter študenti, ki so pripravljali seminarske naloge na ljubljanski in mariborski Filozofski fakulteti, se je zbirka pregovorov znatno povečala, kar je razvidno tudi iz navedenih virov obravnavane zbirke, saj se seznam teh poveča za 106 enot. Prav tako je Stanonikova pospešila prepisovanje iz periodike in literarnih del.

Ker je bila paremiološka zbirka že precejšnja, paremiografsko delo, tj. zbiranje paremioloških enot, pa je še vedno potekalo, je leta 1997 pridobila prvi raziskovalni projekt, osredinjen na pregovore, z naslovom Slovenski pregovori in rekla, saj je bilo jasno, da je brez projekta tako obsežno zbirko nemogoče urediti. Projekt je trajal tri

<sup>3</sup> Marija Stanonik je bila na Inštitutu za slovensko narodopisje ZRC SAZU zaposlena od 1974 do 2012.

leta (1997–1999). Pregovore je v projektni prijavi opredelila kot »največje enote, ki jih še prenese jezikovni slovar, hkrati pa najmanjša žanrska enota slovstvene folklorea« (Stanonik 1996), za njihovo raziskavo pa, da je potrebno jezikoslovno, folkloristično in etnološko znanje, s čimer je paremiologijo opredelila kot interdisciplinarno vedo (pa čeprav paremiologije same ni poimenovala). Prijaviteljica je pregovore opredelila kot

»narodov[o] izkušnj[o] stoletij, modrost in poezij[o] hkrati., [...] Vpletamo jih v vsakdanji pogovor z okrasnim namenom, za podkrepitev lastne misli ali trditve ali tudi njuno relativiziranje, ko izražamo splošno resnico, nauk ali prispodob« (Stanonik 1997).

Stanonikova (1997) je v prijavi projekta tudi izpostavila, da strokovna javnost že dolgo pogreša znanstveno, komentirano izdajo pregovorov in njihovo »temeljito analizo«, hkrati pa, da je ena najtežjih nalog najti ključ za njihovo ustrezno razmejitev med njimi in razvrstitev znotraj njih.

Osnovna naloga v projektu je bila ugotoviti funkcije pregovorov tako znotraj jezikovnih struktur kot v vsakdanjem življenju. Prek strukture pregovorov je bil namen ugotoviti zakonitosti slovenske besedne ustvarjalnosti in tudi »psihološki naših prednikov«. Cilj je bil zasnovati temeljito razpravo o kratkih folklornih žanrih. Predvidena je bila tudi objava znanstveno analiziranih »pregovorov, rekov in rekel«, ki bi bila v pomoč vsem zainteresiranim: tako pedagoško vključenim kot piščičim, uporabnikom slovenščine v javnosti in medijih, pa tudi psihologom, sociologom ipd., skratka čim večjemu krogu uporabnikov. Eden izmed ciljev projekta je bil uzavestiti slovenske pregovore kot pomembno razsežnost slovenske narodne identitete ter jih predstaviti v dostopni obliki javnosti v vsej njihovi številnosti, kakovosti in polifunkcionalnosti. Predvidena je bila motivna razčlenitev, strukturna analiza, zasledovanje pogostnosti, študij rabe ter nenazadnje njihov izvor. Izpisovanje in urejanje gradiva je trajalo tri leta (1997–1999), sodelavci so izpisovali iz del od 16. stoletja dalje in iz rokopisnih arhivov, kjer so našli različne pisave (bohoričica, metelčica, dajničica, rokopisi na podlagi italijanskega in nemškega črkopisa) – te so otežkočale in upočasnjevale samo izpisovanje. Poleg historičnih virov pa so izpisovali tudi iz sodobne slovenske literature, objavljenih zbirk pregovorov in zbirk folklornih pripovedi, delno tudi iz *Slovarja slovenskega knjižnega jezika*. Ob močno obogateni zbirki slovenskih pregovorov so se odločili, da je čas za sistematično urejanje gradiva in aktivne priprave zbirke slovenskih pregovorov.

Delo je bilo tako obsežno, da je bilo treba prijaviti naslednji projekt, zaradi česar je Marija Stanonik vnovič prijavila in pridobila projekt, ki je vključeval tudi pregovore, imenovan *Informatizacija neoprijemljive dediščine za etnologijo in folkloristiko* (2005–2008). Že naslov projekta nakazuje spremembo v arhivski dejavnosti in povečan pomen digitalizacije. V ospredju projekta je bila ureditev in priprava gradiva za prenos v elektronski medij, pri čemer je bila zaobjeta vsa nesnovna dediščina s slovenskega etničnega območja. Pomemben del je bila klasifikacijska shema, ki je bila usklajena z etnološkimi in folklorističnimi sistematikami. V poročilu po drugem letu projekta je zapisano, da gre h koncu »prva faza digitaliziranja pregovorov in da se predvideva digitalna oblika združevanja njihovih variant in razvrščanja po tematskih skupinah« (Stanonik 2004), kar kaže na ponavljanje in variantnost paremioloških enot. V poročilu je izpostavljena problematičnost navajanja virov, saj je bil seznam teh izredno

obsežen. V računalnik je bilo v tem projektu prepisanih »4400 rokopisov pregovorov in paremiološko gradivo iz slovenskih glasil 19. stoletja (*Angelček, Slovenski Gospodar, Slovenski narod, Soča, Šolski prijatelj, Učiteljski tovar(i)š* itn.) ter več sto terenskih zapisov iz Slovenske Čičarije, Vipavske doline, Haloz, Savinjske doline in Kranjske Gore« (Stanonik 2008). Vseh izpiskov je bilo na tisti točki 47.200 iz več kot 715 virov (vsak vir pa naj bi imel nad 150 enot). V poročilu je vodja projekta izpostavila, da je ambicija skupine ne le pasivno prenašanje gradiva v sodobni medij, temveč tudi snovanje in preverjanje pravnih modelov za klasifikacijo tega gradiva ter oziranje k novim teoretično-metodološkim korakom. V zaključku je projekt opredeljen kot sodobni dialog s sledovi pretekle kulture in njenimi transformacijami, ki opredeljujejo tradicijo in njen pomen v vsakdanjem življenju, tradicijo pa so poskušali tematizirati tudi iz prakse, ne le na normativni in idealno-tipični ravni (Stanonik 2008). Projekt je bil velikega pomena tudi zaradi prenosa gradiva v elektronsko obliko in je postavil temelj za nadaljnje delo s sodobnejšimi metodami.

Paremiološko delo se je dve leti pozneje nadaljevalo v projektu Slovenski pregovori kot kulturna dediščina: klasifikacija in redakcija korpusa (2010–2013), ki ga je prav tako vodila Stanonik. V ospredju prijave je bila redakcija in klasifikacija gradiva. Stanonik je poudarila, da v preteklosti ni bilo enotnih meril pri izpisovanju (ali pa jih izpisovalci niso upoštevali), zato je treba vse vire ponovno pregledati in razvozlati (Stanonik 2009). Cilj preverjanja virov je bila priprava in implementacija virov na ta način, da bo programska oprema omogočala avtomatizirano razvrščanje pregovorov po virih in po indeksu izbranih besed, obenem pa bi bilo pri vsakem pregovoru navedeno, v katerih virih se pojavlja. Prav zato je bila ena osrednjih nalog projekta razvoj računalniške infrastrukture za klasifikacijo pregovorov.

Stanonikova je pri analizi gradiva poskušala vzpostaviti koncept klasifikacije s folklorističnega vidika, ki bi imela »računalniško prilagojena pravila«. Korpus se je v tem času povečal še za 8447 izpisov, tako z gradivom iz arhiva ISN kot iz periodike, koledarjev, monografij in celo iz prepisov zbirk pregovorov (Stanonik 2013). Obenem so se v tem času sistematično odpravljale napake iz dotakratnih seznamov pregovorov in virov. Med drugim je omenjeno tudi skeniranje arhiva slavista Janeza Dolenca, ki je v 40 letih poučevanja zbral pregovore z Breginjskega kota, Bovškega in Kobariškega, Tolminskega, Baške grape itd. Stanonik (2013) v projektne poročilu navaja, da je bila l. 2010 razvita osnovna podatkovna infrastruktura. Pregovori so bili pretvorjeni v programsko lahko berljivo XML obliko, pripravljene so bile podatkovne strukture za hranjenje osnovnih entitet za klasifikacijo pregovorov (znak, beseda, pregovor, skupina pregovorov, seznam skupin pregovorov). Na teh podatkovnih strukturah so bili implementirani algoritmi za osnovne statistične analize korpusa pregovorov, merjenje podobnosti pregovorov ter za klasifikacijo pregovorov. Ti osnovni algoritmi pa naj bi ustvarili temelje za razvoj algoritmov, ki bodo prilagojeni ciljni klasifikaciji pregovorov: »Klasifikacija naj bi bila narejena na podlagi indeksa besed. V ta namen so pritegnili več ekspertov s področja statističnega modeliranja in klasifikacije (kratkih) besedil. Na področju informacijske infrastrukture in algoritmične podpore projekta so bile implementirane podatkovne strukture za obdelavo podatkov o pregovorih ter

algoritmi za merjenje podobnosti pregovorov in oblikovanje osnovnih skupin na podlagi podobnosti« (Stanonik 2013).

V okviru razvoja programske opreme so definirali XML format za zapis pregovorov in pripravili podatkovne strukture za učinkovito obravnavo pregovorov in skupin/množic pregovorov; ti so bili analizirani glede na vsebovane besede in identificirane ekvivalentne besede. Identificirane so bile množice pregovorov in na osnovi tega je bil izdelan graf celotnega nabora pregovorov. Vodja projekta že v vmesnem poročilu opozarja, da je dopolnjevanje korpusa z novim gradivom presešlo predvideni časovni okvir in ta delovna faza bi se lahko še kar nadaljevala, ne le na podlagi sodobnega, aktualnega gradiva, temveč tudi zaradi popravkov starejših izpisov in zapisov gradiva. Zato se je takratna projektna skupina odločila, da zaključi pri 58.572 enotah paremiološkega gradiva. Na podlagi te količine in zbirke enot naj bi stekla zanesljiva klasifikacija in sprotna digitalizacija. Navedeno potrjuje in poudarja dejstvo, da je paremiografski del projektov nikoli dokončano delo; vedno bo odprt za vpis novega gradiva, prav tako bo vedno prostor za nove popravke. V projektu so razvili avtomatično klasifikacijo pregovorov, razvito v Laboratoriju za umetno inteligenco Inštituta Jožef Stefan, ki je bila primerna za začetno razvrščanje pregovorov, vendar so načrtovali dopolnitve programa zaradi zahtevnega nivoja natančnosti kvalifikacije.

Stanonikova je poudarila, da je priprava podatkov (pregovorov in virov) za računalniško obdelavo, tj. »poenotenje zapisa in njegovo dopolnjevanje najzamudnejša, vendar neogibna faza pri redakciji gradiva« (Stanonik 2013). Poleg tega pa je izpostavila eno najtežjih nalog, tj. klasifikacijo gradiva. Vsi modeli, ki so jih izbrali, so se izkazali za neustrezne, zato so izbrali induktivno izhodišče: v ta namen so izločili 10 % gradiva s predpostavko, da je to dovoljšen vzorec za oblikovanje klasifikacijske sheme, ki je v končni fazi izločila 260 gesel s pripadajočimi ključnimi besedami. Vsekakor je bil glavni rezultat tega projekta znanstveno zasnovana zbirka slovenskih pregovorov, ki ob koncu projekta žal ni bila širše dostopna. Poleg tega pa je zelo pomemben rezultat tega projekta članek Marije Stanonik *Slovenski pregovori kot kulturna dediščina: klasifikacija in redakcija korpusa*, ki je bila objavljen v znanstveni reviji *Traditiones* (2015), kjer natančno opiše sam projekt in paremiološko temeljito analizira gradivo iz zbirke, hkrati pa poskusi opredeliti pregovor.

Sedem let po končanem zadnjem projektu smo znova prijavili projekt, tokrat z naslovom Tradicionalne paremiološke enote v dialogu s sodobno rabo (2021–2023). Paremiografski cilj projekta je narediti dostopno zbirko slovenskih pregovorov. V okviru te ideje smo zbirko s pomočjo sodelavcev na Inštitutu Jožefa Štefana objavili kot korpus na spletišču Clarin.si. Korpus je namreč zelo primeren za nadaljnje paremiološko raziskovanje, saj omogoča hitro in enostavno iskanje, hkrati pa je dostopen širši zainteresirani javnosti.

Zbirka se je po zadnjem izvedenem projektu nahajala v obliki XLS in DOC, zato smo prečistili gradivo, tj. izločili neparemiološke enote, izločili dvojnike (iz nekaj več kot 63.000 enot smo izluščili 36.394 enot, tem pa smo dodali še dobrih 200 enot iz

zbirke Janeza Dolenca, ki jih je pripravila Barbara Ivančič Kutin), pregledali in izločili odvečne viře, zbirko smo ponovno pretvorili v XML obliko, nato pa jo je računalniška ekipa pod vodstvom Tomaža Erjavca implementirala na repozitorij Clarin.si (več o tem v Babič, Erjavec 2022), kjer je moč do znanstvene zbirke pregovorov iz arhiva ISN ZRC SAZU dostopati tako doma kot v tujini, kar je velik korak v slovenski paremiografiji in paremiologiji.

### 3 Sklep

Zbirke pregovorov, predvsem pa njihova znanstvena obdelava in prezentacija, so vedno delo več generacij in rezultat dela, ki presega stoletja. Na podlagi prvih objav so nastale prve zbirke, na temeljih prvih zbirk so grajene vse nadaljnje. Tako je tudi z zbirko pregovorov na Inštitutu za slovensko narodopisje ZRC SAZU. Ideja o inštitutski zbirki se je porodila ob nastanku Komisije za slovensko narodopisje, sistematično pa je začela nastajati s premikom k sistematičnemu zbiranju in raziskovanju paremioloških enot in s projekti Marije Stanonik, ki ni le pospešila prepisovanja iz starejših virov, temveč je s pedagoškim delom vzpodbudila zbiranje pregovorov na terenu. Seminarske in diplomske naloge so prinesle precej gradiva. Marija Stanonik je v treh projektih, pa tudi v vmesnem času, počasi, a vztrajno gradila korpus in pripadajoče viře, poskrbela za njihovo preverbo in popravke. V teku projektov je vedno znova ugotavljala, da je virov in gradiva nepregledno veliko in da je na neki točki treba začrtati mejo, hkrati pa se je zavedala, da je zbirka pregovorov rastoča zbirka, tako z vidika novih enot kot tudi dodajanja virov k obstoječim enotam. Treba je poudariti, da je prispevek Stanonikove k slovenski paremiografiji in paremiologiji s tem delom izreden. Na podlagi njenega dela smo lahko naredili korak dlje in zbirko digitalizirali po sodobnih smernicah. Iz zbirke, ki smo jo v formatu XLS hranili na Inštitutu za slovensko narodopisje ZRC SAZU, smo izločili neparemiološke enote in poenotili viře. Pregledano in redigirano gradivo je v XML formatu objavljeno na repozitoriju Clarin.si.

Kljub velikemu koraku naprej v slovenski paremiografiji pa delo ne more biti zaključeno. Vedno znova se pojavljajo novi viri in novi pregovori, ki jih nameravamo dodajati, s čimer bo zbirka pregovorov postala rastoča.

### VIRI IN LITERATURA

Saša BABIČ, 2011: Folklorni obrazci v retorskih besedilih iz 18. stoletja. *Neznano in pozabljeno iz 18. stoletja na Slovenskem*. Ur. Miha Preinfalk. Ljubljana: Zgodovinski inštitut Milka Kosa ZRC SAZU: Slovensko društvo za preučevanje 18. stoletja. 219–26. [Na spletu](#).

Saša BABIČ, Tomaž ERJAVEC, 2022: Izdelava in analiza digitalizirane zbirke paremioloških enot. *Zbornik konference Jezikovne tehnologije in digitalna humanistika =*



- Proceedings of the Conference on Language Technologies and Digital Humanities*. Ljubljana: Inštitut za novejšo zgodovino.
- Etbin BOJC, 1974: *Pregovori in reki na Slovenskem*. Ljubljana: DZS.
- Katharine EVERETT GILBERT, Helmut KUHN, 1967: *Zgodovina estetike*. Ljubljana: DZS.
- Francè KIDRIČ, 2013: Mihelič, Janez (1750–1792). *Slovenska biografija*. Ljubljana: SAZU, ZRC SAZU. [Na spletu](#).
- Peter GRZYBEK, 2008: Fundamentals of Slovenian Paremiology. *Traditiones* 37/1. 23–46. [Na spletu](#).
- Fran KOCBEK, 1887: *Pregovori, prilike in reki*. Ljubljana: A. Trstenjak.
- Fran KOCBEK, Ivan ŠAŠELJ, 1943: *Slovenski pregovori, reki in prilike*. Celje: Družba sv. Mohorja.
- Letopisi ZRC SAZU 1947–1981.
- Letopisi ZRC SAZU 1981–1997.
- Stanko PREK, 1972: *Ljudska modrost trden je most: slovenski ljudski pregovori in reki*. Maribor: Obzorja.
- Marija STANONIK, 1996: *Slovenski pregovori in rekla (1997–1999)*. Projektna prijava.
- Marija STANONIK, 2004: *Informatizacija neoprijemljive dediščine za etnologijo in folkloristiko (2005–2008)*. Projektna prijava.
- Marija STANONIK, 2008: *Informatizacija neoprijemljive dediščine za etnologijo in folkloristiko (2005–2008)*. Projektno poročilo.
- Marija STANONIK, 2009: *Zgodovina slovenske slovstvene folklore. Od srednjega veka do sodobnosti*. Ljubljana: Slovenska matica.
- Marija STANONIK, 2009: *Slovenski pregovori kot kulturna dediščina: klasifikacija in redakcija korpusa (2010–2013)*. Projektna prijava.
- Marija STANONIK, 2013: *Slovenski pregovori kot kulturna dediščina: klasifikacija in redakcija korpusa (2010–2013)*. Projektno poročilo.
- Marija STANONIK, 2015: Slovenski pregovori kot kulturna dediščina. *Traditiones* 77/3. 171–214. [Na spletu](#).
- Slovenska biografija*: Mihelič, Janez, 6. 1. 2022. [Na spletu](#).
- Viljem URBAS, 1870: *O pregovorih in prilikah, posebno slovenskih*. Jber. d. OberRealschule, Gor.

## ABSTRACT

The collection of Slovenian paremiological units in the Institute of Slovenian Ethnology ZRC SAZU has been in the making since 1947, when the Commission for Slovenian Ethnology was established and concerted work on the collection began. The collection was started on the basis of copying proverbial units from rhetorical texts from the 17th and 18th centuries (Jurij Dalmatin, Matija Kastelec, Janez Svetokriški, Marko Pohlin, Urban Jarnik, etc.), grammars and dictionaries (printed and manuscripts) from the 18th century and on (by authors such as Matija Kastelec, Bernardo Marburgensis, Ožbolt Gutsman, Valentin Vodnik, Peter Danjko, Maks Pleteršnik, etc.), as well as the enormously important paremiographic work in the 19th century (with the emphasis on Stanko Vraz, Matija Vertovec, Matija Majar-Ziljski, Matevž Ravnikar-Poženčan, Anton Murko, etc.), including all individual collecting and collecting

activities (by Karl Štrekelj: The request for folk material, 1887). The first independent collection was compiled and edited by Fran Kocbek in 1887, although the newspaper *Pisanice* reported on an impressive collection of proverbs by Janez Mihelič as early as 1779; the collection is considered lost.

Although the older sources were recognised as important for the paremiological collection, it was obvious that contemporary material was also a significant part: various folkloristic fieldwork, records of proverbs, and transcriptions from the print media provided a large number of additional paremiological units. Most of it was added under the dedicated supervision of Marija Stanonik: She facilitated intensive work on the collection with three state projects beginning in 1997 and in collaboration with students (folklore seminar and bachelor theses). Her goal was to create a scientific collection of Slovenian proverbs, including structured sources and categorization of proverbs. As trends in folklore collections changed with digitization, Stanonik also tried to follow the trend. In addition to collecting proverbs through her own fieldwork, student collaborations, and newspaper collection efforts, she organised digital transcription of material in the late 20th and early 21st centuries. At the end of each project, Stanonik noted that such a collection is a long-term project and never-ending work: there are always more proverbs to transcribe or note down, both new units and new contexts of already recorded units. The structured sources from their processing give detailed insight into the continuous use of paremiological units and at the same time the possibility to search for the proverb in context.

All the work done on the collection is essential for the current collection: after further editing of the collection in Excel form, it was converted to xml, and then implemented into the Clarin.si repository. The collection meets today's standards of a modern scientific collection of paremiological units with the indicated sources. Despite great progress in Slovenian paremiography, however, the work cannot be completed. New sources and new proverbs keep appearing for us to add, and the collection of proverbs will grow.