




Uredili / Edited by
Saša Babič, Nataša Jakop, Rok Mrvič

**Paremiologija med tradicijo
in sodobnostjo / Paremiology
between tradition and innovation**





**Paremiologija med tradicijo
in sodobnostjo / Paremiology
between tradition and innovation**



Založba ZRC

Paremiologija med tradicijo in sodobnostjo
Paremiology between tradition and innovation

Recenzentki / Reviewers: Natalija Ulčnik, Urška Valenčič Arh
Uredniki / Editors: Saša Babič, Nataša Jakop, Rok Mrvič
Jezikovni pregled / Proofreading: Nataša Jakop, Marko Petrovič
Prevod / Translation: Irena Duša Draž
Oblikovanje in prelom / Design: Jernej J. Kropelj
Izdajatelj / Issued by: ZRC SAZU, Inštitut za slovensko narodopisje
Zanj / Represented by: Saša Babič
Založnik / Published by: Založba ZRC, ZRC SAZU
Zanj / Represented by: Oto Luthar
Glavni urednik / Editor-in-Chief: Aleš Pogačnik
Tisk / Printed by: Cicero Begunje d. o. o.
Naklada / Print run: 250 izvodov

Kataložna zapisa o publikaciji (CIP) pripravili v
Narodni in univerzitetni knjižnici v Ljubljani

ISBN 978-961-05-0885-4
COBISS.SI-ID 201465091

ISBN 978-961-05-0886-1 (PDF)
COBISS.SI-ID 201448963

Prva izdaja, prvi natis. / First edition, first print.
Ljubljana 2024

Prva e-izdaja knjige (pdf) je prosto dostopna pod pogoji licence Creative Commons 4.0 CC BY-NC-SA. / First e-edition of the book is freely available in e-form (pdf) under the Creative Commons 4.0 CC BY-NC-SA: <https://doi.org/10.3986/9789610508861>.

Knjiga je nastala v okviru programa *Etnološke, antropološke in folkloristične raziskave vsakdanjika* (P6-0088, 2022–2027) in projekta *Tradicionalne paremiološke enote v dialogu s sodobno rabo* (J6-2579, 2020–2024), ki ju sofinancira Javna agencija za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije.

The book was created as part of the research programme *Ethnological, Anthropological and Folklore Studies Research on Everyday Life* (P6-0088, 2022–2027) and the project *Traditional Paremiological Units in Dialogue with Contemporary Use* (J6-2579, 2020–2024), both co-financed by the Slovenian Research and Innovation Agency.

Uredili Saša Babič, Nataša Jakop in Rok Mrvič
Edited by Saša Babič, Nataša Jakop and Rok Mrvič

**Paremiologija med tradicijo
in sodobnostjo / Paremiology
between tradition and innovation**

Ljubljana 2024

Kazalo / Table of contents

UVOD / INTRODUCTION

- 7 **Saša Babič, Nataša Jakop**
Paremiologija med tradicijo in sodobnostjo / Paremiology between tradition and innovation

I DIGITALIZACIJA PAREMIOLOŠKIH ENOT / DIGITISATION OF PAREMIOLOGICAL UNITS

- 15 **Ajda Pretnar Žagar**
Digitalne tehnologije v folkloristiki / Digital technologies in folklore studies
- 31 **Tomaž Erjavec**
The Collection of Slovenian paremiological units Pregovori: Encoding, publication and quantitative overview / Zbirka slovenskih paremioloških enot Pregovori: kodiranje, objava in kvantitativni pregled
- 49 **Outi Lauhakangas**
Typologies and digital databases of proverbs as research tools / Tipologije in digitalne zbirke podatkov pregovorov kot raziskovalna orodja
- 63 **Nikola Bakarić, Davor Nikolić**
Dataset of stylistic features of Croatian folklore genres / Nabor podatkov o slogovnih značilnostih hrvaških folklornih žanrov

II FOLKLORISTIČNA MEDŽANRSKOST / INTERGENERIC RELATIONS IN FOLKLORE

- 75 **Barbara Ivančič Kutin**
Žanrske, pomenske in zgradbene značilnosti bovških pregovorov in (primerjalnih) frazemov z leksemsko sestavino *rit* / Genre, semantic, and structural characteristics of Bovec proverbs and (comparative) idioms with the lexeme component *rit* (*ass*, *butt*)
- 93 **Monika Kropelj Telban**
The fox in Slovenian proverbs and sayings / Lisica v slovenskih pregovorih in rekih
- 107 **Rok Mrvič**
Slovensko pregovorno izročilo o besednih zvezah: pregled kontekstualnih dejavnikov v izbranih paremioloških enotah / Slovene proverbial lore on verbal bonds: An overview of contextual factors in selected paremiological units

III ETNOLINGVISTIČNA ANALIZA PAREMIOLOŠKIH ENOT / ETHNOLINGUISTIC ANALYSIS OF PAREMIOLOGICAL UNITS

- 125 **Saša Babič**
The naked truth for a good ending: The end in Slovenian paremiological units / Gola resnica za dober konec: konec v slovenskih paremioloških enotah
- 137 **Katarina Šrimpf Vendramin**
Clothing in proverbs: Between symbol and meaning / Oblačila v pregovorih: med simbolom in pomenom
- 151 **Vanja Huzjan**
Ethnological analysis of lexeme 'otrok' (child) in the Collection of Slovenian paremiological units on Clarin.si / Etnološka analiza leksema 'otrok' v Zbirki slovenskih paremioloških enot na Clarin.si

IV SODOBNA RABA PAREMIOLOŠKIH ENOT IN DIGITALNIH VIROV / CONTEMPORARY USE OF PAREMIOLOGICAL UNITS AND DIGITAL SOURCES

- 183 **Matej Meterc**
Dvajset najbolj poznanih slovenskih pregovorov in sorodnih paremioloških izrazov: podatki v paremiološkem slovarju SPP in zbirki Pregovori 1.1 / The twenty best-known Slovenian proverbs and related paremiological expressions: Data in the SPP paremiological dictionary and the Proverbs 1.1 collection
- 201 **Nataša Jakop**
Proverbs about wealth and poverty on social media: Usage and functions / Pregovori o bogastvu in revščini na družbenih omrežjih: raba in funkcije
- 217 **Dan Podjed**
'You are where you post': Changing identities through online social networks / »Si tisto, kjer objavljaš«: spreminjanje identitet na spletnih družbenih omrežjih

AVTORJI / AUTHORS 231



Paremiologija med tradicijo in sodobnostjo

Paremiology between tradition and innovation

Jezik je ohranjevalec in nosilec kulture, s katerim človeštvo ustvarja in vključuje refleksije o samem sebi (Pitkin 1972; Bartmiński 2005; Tolstaja 2015). Enako kot jezik odseva naš svet, ga tudi oblikuje: način, kako izrazimo misli, vpliva na odziv naše okolice. Oblikuje se lahko v različne oblike, ki so posredovane iz ene generacije v drugo. Pregovor je folklorni žanr, v katerem se folklorni artefakti in entitete oblikujejo v generične tipske obrazce, ti pa se lahko prenašajo skozi prostor (Georges in Jones 1995). Paremiološke enote oz. pregovori so ustaljeni obrazci, nastali v določeni družbi, času in prostoru ter jih uporabljamo v vsakodnevem jeziku. So kratki, stalni in ponavljajoči se izrazi z namenskim pragmatičnim ciljem v različnih dialogih, ki vsebujejo različne kulturne nazore.

Pregovori še niso bili dokončno definirani, vendar jih lahko opišemo kot relativno stalne povedi, ki jih uvrščamo med kratke folklorne obrazce. Pogosto so označeni kot »modrost ljudstva« (Taylor 1931; Mieder 1993), »stara modrost« in »poezija vsakdanjega jezika« (Matičeto 1956). Splošne značilnosti pregovorov so, da so kratki stereotipi na sentenčni ravni, s prenesenim ali generalizirajočim pomenom, so splošno znani in se prenašajo iz ene generacije v drugo (Grzybek 2012). Pregovori so kulturna besedila z velikim semantičnim potencialom (Grzybek 2014), vendar pa se ne razlikujejo le po besedilu, temveč tudi glede na teksturo in kontekst (Dundes 1965). Zaradi prozodičnih značilnosti si jih je lažje zapomniti, dandanes pa zato ponujajo možnosti za nadaljnjo uporabo, na primer pri oglaševanju, sodobnem prenosu mnenj, grafitih ali modifikacijah v različnih medijih. Semiotična kompleksnost pregovorov in prepletenost med sintaktično (kratkost), pragmatično (prenašanje skozi različne generacije) in semantično (stereotipno, splošno znanje) razsežnostjo ponujajo raziskovanje pregovorov kot kulturni znak, ki ohranja zgodovino kulture s prenosom iz roda v rod, hkrati pa sprejema nove funkcije, ki širijo in porajajo nove kontekste (npr. oglasi, šale itd.).

Čeprav pregovori danes spadajo na področje paremiologije, so pogosto tudi raziskovalni predmet folkloristike, sociologije, pedagogike, jezikoslovja, etnolingvistike, semiotike itd. Za prav vsa področja pa je digitalna humanistika tista, ki omogoča raziskovalcem, da sprejmejo nove, korenito drugačne metode raziskovanja in, kar je prav tako pomembno, da so na voljo elektronske zbirke z naprednimi možnostmi iskanja podatkov (Zhitomirsky-Geffet idr. 2015). Korpusno jezikoslovje in trenutno priljubljena »metodologija branja na daljavo« poskuša izkoristiti velike jezikovne vzorce, da bi

pridobili (kvantitativni) vpogled v slovar, uporabo, trende in vizualizacije na področjih jezikovnega interesa. Takšne zbirke je treba skrbno strukturirati, da bi dosegli svoj polni potencial. Vse to so izzivi, s katerimi se soočajo tudi avtorji pričujoče monografije.

Vsebina monografije je razdeljena na štiri tematske sklope. Prvi tematski sklop **DIGITALIZACIJA PAREMIOLOŠKIH ENOT** je namenjen razpravi o teoretičnih in praktičnih vidikih sodobnih metod zbiranja, obdelovanja in analiziranja folklorističnega gradiva v dobi pospešene digitalizacije. **Ajda Pretnar Žagar** v poglavju *Digitalne tehnologije v folkloristiki* prikaže možnosti razvoja digitalne folkloristike glede integracije digitalnih tehnologij v folkloristiko v vseh fazah, od zbiranja, shranjevanja gradiva ter organizacije digitaliziranih podatkov z uporabo statističnih analiz in strojnega učenja za označevanje, segmentacijo in semantično analizo. **Tomaž Erjavec** v poglavju *Zbirka slovenskih paremioloških enot Pregovori: kodiranje, objava in kvantitativni pregled* opiše nastanek, lastnosti, dostopnost, kvantitativni pregled ter uporabno vrednost prve slovenske digitalne in javno dostopne zbirke paremioloških enot Pregovori 1.0. **Outi Lauhakangas** v poglavju *Tipologije in digitalne zbirke podatkov pregovorov kot raziskovalna orodja* sledi ideji o mednarodni tipologiji pregovorov in predstavi praktične možnosti raziskovanja paremiologije na osnovi take digitalne zbirke. **Nikola Bakarić** in **Davor Nikolić** v poglavju *Nabor podatkov o slogovnih značilnostih hrvaških folklornih žanrov* opišeta podatke računalniške analize z uporabo skripta za samodejno zlaganje in ekstrakcijo n-gramov za nekatere hrvaške folklorne žanre.

V drugem tematskem sklopu **FOLKLORISTIČNA MEDŽANRSKOST** se ukvarjamo z žanrsko analizo in interpretacijo različnih tipov paremioloških enot. **Barbara Ivančič Kutin** v poglavju *Žanrske, pomenske in zgradbene značilnosti bovških pregovorov in (primerjalnih) frazemov z leksemsko sestavino rit* prikaže žanrsko raznolikost, pomen, funkcije in jezikovno-strukturne značilnosti bogatega narečnega bovškega paremiološkega gradiva, in sicer na primerih somatskih paremij s sestavino *rit*. **Monika Kropelj Telban** v poglavju *Lisica v slovenskih pregovorih in rekih* interpretira slovenske pregovore o lisici glede na njihovo širše sobivanje v mednarodnem literarnem in umetnostnem kontekstu. **Rok Mrvič** v poglavju *Slovensko pregovorno izročilo o besednih zavezah: pregled kontekstualnih dejavnikov v izbranih paremioloških enotah* opiše poskus novega pristopa k določanju žanrskih lastnosti na osnovi iz paremij izluščenih kontekstualnih dejavnikov, ki tematizirajo obljubljanje, priseganje in sklepanje besednih zavez.

V tretjem tematskem sklopu **ETNOLINGVISTIČNA ANALIZA PAREMIOLOŠKIH ENOT** SO Z entolingvističnega oz. etnološkega vidika analizirane izbrane antropocentrične teme, kot so minljivost, oblačilna kultura in otrok. **Saša Babič** v poglavju *Gola resnica za dober konec: konec v slovenskih paremioloških enotah* raziskuje možnosti etnolingvističnega pristopa in semiotike pri razkrivanju globljih struktur in pomenov ter družbenih stereotipov v povezavi s paremijami, ki tematizirajo konec. **Katarina Šrmpf Vendramin** v poglavju *Oblačila v pregovorih: med simbolom in pomenom* z analizo pregovorov, ki vsebujejo oblačila, raziskuje metaforiko in simboliko slovenske oblačilne kulture.

Vanja Huzjan v poglavju *Etnološka analiza leksema 'otrok' v Zbirki slovenskih pregovorov na Clarin.si* s poglobljeno etnološko analizo opisuje skrivnostni svet socializacije otrok na Slovenskem v preteklih stoletjih.

Četrty sklop **SODOBNA RABA PAREMIOLOŠKIH ENOT IN DIGITALNIH VIROV** prinaša jezikoslovno analizo sodobne rabe in funkcij paremioloških enot oz. razmislek o lastnostih sodobnih medijev in vplivov na njihovo rabo na splošno. **Matej Meterc** v poglavju *Dvajset najbolj poznanih slovenskih pregovorov in sorodnih paremioloških izrazov: podatki v paremiološkem slovarju SPP in zbirki Pregovori 1.1* predstavi vrh slovenskega paremiološkega minimuma, in sicer prvih 20 najbolj poznanih paremij med govorce slovenskega jezika. **Nataša Jakop** v poglavju *Pregovori o bogastvu in revščini na družbenih omrežjih: raba in funkcije* analizira rabo in funkcije pregovorov o bogastvu in revščini v komunikaciji na družbenih omrežjih. **Dan Podjed** v poglavju »*Si tisto, kjer objavljaš*«: spreminjanje identitete na spletnih družbenih omrežjih razmišlja o lastnostih identitete posameznika – sodobnega uporabnika paremij – na družbenih omrežjih z antropološkega vidika.

Knjiga, ki je pred vami, vsebuje preplet digitalne humanistike, paremiografije – tj. gradiva pregovorov, paremiologije, folkloristike, jezikoslovja, antropologije. Prikazani so tako postopki digitalizacije folklorne gradiva, uporaba spleta kot terenski del raziskav, folkloristični vpogledi v medsebojna razmerja žanrov in kot tudi kulturološke in jezikoslovne analize sodobnih rab. Knjiga prinaša poglavja, razdeljena v štiri sklope, prav v vsakem pa se najdejo različni pristopi mednarodno uveljavljenih raziskovalcev.

Language is a vehicle for preserving and transmitting culture, which humanity employs to create and integrate reflections on itself (Pitkin 1972; Bartmiński 2005; Tolstaja 2015). Just as language reflects our world, it also gives it shape. The way we express our thoughts affects the way our surroundings react. Language can take different forms, which are passed on from one generation to the next. This project focuses on a specific language form, the proverb, which is a folklore genre in which folklore artefacts and entities are configurable into generic type sets and transmissible through space (Georges and Jones 1995). Proverbs are the fixed phrases used in everyday language and determined by a particular community, time and place. They are short, fixed and recurring traditional expressions with an intentional pragmatic purpose in different dialogues involving different cultural viewpoints.

Although an unequivocal definition of proverbs remains elusive, they can be described as relatively fixed expressions that are classified as short forms of folklore. They have been labelled “the wisdom of the people” (Mieder 1993), “ancient wisdom” and “the poetry of everyday language” (Matičev 1956). Proverbs are generally characterised

as being short stereotypes on a sentence level, with a figurative or generalising meaning. They are common knowledge and are passed down from generation to generation (Grzybek 2012). Proverbs are cultural texts with great semantic potential (Grzybek 2000), which differ not only in textual aspect, but also in texture and context (Dundes 1965). Their prosodic characteristics make them easier to remember and offer possibilities for further use, for example in advertising, opinion-forming, graffiti or modified in various media. The semiotic complexity of proverbs and the close interactions between their syntactic (brevity), pragmatic (intergenerational transmission and persistence) and semantic (stereotypy, common knowledge) dimensions offer the possibility of exploring proverbs as a cultural sign that preserves the history of a culture by being transmitted from one generation to the next, while at the same time taking on new functions that expand and create new contexts (e.g. advertisements, jokes, etc.).

Although proverbs nowadays belong to the field of paremiology, they are also often the subject of research in folklore studies, sociology, pedagogy, linguistics, ethnolinguistics, semiotics, etc. The digital humanities are enabling researchers to adopt new, radically different research methods in all fields. Just as importantly, they are making available electronic repositories with advanced data retrieval capabilities (Zhitomirsky-Geffet et al. 2015). Corpus linguistics and the currently popular “distant reading methodology” attempt to exploit large linguistic samples to gain (quantitative) insights into vocabulary, usage, trends and visualisations in areas of linguistic interest. Such collections need to be carefully structured to reach their full potential. All of these are challenges faced by the authors of the present monograph.

The content of the monograph is divided into four thematic sections. The first thematic section entitled **DIGITISATION OF PAREMIOLOGICAL UNITS** discusses the theoretical and practical aspects of modern methods of collecting, processing and analysing folklore material in an era of accelerated digitisation. In the chapter *Digital technologies in folklore studies*, **Ajda Pretnar Žagar** presents the possibilities for developing digital folklore studies in terms of integrating digital technologies at all stages, from collecting and storing materials to organising digitised data using statistical analysis and machine learning for labelling, segmentation and semantic analysis. In the chapter entitled *The Collection of Slovenian paremiological units Pregovori: Encoding, publication and quantitative overview*, **Tomaž Erjavec** describes the origin, characteristics, accessibility, quantitative overview and utility value of the first Slovenian digital and publicly accessible collection of paremiological units, Pregovori 1.0. In the chapter *Typologies and digital databases of proverbs as research tools*, **Outi Lauhakangas** pursues the idea of an international typology of proverbs and presents the practical possibilities of researching paremiology on the basis of such a digital collection. In the chapter entitled *Dataset of stylistic features of Croatian folklore genres*, **Nikola Bakarić** and **Davor Nikolić** describe the data of a computer analysis using a script for the automatic stacking and extraction of n-grams for some Croatian folklore genres.

In the second thematic section, **INTERGENERIC RELATIONS IN FOLKLORE**, we deal with genre analysis and the interpretation of different types of paremiological units. In the chapter entitled *Genre, semantic, and structural characteristics of Bovec proverbs and (comparative) idioms with the lexeme component rit (ass, butt)*, **Barbara Ivančič Kutin** demonstrates the genre diversity, meaning, functions and linguistic-structural features of the rich dialectal paremiological material from the region of Bovec, using examples of somatic paremias with the constituent *rit*. In the chapter *The fox in Slovenian proverbs and sayings*, **Monika Kropelj Telban** interprets Slovenian proverbs about foxes in relation to their broader coexistence within an international literary and artistic context. In the chapter entitled *Slovene proverbial lore on verbal bonds: An overview of contextual factors in selected paremiological units*, **Rok Mrvič** describes an attempt at a new approach to determining genre characteristics on the basis of contextual factors extracted from paremias, by focusing on promising, swearing and making verbal bonds.

In the third thematic section entitled **ETHNOLINGUISTIC ANALYSIS OF PAREMIOLOGICAL UNITS**, selected anthropocentric themes such as transience, dress culture and the child are analysed from an ethno-linguistic or ethnological point of view. In the chapter *The naked truth for a good ending: The end in Slovenian paremiological units*, **Saša Babič** examines the potential of an ethno-linguistic approach and semiotics in revealing underlying structures and meanings, as well as social stereotypes in relation to paremias that address the end. In the chapter *Clothes in proverbs: Between symbol and meaning*, **Katarina Šrampf Vendramin** analyses proverbs that mention items of clothing, in order to explore the metaphors and symbolisms of Slovenian clothing culture. In the chapter entitled *Ethnological analysis of lexeme 'otrok' (child) in the Collection of Slovenian paremiological units on Clarin.si*, **Vanja Huzjan** provides an in-depth ethnological analysis to describe the mysterious world of the socialisation of children in Slovenia in past centuries.

The fourth section entitled **CONTEMPORARY USE OF PAREMIOLOGICAL UNITS AND DIGITAL SOURCES** provides a linguistic analysis of the contemporary use and functions of paremiological units, together with a reflection on the characteristics of contemporary media and their impact on the use of paremiological units in general. In the chapter entitled *The twenty best-known Slovenian proverbs and related paremiological expressions: Data in the SPP paremiological dictionary and the pregovori collection*, **Matej Meterc** presents the pinnacle of the Slovenian paremiological minimum, namely the 20 best-known paremias among speakers of Slovene. In the chapter *Proverbs about wealth and poverty on social media: Usage and functions*, **Nataša Jakop** analyses the use and functions of proverbs on wealth and poverty in communication on social media. In the chapter *'You are where you post': Changing identities through online social networks*, **Dan Podjed** considers the characteristics of a contemporary paremia user's identity on social media, from an anthropological perspective.

This book contains a mixture of digital humanities, paremiography (proverbial material), paremiology, folklore studies, linguistics and anthropology. It presents the processes of digitising folklore material, the use of the web as a fieldwork component of research, folkloristic insights into the interrelationships of genres, as well as cultural and linguistic analyses of contemporary usages. The chapters are divided into four sections, each of which contains different approaches by internationally renowned researchers.

Literatura / References

- Bartmiński, Jerzy, 2005: *Jazykovej obraz mira: očerki po etnolingvistike*. Moskva: Indrik.
- Dundes, Alan, 1965: *The study of folklore*. Englewood Cliffs, NJ: Prentice-Hall.
- Georges, Robert A.; Jones, Michael O., 1995: *Folkloristics: an introduction*. Bloomington: Indiana University Press.
- Grzybek, Peter, 2012: Proverb Variants and Variations: A New Old Problem? V: Outi Lauhakangas, Rui J.B. Soares (Eds.), *Proceedings of the Fifth Interdisciplinary Colloquium on Proverbs*. Tavira: IAP, 136–152.
- Grzybek, Peter, 2014: Semiotic and Semantic Aspects of the Proverb. V: Hrisztova-Gotthardt, Hrisztalina; Varga, Melita Aleksa (eds.), *Introduction to Paremiology: A Comprehensive Guide to Proverb Studies*. Warsaw, Berlin: de Gruyter, str. 68–111.
- Matičeto, Milko, 1956: *Pregovori in uganke; ljudska proza*. Ljubljana: Slovenska matica.
- Mieder, Wolfgang, 1993: *Proverbs are never out of season: Popular wisdom in modern age*. New York, Oxford: Oxford University Press.
- Pitkin, Hanna F., 1972: *The concept of representation*. Berkeley, Los Angeles, London: University of California Press.
- Taylor, Archer, 1931: *The proverb*. Cambridge: Harvard University Press.
- Tolstaja, Svetlana M., 2015: *Obraz mira v tekste i rituale*. Moskva: Universitet Dmitrija Požarskogo.
- Zhitomirsky-Geffet, M.; Prebor, G.; Bloch, O., 2015: *Indexing and retrieval of non-text information*. Berlin, Boston: De Gruyter Saur.



I

Digitalizacija paremioloških enot
Digitisation of paremiological units



Digitalne tehnologije v folkloristiki

Digital technologies in folklore studies

Članek pregleda digitalne tehnologije, ki podpirajo folklorne raziskave, tako da sledi celotnemu ciklu podatkov. Cilj je navdihniti začetnike v digitalni humanistiki in računalniški folkloristiki s predstavitvijo ključnih tehnologij, ki pomagajo digitalnemu raziskovanju. Članek opiše proces zbiranja podatkov, predobdelave, analize in shranjevanja. Zbiranje podatkov vključuje tradicionalno terensko delo, zbiranje obstoječih podatkov s spleta ali preko specializiranih repozitorijev. Koraki zbiranja podatkov so predstavljeni in pojasnjeni v načrtu upravljanja podatkov. Nato se podatki predobdelajo z digitalizacijo analogne vsebine z optičnim prepoznavanjem znakov in pretvorbo govora v besedilo. Predstavljena sta formata TEI XML in CoNNL-U za shranjevanje jezikovnih podatkov ter našete ključne tehnike obdelave naravnega jezika, od katerih jih je mnogo na voljo za slovenščino. Računalniške metode organizirajo digitalizirane podatke, pri čemer se uporabljajo statistične analize in strojno učenje za označevanje, segmentacijo in semantično analizo. Predstavimo brezplačno dostopna spletna orodja za analizo podatkov, ki segajo od manj tehničnih konkordančnikov do bolj naprednih orodij strojnega učenja. Nazadnje obravnavamo shranjevanje podatkov ter naštejemo ključne repozitorije za jezikovne podatke. Kljub naprednim orodjem ostaja človeško razumevanje ključno za kontekstualizacijo ugotovitev in določanje pomena pri raziskovanju folklore.

• **Ključne besede:** digitalna folkloristika, življenjski cikel podatkov, digitalne tehnologije, računalniška folkloristika

This article reviews the digital technologies that support folklore research by tracing the full data cycle. The aim is to inspire beginners in digital humanities and computational folklore studies by presenting key technologies that support digital research. The article describes the process of data collection, pre-processing, analysis and storage. Data collection involves traditional fieldwork, collecting existing data from the web or via specialised repositories. The steps of data collection are presented and explained in the data management plan. The data are then pre-processed by digitising the analogue content through optical character recognition and speech-to-text conversion. The TEI XML and CoNNL-U formats for storing linguistic data are presented and key natural language processing techniques are listed, many of which are available for Slovene. Computational methods organise the digitised data, using statistical analysis and machine learning for labelling, segmentation and semantic analysis. We present freely available online tools for data analysis, ranging from less technical concordancers to more advanced machine learning tools. Finally, we discuss data storage and list key repositories for linguistic data. Despite advanced tools, human understanding remains key to contextualising findings and determining meaning in folklore research.

• **Keywords:** digital folklore, data lifecycle, digital technologies, computational folklore

1 Uvod

Vsako znanstveno raziskovanje temelji na podatkih, bodisi kvalitativnih, kvantitativnih ali nekaj vmes. V folkloristiki zajemajo raziskovalni podatki besedila, avdio posnetke, slike, video posnetke in materialno kulturo ljudskih pravljič, zgodb, šal in drugih primerov ljudskega izražanja. Tipičen cikel podatkov v sodobnem raziskovanju vključuje zbiranje podatkov, predobdelavo (pripravo), analizo in raziskovanje, interpretacijo ter ohranjanje. Veliko teh korakov je bilo nekoč opravljenih ročno, medtem ko jih danes večinoma podpirajo sodobne strojne in programske rešitve. Zvok se ne snema več na trakove, ampak v digitalni obliki. Posnetki se ne prepisujejo več ročno, ampak s pomočjo orodij na osnovi umetne inteligence. Podatki se ne štejejo več ročno, ampak se lahko interaktivno prikažejo. Nazadnje, podatki ne ležijo več (samo) v fizični obliki v arhivu, ampak so shranjeni digitalno.

Podatke običajno pridobivamo z organiziranimi raziskovalnimi procesi (digitaliziranimi ali ne). Vendar to ni edini vir podatkov, ki je na voljo za raziskovanje folklore. V zadnjih letih se uveljavlja digitalna folkloristika. Ta veja folkloristike se osredotoča na raziskovanje spletnega ljudskega izražanja in vključuje študije spletnih šal, memov, GIF-ov, internetnih legend in mitov (Shifman 2013; Chiaro 2018; Yogarajah 2022). Prednost digitalne folkloristike je dostopnost podatkov, ki jih je mogoče enostavno zajeti ali prenesti s spleta. Takšni podatki se imenujejo »podatki iz divjine«, če si izposodimo izraz Susan Halford (Halford 2017; Halford opozarja, da izraz prinaša svoj niz težav). V razdelku 2 se posvetimo pridobivanju podatkov ter načrtu za upravljanje podatkov v digitalni folkloristiki.

Programska orodja zahtevajo, da so podatki na voljo v digitalni obliki. Vsi podatki, avdio in video posnetki, slike in besedilo, so lahko v analogni ali digitalni obliki (zvočni trak vs. zvočna datoteka, papirnata fotografija vs. digitalna fotografija, papir vs. besedilna datoteka). Pretvorba prvih v drugo ni nič kaj enostavna. Dva običajna pristopa za digitalizacijo analognih besedilnih in zvočnih podatkov sta optično prepoznavanje znakov (OCR) in tehnologije pretvorbe govora v besedilo (STT).¹ V razdelku 3 so predstavljeni sodobni pristopi k digitalizaciji folklornih podatkov ter tehnike za obdelavo naravnega jezika.

Folkloristika je po svoji naravi interdisciplinarna veda. Zato je ena najmočnejših članic na področju digitalne humanistike, ki uporablja računalniške pristope za študije humanističnih podatkov. Računalniško raziskovanje folklore sega od ustvarjanja zbirk ljudskih pravljič (Karsdorp idr. 2015; Meder idr. 2023) do uporabe statistike in strojnega učenja za identifikacijo motivov (Thuillard idr. 2018; Eklund idr. 2023). V razdelku 4

¹ Tudi predmeti se vse bolj digitalizirajo s 3D tehnologijo, kar raziskovalcem omogoča študij predmetov na daljavo.

opišemo orodja za računalniško-podprto analizo besedilnih podatkov, od enostavnejših konkordančnikov do orodij na osnovi umetne inteligence.

Shranjevanje podatkov je pomemben sestavni del življenjskega cikla podatkov. Spletne zbirke podatkov spodbujajo ponovno uporabo, primerjalno raziskovanje in razvoj novih analitičnih pristopov. Na voljo je vse več skrbno urejenih zbirk folklore in kulturne dediščine. V razdelku 5 obravnavamo FAIR izhodišča za shranjevanje podatkov, na katerih so osnovane odprte spletne zbirke za raziskovanje folklore. V članku je poudarek na folkloristiki, vendar večina pristopov velja za humanistiko na splošno. Članek obravnava vsako stopnjo cikla podatkov in podaja primere digitalnih tehnologij, ki jih podpirajo. Kjer je mogoče, so navedena najnovejša orodja za slovenski jezik.

2 Stopnja 1: Zbiranje podatkov

Folklorni podatki vključujejo besedila, avdio in video posnetke, slike ter predmete. Predmeti so nekoliko posebni, zato jih bomo izpustili iz pregleda, čeprav sodobni pristopi tridimenzionalnega modeliranja omogočajo digitalizacijo materialne dediščine (Kingsland 2020). V grobem obstajajo trije pristopi k zbiranju podatkov v folkloristiki: s terenskim delom, z uporabo obstoječih virov in preko spleta. Tudi pri terenskem delu se podatki digitalno beležijo v avdio (ali video) formatu. Avdio in video posnetke je mogoče analizirati v njihovi izvorni obliki ali pretvoriti v besedilo, kar bomo obravnavali v naslednjem razdelku. Druga možnost, še posebej v digitalni folkloristiki, je zbiranje podatkov na spletu. Danes obstaja veliko spletnih repozitorijev, ki vsebujejo takoj dostopne strukturirane folklorne podatke (Babič idr. 2022, Fišer idr. 2018, Verdonik in Zwitter Vitez 2011).

Spletni podatki obsegajo razprave na forumih, objave na družbenih omrežjih, bloge in vloge, klepetalnice, niti na Redditu, meme, komentarje in druge vrste spletnih podatkov. Ang idr. (2013) temu pravijo »najdeni podatki«, medtem ko Krawczyk-Wasilewska (2016) to imenuje »e-folklor«. Za Susan Halford (2017) gre za t. i. »podatke iz divjine«, čeprav opozarja, da izraz prinaša svoj niz težav. Glavna značilnost takih podatkov je, da niso bili zbrani z določenim raziskovalnim vprašanjem (ali vprašanji) v mislih.

Takšni podatki še vedno lahko predstavljajo vrednost za raziskovanje folklore. Vsebujejo odgovore na določena vprašanja, na primer, »Kaj je ljudem smešno na spletu?«, »Kako se s širjenjem po spletu spreminjajo spletne pripovedi?« ali »Kateri motivi obstajajo zgolj v internetnih skupnostih?«. So takoj dostopni, še posebej za raziskovalce z nekaj programerskimi spretnostmi. So izvorno digitalizirani in zato je po njih enostavno iskati. Vendar pa imajo spletni podatki več težav, med drugim pristranskost vzorca, zanesljivost in lastništvo.

Do pristranskosti vzorca pride, ko podatki (vzorec) ne predstavljajo populacije, kar pogosto velja za spletne podatke. Nekateri ljudje ne sodelujejo v spletnih razpravah

in zato niso vključeni v takšne študije. Čeprav bi lahko trdili, da le-ti posledično niso populacija študije, je to treba izrecno navesti v raziskovalnem načrtu in upoštevati pri analizi. Poleg tega je zelo težko določiti demografsko porazdelitev vzorca. Nekateri deli populacije so lahko pod- ali preveč zastopani v izbranem podatkovnem vzorcu.

Spletni podatki imajo vprašljivo lastništvo. Hkrati so predmet avtorskih pravic, zaščite zasebnosti in pogojev uporabe. Ko so podatki »najdeni« na spletu, niso podani z obrazcem za soglasje in udeleženci niso seznanjeni, da sodelujejo v raziskavi. Ne morejo se umakniti iz raziskave, preprosto zato, ker bi bilo preverjanje vsakega subjekta logistično nemogoče. Čeprav so spletni podatki javni podatki, če je za dostop do njih potreben račun, je javna narava podatkov že postavljena pod vprašaj, saj jih lahko vidijo le člani določene skupnosti. Spletni podatki so torej podvrženi istim etičnim merilom kot terenski podatki - ohranjanju človeškega dostojanstva, zaščiti udeležencev raziskave, maksimiranju koristi, minimiziranju tveganj ter zagotavljanju spoštovanja in pravičnosti (Halford 2017).

Faza zbiranja podatkov mora vključevati načrt upravljanja podatkov (NUP). Sestavljanje takega načrta lahko raziskovalcem zagotovi, da se bo s podatki učinkovito upravljalo, jih ohranjalo in delilo, kar poveča njihovo vrednost in vpliv. Načrt upravljanja podatkov opredeli vrsto podatkov, ki bodo zbrani (tj. format, obseg in standarde), kako se bodo podatki zbirali (tj. metode in orodja), kako se bodo podatki opisali (tj. metapodatkovne standarde), kje se bodo shranili (tj. varnostne kopije in varnost podatkov) ter kako se bodo delili (tj. repozitoriji ali arhivi). Prav tako načrt opredeli lastniške pravice in licence, etične in pravne vidike glede podatkov ter kdo je odgovoren za kuriranje podatkov.

Načrt za upravljanje podatkov je na prvi pogled videti zastrašujoč, vendar je lahko izredno koristen za raziskovalca, saj mora ta upoštevati vse vidike rokovanja s podatki. Hkrati taki načrti postajajo vseprisotna dobra praksa v mnogih raziskovalnih disciplinah, vključno s folkloristiko. Kung (2022), na primer, predlaga zaokrožen NUP za jezikovne podatke, obstajajo pa tudi specializirane metapodatkovne sheme, ki obravnavajo heterogenost podatkov v folkloristiki (Lourdi idr. 2007).

3 Stopnja 2: Predobdelava podatkov

Terensko delo v folkloristiki običajno privede do zbirke avdio in video datotek. Če se datoteke nanašajo na pesmi in plesne predstave, se lahko datoteke analizirajo neposredno z uporabo tehnik avdio segmentacije (Marolt idr. 2019), pridobivanja glasbene informacije (Wiering idr. 2009), prepoznavanja slik (Lee 2022) ali multimodalne analize (Smits in Wevers 2023). Vendar pa, če nas zanima vsebina, morajo biti datoteke digitalizirane v zaporedje znakov, drugače povedano, v besedilo.

3.1 Avtomatska transkripcija

Besedilo je eden najpogostejših tipov podatkov v digitalni humanistiki. Običajno, ko je ljudski izraz posnet, se avdio datoteka transkribira in postane besedilo. Tehnologija pretvorbe govora v besedilo (STT) je prebojna za sodobno folkloristiko, saj prihrani dragocen čas pri transkripciji. Avtomatsko prepoznavanje govora (ASR) je model strojnega učenja, ki je usposobljen za prepoznavanje besed in fraz v govornem jeziku ter njihovo pretvorbo v besedilo.

Priljubljeni modeli STT vključujejo zasebne Google DeepSpeech, Microsoft Azure Speech to Text in IBM Watson Speech to Text. Trije odprtokodni modeli so Kaldi (Povey idr. 2011), Whisper (Radford idr. 2023) in Slovene Conformer² (Lebar Bajec idr. 2022). Whisper, Slovene Conformer, Azure Speech to Text in DeepSpeech podpirajo slovenščino. Čeprav so sodobni modeli STT nepopolni, imajo povprečno 12-odstotno stopnjo napake pri besedah (Radford idr. 2023), kar običajno zagotavlja uporabno začetno transkripcijo. Služijo lahko kot začetna točka za raziskovalca, ki popravlja napake, namesto da bi celotno besedilo vpisoval ročno.

Pretvorba govora v besedilo je močno odvisna od kakovosti podatkov. Natančnost transkripcijskega modela se zmanjšuje z nižjo akustično kakovostjo, hrupom v ozadju, prekrivanjem govora, besedami zunaj slovarja in narečji ter mešanjem jezikov. Nižja akustična kakovost je običajno prisotna pri starejših posnetkih, medtem ko je hrup v ozadju pogost pri neformalno posnetem gradivu (npr. v lokalni, na festivalu ali na zabavi). Prekrivanje govora se nanaša na problem določanja govorcev, imenovan tudi diarizacija govorca (Park idr. 2022). Pred poskusom analize je treba avdio posnetek razdeliti po govornih, da model STT pravilno deluje. Modeli imajo zmanjšano točnost tudi zaradi slabe učinkovitosti pri besedah izven slovarja in narečjih. Modeli so naučeni na določenem naboru podatkov, ki vsebuje omejeno (čeprav veliko) število besed. Če model naleti na neznano besedo (besedo zunaj slovarja), ne more določiti, kako jo transkribirati. Narečje prav tako spremeni zvok besede onkraj prepoznavanja modela. Nazadnje, modeli STT običajno zahtevajo, da raziskovalec določi jezik avdio posnetka. Če je jezik mešan, je težko zanesljivo transkribirati dialog. Na primer, izposojene angleške besede v slovenskem govoru predstavljajo težavo.

Transkripcija za pridobivanje glasbene informacije obsega transkripcijo instrumenta (na podlagi not), transkripcijo petja (na podlagi not) in transkripcijo besedila (na podlagi besedila). Za folkloristiko je avtomatska transkripcija besedila (ALT) najbolj relevantna. Najzgodnejši modeli so lahko identificirali vsako peto zapeto besedo (Mesáros in Virtanen 2010). Čeprav sodobni modeli dosegajo veliko višjo natančnost (Demirel idr. 2020), ALT še vedno dosega slabše rezultate kot standardni modeli STT zaradi pomanjkanja anotiranih baz učnih podatkov in kompleksnih testnih vzorcev, saj (ljudsko) petje v posnetkih vsebuje več hrupa, nejasno tonalnost, zdrs tona in neenakomerno ritmiko.

² Dostopno na: <https://www.slovenscina.eu/razpoznava/nik>.

Za slovenščino portal slovenscina.eu³ našteva več odprtokodnih jezikovnih tehnologij, vključno s STT modelom (v času pisanja ni modela ALT).

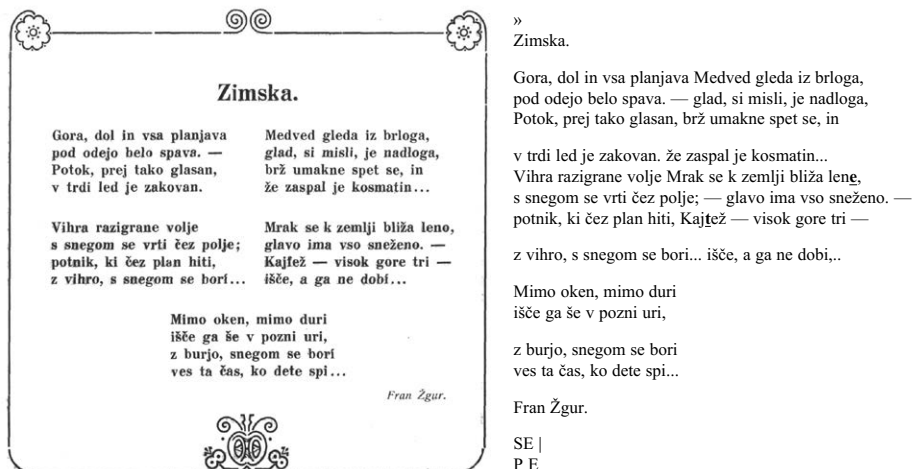


Tabela 1: Primer optičnega prepoznavanja znakov (OCR) z orodjem Tesseract za slovensko pesem (revija Zvonček, 1. 12. 1913). Model ima visoko točnost transkripcije besed, ne upošteva pa dvostolpčne stave besedila. Transkript ima dve črkovni napaki (odebeljeno in podčrtano). Model poskuša transkribirati tudi del besedilnih ornamentov.

Prav tako je mogoče avdio posnetke analizirati neposredno. Avdio posnetek je treba predhodno obdelati z uporabo ločenih tehnik obdelave zvoka. Ena takšna tehnika je segmentacija podatkov, kjer se avdio posnetek razdeli na posamezne enote, kot so govor, solo petje, zborovsko petje in instrumentalna glasba. Segmentacija podatkov omogoča delo s posameznimi avdio enotami in primerjavo le-teh med različnimi posnetki. Obstaja veliko tehnik predobdelave zvoka; opis vseh presega obseg tega članka.

3.2 Digitalizacija tiska

Ko so izvorni podatki že na voljo kot besedilo, a v tiskani obliki, je treba uporabiti drugačno tehnologijo. Natisnjene pravljice, pripovedi, zgodovinska pričevanja in ročno pisani zapiski zahtevajo optično prepoznavanje znakov (OCR) za digitalizacijo. Tehnologija OCR prebere digitalno sliko besedila in sliko pretvori v zaporedje znakov z uporabo modela strojnega učenja (Tabela 1). Najbolj znano odprtokodno OCR orodje je Tesseract (Smith 2007), ki ga je nekaj časa razvijal Google. Kasneje je Google razvil svoje lastno OCR orodje, zapakirano v Google Cloud, medtem ko je Tesseract ostal odprtokoden. ABBYY je še ena priljubljena komercialna rešitev OCR.

³ Dostopno na: <https://slovenscina.eu>.

OCR orodja običajno pričakujejo dobro strukturirano PDF datoteko, vendar pomanjkanje enotnega PDF standarda včasih preprečuje, da bi bili tudi preprosti dokumenti digitalizirani s 100-odstotno natančnostjo. Da bi premagali to težavo, se za določene vrste dokumentov razvijajo specializirani modeli OCR – na primer, za dvostolpčne časopise, mešanico tiskanih in ročno napisanih opomb ali kompleksno postavitvev s številnimi slikami (Berg-Kirkpatrick in Klein 2014).

Ključni korak pri digitalizaciji podatkov je ustrezno organiziranje podatkov in izbira primerne oblike. Avdio, video in slike so shranjeni v binarni obliki,⁴ medtem ko so besedilni podatki običajno kodirani v formatu TEI (TEI Consortium 2023), priznani kodirni shemi v skupnosti digitalnih humanistov. TEI temelji na formatu XML, ki omogoča označevanje delov besedila (Slika 1). Oznake v TEI so vnaprej določene in vključujejo lastnosti, kot so datum, čas, ime, avtor, kraj in naslov. V takšnih podatkih je mogoče programsko iskati dano lastnost, na primer vse kraje.

Medtem ko je TEI predvsem format za shranjevanje podatkov, format CoNLL-U zagotavlja dodatne funkcije za obdelavo naravnega jezika. CoNLL-U besedilo je organizirano tako, da je v vsaki vrstici po ena beseda, ki so ji dodane jezikovne oznake. Predhodno označeni podatki omogočajo specializirano iskanje na podlagi jezikovnih značilnosti, na primer luščenje zgolj lematiziranih samostalnikov.

3.3 Obdelava naravnega jezika

Ko so podatki digitalizirani in pretvorjeni v želeni format, jih je treba pred nadaljnjo analizo predobdelati z uporabo tehnik obdelave naravnega jezika (NLP). Na primer, CoNLL-U že zagotavlja besedilo v visoko predobdelani obliki, zato je ta format še posebej uporaben za nadaljnjo analizo. Sicer je treba uporabiti dodatne tehnike NLP.

NLP predobdelava obsega kombinacijo tokenizacije, lematizacije in oblikoslovno označevanje (POS). Tokenizacija razdeli besedilo na osnovne enote analize, običajno besede. Lematizacija te enote preoblikuje v osnovno obliko, npr. v ednino imenovalnika. Oblikoslovno označevanje dodeli vsaki besedi v danem korpusu oznako besedne vrste na podlagi njene slovnične vloge in konteksta v stavku. Sorodna naloga je prepoznavanje imenskih entitet (NER), ki doda imensko entiteto dani besedi. Na primer, po opisanem postopku bi se stavek »*Jutri grem v Ljubljano*,« pretvoril v »*jutri_Adverb, iti_Verb, v_Adposition, Ljubljana_Noun_loc*«, pri čemer je vsaka beseda pretvorjena v osnovno obliko, pripisana ji je besedna vrsta, beseda »Ljubljana« pa je označena kot kraj.

Alternativa predobdelavi besedila je pretvorba vsake besede v vektorsko predstavitev na podlagi njenega pomena. Ta pristop se imenuje vložitev besed (angl. *embedding*). Pristop vzame vsako besedo iz besedila in ji dodeli vektor na podlagi predhodno naučenega modela vložitev.

⁴ Glej smernice CLARIN: <https://www.clarin.si/repository/xmlui/page/data?locale-attribute=sl>.

Eden prvih lematizatorjev za slovenščino je bil LemmaGen (Juršič idr. 2010), razvit na Inštitutu Jožefa Stefana. Malce kasneje je bilo objavljeno orodje za oblikoslovno označevanje Obelisk (Grčar idr. 2012). Danes so bolj razširjena večnamenska orodja. Cevovod CLASSLA (Ljubešič in Dobrovoljc 2019) podpira tokenizacijo, lematizacijo, oblikoslovno označevanje in prepoznavanje imenskih entitet. Slovenski model Trankit (Krsnik in Dobrovoljc 2023) ponuja podobno funkcionalnost. Za vložitev besed je primerna SloBERTa (Ulčar in Robnik-Šikonja 2021), najnovejši model vložitve besed za standardno slovenščino, in njena mlajša sestra SloBERTa-SIEng za slovenski sleng. **1**

4 Stopnja 3: Analiza in raziskovanje podatkov

Ko so podatki zbrani, organizirani in predobdelani, jih lahko analiziramo. Korpusno jezikoslovje, ki ima analognega predhodnika iz 13. stoletja (McCarthy in O'Keeffe 2010), je danes skoraj popolnoma računalniško. To je študij jezika z uporabo velikih zbirk besedil, imenovanih korpusi (ednina: korpus). Ker so podatki tako obsežni, korpusno jezikoslovje uporablja računalniške in statistične metode za njihovo preučevanje.

Eden prvih korakov pri analizi podatkov je raziskovanje podatkov, ki vključuje opazovanje frekvenc izrazov in meta(podatkov), osnovno statistiko, primerjavo kategorij in podobno. Bodisi za izhodiščno raziskovanje podatkov bodisi za obsežno jezikovno analizo so primerna možnost za začetnike konkordančniki. Konkordančnik je orodje za raziskovanje frekvenc besed in fraz, sopojavitev besed ter iskanje ključnih besed v kontekstu (KWIC). Obstaja veliko programske opreme za lokalno analizo konkordanc, medtem ko postajajo spletni konkordančniki vedno bolj razširjeni. Med komercialnimi orodji je najbolj znan SketchEngine (Kilgarriff idr. 2014), medtem ko je noSketchEngine priljubljena brezplačna različica za upravljanje korpusov.⁵ CLARIN.SI navaja splošno dostopne slovenske konkordančnike.⁶

Naslednji korak za konkordančniki, ki temeljijo na ključnih besedah, je obdelava naravnega jezika in rudarjenje besedil. To vključuje analizo celotnih korpusov in prepoznavanje vzorcev v njih z metodami, ki temeljijo na umetni inteligenci, kot so tematsko modeliranje, gručenje ali klasifikacija dokumentov, analiza sentimenta in podobno.

Med bolj priljubljenimi spletnimi orodji za analizo besedil je Voyant Tools⁷ (Miller 2018), brezplačno orodje za analizo besedil. Ponuja več statističnih vizualizacij, kot so oblaki besed, grafi sopojavitev, spremembe frekvenc besed in analize trendov. Glavna prednost Voyant-a, poleg tega, da je brezplačno dostopna spletna storitev, je njegova enostavnost uporabe, kjer raziskovalec naloži dokumente in interaktivno opazuje statistiko in grafe.

⁵ Seznam konkordančnikov je dostopen na: <https://corpus-analysis.com/tag/concordancer>. Stran navaja tudi druga orodja za analizo korpusov.

⁶ Slovenski konkordančniki: <https://www.clarin.si/info/concordances>.

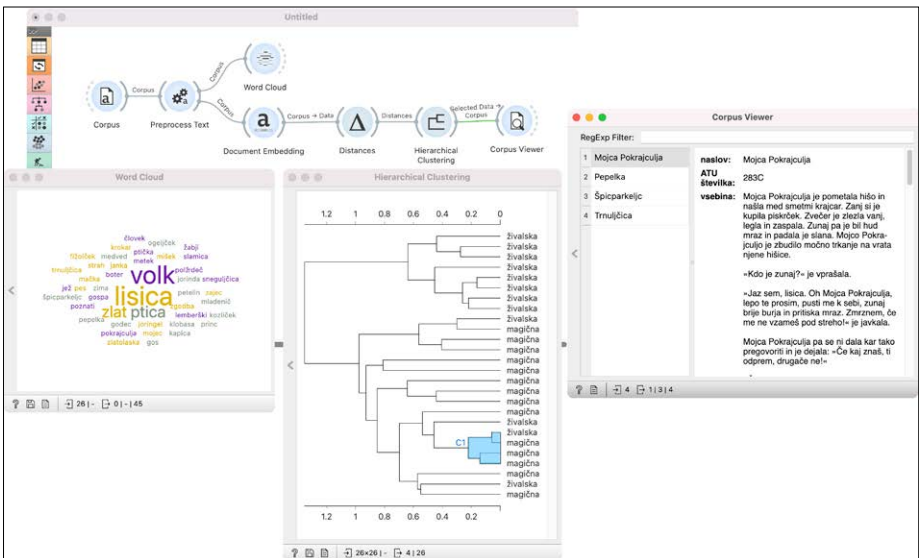
⁷ Dostopno na: <https://voyant-tools.org/>.


```

1 </seg>
2 <bibl corresp="#bib23.1">
3 <date type="reprint" when="1999">1999</date> (<date type="firstEdition" when="1789">1789</date>
4 <biblScope unit="page">113</biblScope>
5 </bibl>
6 <bibl corresp="#bib7.1">
7 <date when="1789">1789</date>
8 <biblScope unit="page">408</biblScope>
9 </bibl>
10 </ab>
11 <ab xml:id="PREG-00-00012" n="13">
12 <seg xml:lang="sl-bohoric" xml:id="PREG-00-00012.dipl" type="dipl">Sarezen kruh fe rad rieshe.</seg>
13 <seg xml:lang="sl" xml:id="PREG-00-00012.crit" type="crit">Sarečen kruh se rad rieže.</seg>
14 <seg xml:id="PREG-00-00012.crit.norm" type="crit.norm" xml:lang="sl">Zarečen kruh se rad reže.</seg>
15 <seg xml:id="PREG-00-00012.crit.norm.ana" type="crit.norm.ana" xml:lang="sl">
16 <s xml:id="PREG-00-00012.crit.s1">
17 <w norm="Zarečen" ana="mie:Apptmsn" msd="UPosTag=ADJCase=NomiDefinite=IndlDegree=PoslGender=MascINumber=SinglVerbForm=Part"
18 lemma="zarečen" xml:id="PREG-00-00012.crit.s1.11">Sarečen</w>
19 <w ana="mie:Ncmsn" msd="UPosTag=NOUNCase=NomiGender=MascINumber=Sing" lemma="kruh" xml:id="PREG-00-00012.crit.s1.12">kruh</w>
20 <w ana="mie:Px-----y" msd="UPosTag=PRONIPronType=PrslReflex=YeslVariant=Short" lemma="se" xml:id="PREG-00-00012.crit.s1.13">se</w>
21 <w ana="mie:fgp" msd="UPosTag=ADVIDegree=Pos" lemma="rad" xml:id="PREG-00-00012.crit.s1.14">rad</w>
22 <w norm="reže" ana="mie:Vmpf3s" msd="UPosTag=VERBIAspect=ImplMood=IndlNumber=SinglPerson=3lTense=PreslVerbForm=Fin"
23 lemma="režati" joins="right" xml:id="PREG-00-00012.crit.s1.15">reže</w>
24 <pc ana="mie:Z" msd="UPosTag=PUNCT" xml:id="PREG-00-00012.crit.s1.16"></pc>
25 <linkGrp type="UD-SYN" targFunc="head argument" corresp="#PREG-00-00012.crit.s1">
26 <link ana="ud-syn.amod" target="#PREG-00-00012.crit.s1.12 #PREG-00-00012.crit.s1.11"/>
27 <link ana="ud-syn.subj" target="#PREG-00-00012.crit.s1.15 #PREG-00-00012.crit.s1.12"/>
28 <link ana="ud-syn.expl" target="#PREG-00-00012.crit.s1.15 #PREG-00-00012.crit.s1.13"/>
29 <link ana="ud-syn.advmod" target="#PREG-00-00012.crit.s1.15 #PREG-00-00012.crit.s1.14"/>
30 <link ana="ud-syn.root" target="#PREG-00-00012.crit.s1 #PREG-00-00012.crit.s1.15"/>
31 <link ana="ud-syn.punct" target="#PREG-00-00012.crit.s1.15 #PREG-00-00012.crit.s1.16"/>
32 </linkGrp>
33 </s>

```

Slika 1: Primer pregovora (Babič 2022), označenega po shemi TEL.



Slika 2: Primer delotoka hierarhičnega gručenja pravljic bratov Grimm v orodju Orange. Delotok prikazuje gručo (C1), v kateri so različne oznake ATU. Moja Pokrajculja je živalska pravljica, ostale pa so magične pravljice. Kljub temu si te pravljice delijo podobne besede.

Za pristop s poudarkom na umetni inteligenci orodja za vizualno programiranje ponujajo nizek vstopni prag za raziskovalce, ki želijo spoznati rudarjenje besedil. Ta orodja namreč vizualno prikažejo delotok, dajejo takojšnje (in vmesne) rezultate, spodbujajo interakcijo in so na splošno zasnovana, da so preprosta. Tak primer je odprtokodno orodje za rudarjenje podatkov Orange⁸ (Demšar idr. 2013), kjer uporabnik sestavlja analitične delotoke s povezovanjem komponent. Modularno zlaganje komponent omogoča prilagodljive, a sposobne delotoke, ki ne zahtevajo programerskih veščin (Slika 2). Rezultati so predstavljeni v interaktivnih vizualizacijah, ki omogočajo raziskovanje podatkov. Orodje ponuja razširitev Text za analizo besedil. Podobna orodja za vizualno programiranje so KNIME, RapidMiner in Weka.

Za analizo folklornih podatkov obstaja veliko tehnik na osnovi umetne inteligence, kot so analiza sentimenta, časovna analiza, razreševanje koreferenc, dodeljevanje razrednih oznak (klasifikacija) ali samodejno luščenje tem (tematsko modeliranje) iz dokumentov. Ker zahtevajo računalniške in znanstvene veščine, lahko navdušen digitalni humanist začne z zmogljivimi konkordančniki in nadaljuje od tam. Projekti digitalnih humanistov so izrazito primerni za interdisciplinarno sodelovanje.

5 Stopnja 5: Ohranjanje podatkov

Zadnja stopnja v življenjskem ciklu podatkov je shranjevanje in ohranjanje podatkov. Dobre prakse deljenja podatkov omogočajo ponovljivost rezultatov in ponovno uporabo podatkov, kar nas pripelje nazaj k prvemu koraku – zbiranju podatkov. Seveda so primerni za deljenje le nekateri raziskovalni podatki, še posebej, če vključujejo občutljive informacije. Kljub temu se večina podatkov v folkloristiki lahko deli v neki obliki.

FAIR (Wilkinson idr. 2016) je najbolj znan in široko sprejet standard za upravljanje podatkov. To je kratica za najdljivost (*findability*), dostopnost (*accessibility*), interoperabilnost (*interoperability*) in ponovno uporabnost (*reusability*). To so štiri splošna načela, ki povzemajo idejo, da mora biti raziskovalne podatke enostavno najti, pridobiti in delati z njimi. To pomeni, da morajo podatki imeti trajne identifikatorje, bogate metapodatke, biti indeksirani v iskalnem viru (najdljivost), jih je mogoče pridobiti z odprtim protokolom (dostopnost), so v splošno sprejetem formatu (interoperabilnost) in imajo podrobne (meta)podatke in jasne, odprte licence (ponovna uporabnost). Uporaba FAIR načel za humanistiko je opisana v Harrower idr. (2020).

FAIR načela so splošno vključena v odprte podatkovne repozitorije. Ti repozitoriji, na primer Danski Folkloristični Makroskop (Tangherlini 2013) ali bolj splošni Virtualni Jezikovni Observatorij, vključujejo strukturirane podatke iz folkloristike.

⁸ Dostopno na: <https://orangedatamining.com/>.

Slovenska nacionalna veja evropske infrastrukture CLARIN, CLARIN.SI⁹, omogoča dostop do različnih jezikovnih virov predvsem (a ne izključno) v južnoslovanskih jezikih. Vsi podatki niso v besedilni obliki – podatki iz folkloristike vključujejo tudi avdio (Matsuura idr. 2020) in video datoteke¹⁰ (Nonhebel idr. 2004), medtem ko so digitalni podatki kulturne dediščine dostopni na spletnem portalu Europeana (Purday 2009). Za shranjevanje podatkov je primeren Zenodo, multidisciplinarni repozitorij odprtih raziskovalnih podatkov, ki podpira samoarhiviranje in splošno upravljanje in shranjevanje podatkov (za pregled glej Peters idr. 2017).

6 Zaključek

Seth Long trdi, da podatki v digitalni humanistiki kličejo po ponovni uporabi (Long in Williams 2014). Obstoječi korpusi in podatkovni nizi se lahko uporabijo za primerjalno analizo, analizo z modernimi pristopi, dopolnijo z dodatnim terenskim delom ali pa se obravnavajo na nove načine. Spletni podatki lahko usmerjajo raziskave z ustvarjanjem raziskovalnih vprašanj, dopolnjevanjem terenskih podatkov ali kot samostojen raziskovalni material.

Kje naj torej začne navdušeni digitalni humanist ali računalniški folklorist? Najprej s podatki. Digitalizacija podatkov iz folkloristike je ključna za zagotavljanje boljše kakovosti podatkov, kopiranje in razmnoževanje brez izgube kakovosti, lažje iskanje prek indeksiranja ter lažji dostop do podatkov za primerjalne raziskave in ponovno uporabo podatkov. Mnogi raziskovalci si prizadevajo za zagotavljanje široko dostopnih podatkov iz folkloristike, nekatera od teh prizadevanj so privedla do zglednih repozitorijev. Nekatere institucije s celovitimi spletnimi aplikacijami, ki omogočajo indeksirano iskanje, časovno usklajeno transkripcijo in vizualizacijo (Boyd 2019), ponujajo več kot zgolj običajen digitalni dostop. To je tisto, kar Timothy Tangherlini imenuje »folklorni makroskop«, tj. sistem, ki omogoča celostno uporabo institucionalnih zbirk (Tangherlini 2013). Primer takšnega makroskopa je Sinhronizator Metapodatkov Ustnih Zgodovin, ki poravnava avdio datoteke s časovnimi žigi transkriptov in izboljšuje segmente z metapodatki (tema, ključne besede, povzetek) (Boyd 2019).

Digitalni podatki zahtevajo digitalno usposobljeno osebje in ustrezne zmogljivosti za shranjevanje podatkov, kar ne sme biti samoumevno (glej Mosweu 2011). Usposabljanje na področju korpusnega jezikoslovja, rudarjenja besedil in obdelave naravnega jezika je osnova za napredno besedilno raziskovanje. Vendar to ne vključuje nujno učenja programiranja, kot je razvidno v razdelku 4. Mnoga orodja imajo nizko vstopno

⁹ Dostopno na: <https://www.clarin.si>.

¹⁰ Video datoteke se tipično uporablja za znakovni jezik.

raven, kjer strokovnjaki za folkloristiko lahko uporabijo svoje široko domensko znanje za raziskovanje podatkov na nove načine.

Z obilico digitalnih orodij, ki so na razpolago raziskovalcu, se pojavi vprašanje, kaj potem ostane vlogi raziskovalca? Pazljiv bralec bo opazil, da v članku manjka ena stopnja cikla podatkov; stopnja 4: interpretacija podatkov. Čeprav so računalniki in sodobne tehnike na osnovi umetne inteligence odlični pri organizaciji in čiščenju podatkov, so daleč od človeške sposobnosti interpretacije podatkov. Raziskovalec razume širši kontekst raziskovalnega problema, zgodovinske temelje, razmerja moči, socialno hierarhijo in jezikovne posebnosti. Sinteza izhaja iz združevanja obstoječega znanja z novimi vpogledi. Digitalna orodja so ravno to – orodja, ki nam pomagajo prepoznati zanimive vzorce ali odstopanja. Identifikacija vzročnosti, razlaga ugotovitev in določanje pomena ostaja v celoti v človeški domeni.

Zahvala

Članek je nastal v okviru financiranja Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije P6-0436: *Digitalna Humanistika: viri, orodja in metode (2022–2027)*.

Viri in literatura

- Ang, Chee Siang; Bobrowicz, Ania; Schiano, Diane J.; Nardi, Bonnie, 2013: Data in the Wild: Some Reflections. *Interactions* 20/2, 39–43. DOI: <https://doi.org/10.1145/2427076.2427085>.
- Babič, Saša; idr., 2022: Collection of Slovenian paremiological units Pregovori 1.0, Slovenian language resource repository CLARIN.SI. Na spletu: <http://hdl.handle.net/11356/1455> (dostop 13. 2. 2024).
- Berg-Kirkpatrick, Taylor; Klein, Dan, 2014: Improved Typesetting Models for Historical OCR, V: *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. Baltimore: ACL, 118–23.
- Boyd, Doug, 2019: OHMS: Enhancing Access to Oral History for Free. *The Oral History Review* 40/1, 95–106. DOI: <https://doi.org/10.1093/ohr/oht031>.
- Chiaro, Delia, 2018: *The language of jokes in the digital age: Viral humour*. Oxon: Routledge.
- Demirel, Emir; Ahlbäck, Sven; Dixon, Simon, 2020: Automatic Lyrics Transcription Using Dilated Convolutional Neural Networks with Self-Attention, V: *International Joint Conference on Neural Networks*. Glasgow: IEEE, 1–8.
- Demšar, Janez; Curk, Tomaž; Erjavec, Aleš idr., 2013: Orange: Data Mining Toolbox in Python. *Journal of Machine Learning Research* 14/avg, 2349–2353.
- Eklund, Johan; Hagedorn, Josh; Darányi, Sándor, 2023: Teaching Tale Types to a Computer: A First Experiment with the Annotated Folktales Collection. *Fabula* 64/1–2, 92–106. DOI: <https://doi.org/10.1515/fabula-2023-0005>.
- Fišer, Darja; Ljubešić Nikola; Erjavec Tomaž, 2018: The Janes project: language resources and tools for Slovene user generated content. *Language Resources & Evaluation*. DOI: <https://doi.org/10.1007/s10579-018-9425-z>.

- Grčar, Miha; Krek, Simon; Dobrovoljc, Kaja, 2012: Obeliks: statistični oblikoskladenjski označevalnik in lematizator za slovenski jezik (Obeliks: statistical morphosyntactic tagger and lemmatizer for Slovene). V: *Proceedings of the 8th Language Technologies Conference C*. Ljubljana: IJS, 89–94.
- Halford, Susan, 2017: The Ethical Disruptions of Social Media Data: Tales from the Field, V: Woodfield, Kandy (ur.), *Ethics of Internet-mediated Research and Using Social Media for Social Research*. Bristol: Emerald, 13–25.
- Harrower, Natalie; Immenhauser, Beat; Lauer, Gehrard; Maryl, Maciej; Orlandi, Tito; Rentier, Bernard; Wandl-Vogt, Eveline, 2020: *Sustainable and FAIR Data Sharing in the Humanities*. Berlin, Germany: ALLEA - All European Academies. DOI: <https://doi.org/10.7486/DRI.tq582c863>.
- Juršič, Matjaž; Mozetič, Igor; Erjavec, Tomaž; Lavrač, Nada, 2010: Lemmagen: Multilingual lemmatisation with induced ripple-down rules. *Journal of Universal Computer Science* 16/9, 1190–1214. DOI: <https://doi.org/10.3217/jucs-016-09-1190>.
- Karsdorp, Folgert; van der Meulen, Marten; Meder, Theo; van den Bosch, Antal, 2015: MOMFER: A Search Engine of Thompson's Motif-Index of Folk Literature. *Folklore* 126/1, 37–52. DOI: <https://doi.org/10.1080/0015587X.2015.1006954>.
- Kilgarrieff, Adam; Baisa, Vít; Bušta, Jan. idr., 2014: The Sketch Engine: ten years on. *Lexicography ASIALEX* 1, 7–36. DOI: <https://doi.org/10.1007/s40607-014-0009-9>.
- Kingsland, Kaitlyn, 2020: Comparative Analysis of Digital Photogrammetry Software for Cultural Heritage. *Digital Applications in Archaeology and Cultural Heritage* 18. DOI: <https://doi.org/10.1016/j.daach.2020.e00157>.
- Krawczyk-Wasilewska, Violetta, 2016: *Folklore in the Digital Age: Collected Essays*. Łódź: Jagiellonian University Press.
- Krsnik, Luka; Dobrovoljc, Kaja, 2023: The Trankit model for linguistic processing of standard Slovenian, Slovenian language resource repository CLARIN.SI. Na spletu: <http://hdl.handle.net/11356/1870> (dostop 16. 2. 2024).
- Kung, Susan S., 2022: Developing a data management plan. V: Berez-Kroeker, Andrea L.; McDonnell, Bradley; Koller, Eve; Collister, Lauren B. (ur.), *The Open Handbook of Linguistic Data Management*. Cambridge, Massachusetts: The MIT Press, 101–115.
- Lebar Bajec, Iztok; Bajec, Marko; Bajec, Žan; Rizvič, Mitja, 2022: Slovene Conformer CTC BPE E2E Automated Speech Recognition model RSDO-DS2-ASR-E2E 2.0, Slovenian language resource repository CLARIN.SI. Na spletu: <http://hdl.handle.net/11356/1737> (dostop 13. 2. 2024).
- Lee, Benjamin Charles Germain, 2022: The Digital Humanities and the Ladino Press: Using Machine Learning to Extract and Analyze Visual Content in Historic Ladino Newspapers. *Studies in Digital History and Hermeneutics*, 191. DOI: <https://doi.org/10.1515/9783110744828-010>.
- Ljubešić, Nikola; Dobrovoljc, Kaja, 2019: What does Neural Bring? Analysing Improvements in Morphosyntactic Annotation and Lemmatisation of Slovenian, Croatian and Serbian. V: *Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing*. Florence: Association for Computational Linguistics, 29–34.
- Long, Seth; Williams, Sierra, 2014: “Re-Purposing” Data in the Digital Humanities: Data Beg to Be Taken from One Context and Transferred to Another. Na spletu: <https://blogs.lse.ac.uk/impactofsocialsciences/2014/04/02/re-purposing-data-in-the-digital-humanities/> (dostop 2. 4. 2014).

- Lourdi, Irene; Papatheodorou, Christos; Nikolaidou, Mara, 2007: A multi-layer metadata schema for digital folklore collections. *Journal of Information Science* 33/2, 197–213. DOI: <https://doi.org/10.1177/0165551506070711>.
- Marolt, Matija; Bohak, Ciril; Kavčič, Alenka; Pesek, Matevž, 2019: Automatic Segmentation of Ethnomusicological Field Recordings. *Applied Sciences* 9/3, 439. DOI: <https://doi.org/10.3390/app9030439>.
- Matsuura, Kohei; Ueno, Sei; Mimura, Masato; Sakai, Shinsuke; Kawahara, Tatsuya, 2020: Speech Corpus of Ainu Folklore and End-to-End Speech Recognition for Ainu Language, V: *Proceedings of the Twelfth Language Resources and Evaluation Conference*. Marseille: European Language Resources Association, 2622–28.
- McCarthy, Michael; O’Keeffe, Anne, 2010: Historical perspective: What are corpora and how have they evolved? V: O’Keeffe, Anne; McCarthy, Michael (ur.), *The Routledge Handbook of Corpus Linguistics*. London: Routledge, 3–13.
- Meder, Theo; Himstedt-Vaid, Petra; Meyer, Holger, 2023: Collecting International Narrative Heritage in a Multilingual Search Engine. *Fabula* 64/1–2, 107–27. DOI: <https://doi.org/10.1515/fabula-2023-0006>.
- Mesaros, Annamaria; Virtanen, Tuomas, 2010: Automatic Recognition of Lyrics in Singing. *EURASIP Journal on Audio, Speech, and Music Processing* 2010/1. DOI: <http://dx.doi.org/10.1155/2010/546047>.
- Miller, Alissa, 2018: Text Mining Digital Humanities Projects: Assessing Content Analysis Capabilities of Voyant Tools. *Journal of Web Librarianship* 12/3, 169–197. DOI: <https://doi.org/10.1080/19322909.2018.1479673>.
- Mosweu, Tshepho, 2011: Digitising the Oral History Collection at Botswana National Archives and Records Services: Problems and Prospects. *Journal of the South African Society of Archivists* 44, 124–30.
- Nonhebel, Annika; Crasborn, Onno; van der Kooij, Els, 2004: Sign Language Transcription Conventions for the ECHO Project. *Version* 9, 20.
- Park, Tae Jin; Kanda, Naoyuki; Dimitriadis, Dimitrios; Han, Kyu J.; Watanabe, Shinji; Narayanan, Shrikanth, 2022: A Review of Speaker Diarization: Recent Advances with Deep Learning. *Computer Speech & Language* 72. DOI: <https://doi.org/10.1016/j.csl.2021.101317>.
- Peters, Isabella; Kraker, Peter; Lex, Elisabeth; Gumpenberger, Christian; Gorraiz, Juan Ignacio, 2017: Zenodo in the Spotlight of Traditional and New Metrics. *Frontiers in Research Metrics and Analytics* 2. DOI: <http://dx.doi.org/10.3389/frma.2017.00013>.
- Povey, Daniel; Ghoshal, Arnab; Boulianne, Gilles; Burget, Lukas; Glembek, Ondrej; Goel, Nagendra; Hannemann, Mirko; idr., 2011, The Kaldi Speech Recognition Toolkit, V: *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding*. Hawaii: IEEE Signal Processing Society.
- Purday, Jon, 2009: Think Culture: Europeana. Eu from Concept to Construction. *The Electronic Library* 27/6, 919–937. DOI: <https://doi.org/10.1108/02640470911004039>.
- Radford, Alec; Kim, Jong Wook; Xu, Tao; Brockman, Greg; Mcleavey, Christine; Sutskever, Ilya, 2023: Robust Speech Recognition via Large-Scale Weak Supervision, V: Krause, Andreas; Brunskill, Emma; Cho, Kyunghyun; Engelhardt, Barbara; Sabato, Sivan; Scarlett, Jonathan (ur.), *Proceedings of the 40th International Conference on Machine Learning*. 28492–518.
- Shifman, Limor, 2013: *Memes in digital culture*. Cambridge, Massachusetts: MIT press.
- Smith, Ray, 2007: An Overview of the Tesseract OCR Engine, V: *Ninth International Conference on Document Analysis and Recognition (ICDAR 2007)*. Curitiba: IEEE, 629–633.

- Smits, Thomas; Wevers, Melvin, 2023: A Multimodal Turn in Digital Humanities: Using Contrastive Machine Learning Models to Explore, Enrich, and Analyze Digital Visual Historical Collections. *Digital Scholarship in the Humanities* 38/3, 1267–1280. DOI: <https://doi.org/10.1093/lc/fqad008>.
- Tangherlini, Timothy R., 2013: The Folklore Macroscopic: Challenges for a Computational Folkloristics. *Western Folklore*, 7–27.
- TEI Consortium. 2023: TEI P5: Guidelines for Electronic Text Encoding and Interchange. Na spletu: <https://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html> (dostop 16. 11. 2023).
- Thuillard, Marc; Le Quellec, Jean-Loïc; d'Huy, Julien, 2018: Computational Approaches to Myths Analysis: Application to the Cosmic Hunt. *Nouvelle Mythologie Comparée/New Comparative Mythology* 4, 123–154.
- Ulčar, Matej; Robnik-Šikonja, Marko, 2021: SloBERTa: Slovene monolingual large pretrained masked language model. V: *Proceedings of SI-KDD within the Information Society 2021*. 17–20.
- Van Uytvanck, Dieter; Stehouwer, Herman; Lampen, Lari, 2012: Semantic Metadata Mapping in Practice: The Virtual Language Observatory, V: Calzolari, Nicoletta; Choukri, Khalid; Declerck, Thierry, idr. (ur.), *LREC 2012: 8th International Conference on Language Resources and Evaluation*. Istanbul: European Language Resources Association (ELRA), 1029–34.
- Verdonik, Darinka; Zwitter Vitez, Ana, 2011: Slovenski govorni korpus Gos. Ljubljana: Trojina, zavod za uporabo slovenistiko.
- Wiering, Frans; Veltkamp, Remco C.; Garbers, Jörg; Volk, Anja; van Kranenburg, Peter; Grijp, Louis P., 2009: Modelling Folksong Melodies. *Interdisciplinary Science Reviews* 34/2–3, 154–71. DOI: <https://doi.org/10.1179/174327909X441081>.
- Wilkinson, Mark; Dumontier, Michel; Aalbersberg, IJsbrand J. idr., 2016: The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3. DOI: <https://doi.org/10.1038/sdata.2016.18>.
- Yogarajah, Yathukulan, 2022: 'Hodling' on: Memetic Storytelling and Digital Folklore within a Cryptocurrency World. *Economy and Society* 51/3, 467–88. DOI: <https://doi.org/10.1080/03085147.2022.2091316>.



The Collection of Slovenian paremiological units Pregovori: Encoding, publication and quantitative overview

Zbirka slovenskih paremioloških enot Pregovori: kodiranje, objava in kvantitativni pregled

The article discusses the digital collection of paremiological units called Pregovori (Proverbs), the first large and openly available dataset of this type for the Slovenian language. We discuss its encoding in TEI XML, including linguistic annotation, and elaborate on its availability in the repository of the Slovenian research infrastructure CLARIN.SI, along with its integration into concordancers that enable on-line analyses of the Collection. We then give a short quantitative overview of the Collection, followed by the use of its constituent paremiological units in metaFida, the largest corpus of the Slovenian language. Finally, we present our conclusions and give directions for further work.

• **Keywords:** proverbs, digital database, text encoding initiative, language corpus, linguistic annotation, concordancers, open access

Članek obravnava digitalno zbirko paremioloških enot Pregovori, ki je prva obsežna in javno dostopna zbirka tovrstnih podatkov za slovenski jezik. Obravnavamo njeno kodiranje v TEI XML, vključno z jezikovno anotacijo, in podrobneje predstavljamo njeno dostopnost v repozitoriju slovenske raziskovalne infrastrukture CLARIN.SI ter njeno vključitev v konkordančnike, ki omogočajo spletne analize zbirke. Nato podamo kratek kvantitativni pregled Zbirke, ki mu sledi uporaba njenih sestavnih paremioloških enot v metaFidi, največjem korpusu slovenskega jezika. Na koncu predstavimo svoje zaključke in podamo usmeritve za nadaljnje delo.

• **Ključne besede:** pregovori, digitalna podatkovna zbirka, pobuda za kodiranje besedil, jezikovni korpus, jezikovna anotacija, konkordančniki, odprti dostop

1 Introduction

The Open Science paradigm, which advocates for the open availability of research publications and datasets, is increasingly gaining ground, even in the humanities. In this paper, we demonstrate this principle on a large dataset of Slovenian paremiological units and show how it is structured, encoded and made available for download and on-line analysis within the framework of the Slovenian research infrastructure for language resources and technologies CLARIN.SI.

The collection of paremiological units (henceforth the Collection) was initially prepared at the Institute of Slovenian Ethnology at the Scientific and Research Centre of the Slovenian Academy of Sciences and Arts. This institute serves as an archive for ethnological materials and saw a large augmentation of paremiological units during the 1990s and later years, thanks to a series of dedicated projects (Stanonik 1996; 2004; 2009; 2015). The Collection was digitised here, first by manual input into Word and later transferred to Excel.

The process of cleaning up the Excel spreadsheets and converting them to XML has already been detailed in Babič and Erjavec (2022). Here, we only mention that the input for this conversion consisted of two TSV (tab-separated values) tabular files exported from Excel. One file contained the paremiological units and the identifiers (IDs) of the bibliographic sources they were observed in (along with the page number of the mention of the paremiological unit in the source). The second file contained the list of these bibliographic units with their respective IDs. The basic conversion to XML was then relatively straightforward, particularly after the XML encoding had been defined. The XML encoded Collection is the point of departure for this paper. It should also be mentioned that the digitised collection has been published in two versions to date: the “Collection of Slovenian paremiological units Pregovori 1.0” (Babič et al. 2022), and the “Collection of Slovenian paremiological units Pregovori 1.1” (Babič et al. 2023), which corrected some errors from version 1.0 and also slightly extended its scope. In this paper we provide an overview of version 1.1.

The rest of this paper is structured as follows: Section 2 discusses the encoding of the Collection, including its metadata, the bibliographical sources and the paremiological units, as well as the automatic linguistic analysis of the units. Section 3 overviews the publication of the Collection, detailing its availability for download from the CLARIN.SI repository and access through the CLARIN.SI concordancers. Section 4 gives a quantitative overview of the Collection, first in terms of its size and other characteristics, and then in terms of the usage of its paremiological units in a large Slovenian language corpus. Finally, Section 5 gives the conclusions and some directions for further work.

2 Encoding of the Collection

The Text Encoding Initiative (TEI) Guidelines (TEI Consortium, 2020) is a robust and comprehensive framework for encoding and analysing textual materials. TEI originated in the mid-1980s when the academic community recognised the need for a standardised approach to representing and exchanging textual information in the burgeoning field of digital humanities. At the core of the TEI mission was the development of guidelines to enable scholars, librarians and archivists to create machine-readable texts capable of capturing the richness and complexity of human expression.

The primary goal of the TEI is to provide a set of guidelines and standards for encoding texts in a way that is both human-readable and machine-readable. TEI markup, expressed in XML (eXtensible Markup Language), allows for the identification and description of various elements within a text. This granular encoding facilitates sophisticated analysis, searching and presentation of textual content, transcending the limitations of traditional print media. One of the notable features of TEI is its adaptability. The guidelines are designed to accommodate a diverse range of texts, from

literary works and historical documents to linguistic corpora and multimedia resources. This flexibility has made TEI a valuable tool for scholars across different disciplines, fostering interdisciplinary research and collaboration.

TEI has had a profound impact on the field of digital humanities. Scholars and institutions around the world have adopted TEI Guidelines as a standard for encoding and exchanging textual data. This widespread acceptance has facilitated the creation of digital archives, electronic editions of texts and innovative research projects that leverage the capabilities of TEI-encoded materials. Moreover, TEI has contributed to the democratisation of access to cultural heritage. By encoding and digitising historical documents, libraries, museums and archives can make their collections available to a global audience, fostering a new era of scholarship and exploration. TEI-encoded texts also enable advanced computational analysis, opening doors to new research methodologies and insights.

For these reasons, the Collection is encoded in TEI. In particular, it is modelled as one TEI document (XML element *TEI*), which then contains the TEI header (element *teiHeader*) giving the metadata (including the bibliographical sources) of the Collection, followed by the text (element *text*), i.e. the paremiological units constituting the Collection. The next subsection explains and illustrates the encoding of the TEI header, while the following subsection delves into the encoding of paremiological units. We conclude the Section by explaining the automatic linguistic processing of the Collection and how it is encoded in TEI.

2.1 The TEI header

We illustrate the start of the document in Figure 1, giving the overall *TEI* root XML tag with the attributes for the TEI namespace, and the ID and language of the resource. The TEI header is the first element inside *TEI*, and it is illustrated next, specifically the file description (*fileDesc*), which encompasses the title statement (*titleStmnt*) containing the basic bibliographic metadata of the document.

The TEI header contains other substantial metadata. Here, we highlight only two notable elements of the Collection's metadata. The first is the taxonomy of bibliographic units that were the source of the proverbs. The bibliographic units were classified into 18 types (the complete list is given in Table 1) and these types are encoded in the TEI header class declarations (*classDecl*), specifically within the *taxonomy* element. The encoding of the first six types is given in Figure 2.

As can be seen, each category is assigned an ID and then given the Slovenian description of the category in the category description (*catDesc*) element.

The second element illustrated is the source description (*sourceDesc*), specifically its list of bibliographic sources (*listBibl*), with each source encoded in the bibliography element (*bibl*), as illustrated for a few types of sources in Figure 3.

```

<TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="sl" xml:id="pregovori">
  <teiHeader>
    <fileDesc>
      <titleStm>
        <title type="main" xml:lang="sl">Zbirka slovenskih pregovorov</title>
        <title type="main" xml:lang="en">Collection of Slovenian proverbs</title>
      <principal>
        <persName>Saša Babič</persName>
      </principal>
      <respStm>
        <persName>Marija Stanonik</persName>
        <resp xml:lang="sl">Zbiranje gradiva</resp>
        <resp xml:lang="en">Collection of the data</resp>
      </respStm>
      <respStm>
        <persName>Saša Babič</persName>
        <persName>Miha Pečec</persName>
        <resp xml:lang="sl">Čiščenje in pretvorba izvirnega gradiva v enotni zapis</resp>
        <resp xml:lang="en">Clean-up and conversion of source data to a common
          format</resp>
      </respStm>
    </fileDesc>
  </teiHeader>

```

Figure 1: The beginning of the Collection TEI document.

```

<classDecl>
  <taxonomy xml:id="bibls">
    <desc xml:lang="sl"><term>Vrste virov</term></desc>
    <category xml:id="bibl.fiction">
      <catDesc xml:lang="sl">Leposlovje in literarjenje</catDesc>
    </category>
    <category xml:id="bibl.museum">
      <catDesc xml:lang="sl">Muzejske zbirke</catDesc>
    </category>
    <category xml:id="bibl.random">
      <catDesc xml:lang="sl">Naključni viri</catDesc>
    </category>
    <category xml:id="bibl.folklore">
      <catDesc xml:lang="sl">Objavljena slovstvena folklor</catDesc>
    </category>
    <category xml:id="bibl.news">
      <catDesc xml:lang="sl">Periodika - časopisi, revije in glasila</catDesc>
    </category>
    <category xml:id="bibl.yearbook">
      <catDesc xml:lang="sl">Periodika - letopisi</catDesc>
    </category>
  </taxonomy>

```

Figure 2: The taxonomy of bibliographical types.

```

<listBibl xml:lang="sl">
  <bibl xml:id="bibl.1" n="15" corresp="#bibl.dictionary">Bohorič, Adam,
  <date type="reprint" when="1970">1970</date>
  (<date type="firstEdition" when="1584">1584</date>): Arcticae
  horulae succisivae. Faksimile. Mladinska knjiga. Ljubljana.</bibl>
  ...
  <bibl xml:id="bib57.28" n="666/92" corresp="#bibl.news">Kmetijske in
  rokodelske novice. 21. 10. (<date when="1846">1846</date>).
  Št. 42. Letn. 4. Jožef Blaznik. Ljubljana.</bibl>
  ...
  <bibl xml:id="bib242.2" n="90/2" corresp="#bibl.fiction">Kovič, Kajetan,
  (<date when="2005">2005</date>): Jutranji sprehajalec: pripoved o hoji.
  Študentska založba. Ljubljana.</bibl>
  ...
  <bibl xml:id="bib511.71" n="392" corresp="#bibl.oral">Ivančič, Jožef. Čezsoča.
  (<date when="2001">2001</date>).</bibl>
  ...
</listBibl>

```

Figure 3: Examples of bibliographic sources for the proverbs.

Each bibliographic element is assigned an ID, with the generic label (*n*) encoding the element's identifier in the physical collection at the Institute of Slovenian Ethnology. The elements also contain the correspondence (*corresp*) pointer to the categories of the previously mentioned taxonomy, effectively classifying the bibliographic source into one of the 18 defined types.

The content of the bibliographical units is semi-structured, i.e. they contain the text describing the unit, with the date(s) of publication marked up (*date*). The year attribute (*when*) specifies either the publication year of the source or, in the case of serial publication such as “1680–1685”, denotes the start and end dates of its publications (attributes *from* and *to*). Assigning temporal information to controlled attribute values enables automatic processing. For example, it enables the selective retrieval of paremiological units gathered from sources within a specific range of years. If the source is a reprint, this date is distinguished from that of the first edition of the work by the value of the *type* attribute.

TEI allows much more detailed and structured bibliographic descriptions. However, authors and titles are given in a variety of ways, and encoding these additional distinctions would require the extensive list of sources to be marked up manually.

2.2 Markup of the proverbs

In the *text* element of the TEI document, the individual proverbs are marked up as anonymous blocks (*ab*), where each paremiological unit is presented both in the orthography of the source (referred to as the diplomatic transcription), and modernised (standardised) in terms of its alphabet (critical transcription). Each unit is accompanied by bibliographic sources, as illustrated in Figure 4.

The two transcriptions are encoded as segments (*seg*) of two types. Each segment is also marked with its language, where *sl* is the ISO two-letter code for (contemporary) Slovenian, while *sl-bohoric* is the IANA code for the Bohorič alphabet,¹ which was employed until about the 1850s.

```
<ab xml:id="PREG-00-00001" n="1">
  <seg xml:lang="sl-bohoric" xml:id="PREG-00-00001.dipl" type="dipl">Bres muje fe zhreul ne obuje.</seg>
  <seg xml:lang="sl" xml:id="PREG-00-00001.crit" type="crit">Brez muje se čreul ne obuje.</seg>
  <bibl corresp="#bib14.1">
    <date when="1833">1833</date>
    <biblScope unit="page">202</biblScope>
  </bibl>
  <bibl corresp="#bib23.1">
    <date type="reprint" when="1999">1999</date> (<date type="firstEdition" when="1789">1789</date>
    <biblScope unit="page">51</biblScope>
  </bibl>
  <bibl corresp="#bib7.1">
    <date when="1789">1789</date>
    <biblScope unit="page">524</biblScope>
  </bibl>
</ab>
```

Figure 4: Examples of bibliographic sources for the proverbs.

¹ <https://www.iana.org/assignments/lang-subtags-templates/bohoric.txt>

Short editorial notes from the source materials have also been preserved in the diplomatic transcription. They usually give the time of year or date, for example, for saints' days:

(1) *Pred svetim Jakobom<note>25. julij</note> tri dni lepó, rž prav redno dozorela bo.*

(“*Before St. Jakob’s<note>July 25th </note> three days nice, rye ripens in a trice.*”).

The bibliographic elements (*bibl*) point to (*corresp* attribute) the bibliographic sources of the units (*sourceDesc/listBibl/bibl*). The page(s) where the unit can be found in the source are given in the scope of bibliographic reference (*biblScope*). The date(s) of publication are also given for convenience, even though they are redundant, as they are also given in the bibliographical source in the TEI header.

2.3 Linguistic annotation

While encoded in TEI, the information presented so far is essentially the same as in the two tables of bibliographical sources and paremiological units, which served as the basis for the TEI conversion. The Collection is also distributed in this form, for those only interested in the plain text of the paremiological units. However, another variant of the Collection has also been prepared. It is largely identical to the “plain text” version but has added segments that contain the automatically linguistically annotated text of the critical transcriptions. This version, with its added linguistic annotation, enables better searching through the collection (e.g. by lemma rather than word form), as well as searching via abstract linguistic categories, such as the part-of-speech or syntactic relations.

In the first stage of the linguistic processing, the critical transcriptions of the proverbs were tokenised, i.e. split into words and punctuation symbols. In cases where the proverb comprised more than one sentence they were also segmented into sentences. The second stage involved modernising the identified words. The reason for this is two-fold. Firstly, words used to be written differently in historical Slovenian, differing not only from the way they are written today but the spelling also depended on the author and the exact time period, as it had not yet been standardised. This variation means it would otherwise be much more difficult to search for specific words in the Collection, so they have been modernised and their spelling unified. Secondly, the models used for further linguistic analysis have been trained on datasets of contemporary standard Slovenian and work much better on modernised words than when applied to words with archaic spellings.

For modernising the words, we used the open-source normalisation tool² cSMTiser (Scherrer and Ljubešić 2016), which is based on the Statistical Machine Translation (SMT) tool Moses (Koehn 2010). However, while Moses and other SMT tools translate words in a sentence, cSMTiser translates individual letters (characters) in a word. The

² <https://github.com/clarinsi/csmtiser>

tool was trained to carry out modernisation using the manually modernised corpus of historical Slovenian goo300k (Erjavec 2015). This is a process that is similar to what was previously undertaken for modernising the words in the collection of historical Slovenian novels as part of the ELTeC corpus (Schöch et al. 2021). The tool was then used to modernise the word tokens in the critical transcriptions. It should be noted that like other tools based on machine learning principles, cSMTiser knows how to handle unknown words (i.e. words not present in the training data). However, it also makes mistakes and does not always correctly translate archaic words into their modern-day standard spelling.

Once the words had been normalised, the resulting sentences were further annotated with the open-source tool-chain CLASSLA-Stanza³ (Ljubešić and Dobrovoljc 2019; Terčon and Ljubešić 2023), which adds the following information to each sentence token:

- ♦ The lemma or base form of the word, an essential piece of information that simplifies searching for the highly inflective Slovenian language.
- ♦ The morphosyntactic description (MSD) according to the MULTEXT-East specifications⁴ (Erjavec 2012). For example, *Ncmsg* for the morphosyntactic features *Noun*, *Type = common*, *Gender = masculine*, *Number = singular*, *Case = genitive*. It should be noted that conversion tables exist for mapping from MSDs to their features, as well as to MSDs and their features in Slovenian. For example, the Slovenian equivalent of *Ncmsg* is the MSD *Somer* and the features *samostalnik*, *vrsta = občno_ime*, *spol = moški*, *število = ednina*, *sklon = rodilnik*.
- ♦ The morphological features according to the Universal Dependencies framework (de Marneffe et al. 2021)⁵ for Slovenian (Dobrovoljc et al. 2017), e.g. *NOUN*, *Case = Gen*, *Gender = Masc*, *Number = Sing*. These features are similar to those of MULTEXT-East, however, the features have different names and there are sometimes some systematic differences. The reason why both are present is that the MULTEXT-East MSDs have a long tradition of use in Slovenian corpora, while the Universal Dependencies are now becoming a de-facto standard for many languages including Slovenian.
- ♦ The syntactic parse of the sentence, also according to the Universal Dependencies formalism. The parse is dependency-based, i.e. head/argument pairs of sentence tokens are linked and the link is associated with its dependency label.

Figure 5 illustrates the added linguistic information for one paremiological unit, which is encoded in the segment typed as being the normalised form of the critical

³ <https://github.com/clarinsi/classla>

⁴ <https://github.com/clarinsi/mte-msd>, <http://nl.ijs.si/ME/V6/msd/html/msd-sl.html>

⁵ <https://universaldependencies.org/>

transcription (*type="crit.norm"*). As can be noted, all the words in this example are identical to the ones in the critical transcription, which are indeed spelled the same in the contemporary standard, except for “čevlj” as the modernised form of “čreul”. This modernisation illustrates the already mentioned fact that cSMTiser also makes mistakes, as the correct form should in fact be “čevelj”. Nevertheless, the automatically modernised form is closer to the standard than the one in the critical transcription.

The second added segment (only the beginning of which is shown to save space) is presented in its normalised and linguistically analysed critical transcription (*type = "crit.norm.ana"*). This segment contains one or more sentences (element *s*), each of which then contains tokens – either words (*w*) or punctuation symbols (*pc*, not shown here). The words themselves are presented as they appear in the critical transcription. Any variations in their normalised form are indicated by the value of the normalised form (*norm*) attribute. The MULTEXT-East MSD is given as the value of the analysis (*ana*) attribute, serving as a pointer to the decomposition of the MSD into features. The Universal Dependencies morphological features are given as the value of the MSD (*msd*) attribute, and the lemma as the value of the *lemma* attribute. Each token also has an ID, which is necessary for encoding the syntactic analysis, further explained below.

A complication arises in cases where one word is normalised into several contemporary words. For example, “nevboga” (“does not obey”) is now written as “ne uboga” and the “not” is no longer a prefix but a word. The encoding of such a 1-2 mapping is illustrated in Figure 6 (the IDs of the various elements have been removed to improve readability).

The word in focus, “nevboga” is marked up using the word element (*w*), as usual. However, it does not have the normalisation attribute or any other linguistic analysis attributes. Instead, it contains two-word elements, which do have the normalisation attribute (“ne” + “uboga”) along with all the other linguistic analysis attributes. However, these two words have no content.

The example also illustrates the use of the punctuation character (*pc*) element for the comma, as well as the use of the join attribute (*join="right"*) on the penultimate token, meaning that there is no space between “uboga”/“nevboga” and the following comma.

Finally, the syntactic analysis is encoded inside its sentence (*s*) element, in a stand-off fashion. This means it is not encoded as attributes on the tokens, but rather in a special link group (*linkGrp*) element. Each constituent link (*link*) in this group points to the IDs of the relevant tokens, as illustrated in Figure 7. The example shows the last token of the sentence, i.e. the final period, followed immediately by the link group.

The link group is presented as encoding Universal Dependencies syntax (*type="UD-SYN"*) and its target function (*targetFunc="head argument"*), which describes the function of each of the values of the *target* attribute in the contained links, stating that these are the head and argument of the syntactic relation. It also states that the syntactic parse corresponds (*corresp*) to the sentence with the ID *PREG-00-00001.crit.s1*.


```

<ab xml:id="PREG-00-00001" n="1">
  <seg xml:lang="sl-bohoric" xml:id="PREG-00-00001.dipl" type="dipl">Brez muje fe zhreul ne
  obuje.</seg>
  <seg xml:lang="sl" xml:id="PREG-00-00001.crit" type="crit">Brez muje se čreul ne
  obuje.</seg>
  <seg xml:id="PREG-00-00001.crit.norm" type="crit.norm" xml:lang="sl">Brez muje se čevlj ne
  obuje.</seg>
  <seg xml:id="PREG-00-00001.crit.norm.ana" type="crit.norm.ana" xml:lang="sl">
    <s xml:id="PREG-00-00001.crit.sl">
      <w ana="mte:Sg" msd="UPostag=ADP|Case=Gen" lemma="brez"
        xml:id="PREG-00-00001.crit.sl.t1">Brez</w>
      <w ana="mte:Ncfsg" msd="UPostag=NOUN|Case=Gen|Gender=Fem|Number=Sing" lemma="muja"
        xml:id="PREG-00-00001.crit.sl.t2">muje</w>
      <w ana="mte:Px-----y" msd="UPostag=PRON|PronType=Prs|Reflex=Yes|Variant=Short"
        lemma="se" xml:id="PREG-00-00001.crit.sl.t3">se</w>
      <w norm="čevlj" ana="mte:Ncmsn" msd="UPostag=NOUN|Case=Nom|Gender=Masc|Number=Sing"
        lemma="čevlj" xml:id="PREG-00-00001.crit.sl.t4">čreul</w>
    </s>
  </seg>
</ab>

```

Figure 5: Example of a linguistically annotated proverb.

```

<ab xml:id="PREG-00-00213" n="313">
  <seg xml:lang="sl-bohoric" type="dipl">Kdur itariše nevboga, nima írezhe od Boga.</seg>
  <seg xml:lang="sl" type="crit">Kdur stariše nevboga, nima sreče od Boga.</seg>
  <seg type="crit.norm" xml:lang="sl">Kdor starše ne uboga, nima sreče od Boga.</seg>
  <seg type="crit.norm.ana" xml:lang="sl">
    <s>
      <w norm="kdor" ana="mte:Pr msn"
        msd="UPostag=PRON|Case=Nom|Gender=Masc|Number=Sing|PronType=Rel"
        lemma="kdor">Kdur</w>
      <w norm="starše" ana="mte:Ncmpa" msd="UPostag=NOUN|Case=Acc|Gender=Masc|Number=Plur"
        lemma="starš">stariše</w>
      <w>nevboga
        <w norm="ne" ana="mte:Q" msd="UPostag=PART|Polarity=Neg" lemma="ne"/>
        <w norm="uboga" ana="mte:Vmbr3s"
          msd="UPostag=VERB|Mood=Ind|Number=Sing|Person=3|Tense=Pres|VerbForm=Fin"
          lemma="ubogati" join="right"/>
      </w>
      <pc ana="mte:Z" msd="UPostag=PUNCT">,</pc>
      ...
    </s>
  </seg>
</ab>

```

Figure 6: Example of a 1-2 mapping between the diplomatic and critical words.

```

<pc ana="mte:Z" msd="UPostag=PUNCT" xml:id="PREG-00-00001.crit.sl.t7">.</pc>
<linkGrp type="UD-SYN" targFunc="head argument" corresp="#PREG-00-00001.crit.sl">
  <link ana="ud-syn:case"
    target="#PREG-00-00001.crit.sl.t2 #PREG-00-00001.crit.sl.t1"/>
  <link ana="ud-syn:obl"
    target="#PREG-00-00001.crit.sl.t6 #PREG-00-00001.crit.sl.t2"/>
  <link ana="ud-syn:expl"
    target="#PREG-00-00001.crit.sl.t6 #PREG-00-00001.crit.sl.t3"/>
  <link ana="ud-syn:nsubj"
    target="#PREG-00-00001.crit.sl.t6 #PREG-00-00001.crit.sl.t4"/>
  <link ana="ud-syn:advmod"
    target="#PREG-00-00001.crit.sl.t6 #PREG-00-00001.crit.sl.t5"/>
  <link ana="ud-syn:root" target="#PREG-00-00001.crit.sl #PREG-00-00001.crit.sl.t6"/>
  <link ana="ud-syn:punct"
    target="#PREG-00-00001.crit.sl.t6 #PREG-00-00001.crit.sl.t7"/>
</linkGrp>
</s>

```

Figure 7: Example of a dependency parse of a proverb.

Each of the link elements in the link group then encodes a dependency syntactic relation between two tokens. The target attribute contains a pair of pointers, the first to the head and the second to the argument of the relation, while the syntactic dependency type is encoded in the linguistic analysis (*ana*) attribute. As with the MULTTEXT-East MSDs for tokens, the value of this attribute serves as a pointer to the taxonomy of the Universal Dependencies syntactic relations.

It should be noted that the TEI header of the linguistically analysed Collection has a few elements not present in the plain-text version. These include the taxonomy of the Universal Dependencies syntactic relations, the definitions of extended pointers (element *prefixDef*), such as the *ud-syn* prefix in the syntactic analysis above, and the application information (*appInfo*) detailing which tools have been used to perform the linguistic analysis.

3 Publishing the Collection

Our aim was to publish the Collection in open access, with minimal barriers to its use by other researchers and interested parties, thereby contributing to the principles of Open Science. Furthermore, we also wanted to make the Collection immediately available for analysis, rather than simply offering its download. In this Section, we discuss these two methods of distribution.

3.1 The downloadable dataset

We published the Collection in the repository of the Slovenian node of the research infrastructure for language resources and technologies CLARIN.SI (Babič et al. 2023), where it is available for download under the Creative Commons – Attribution (CC BY) licence. The repository offers long-term access to the data, gives it a permanent identifier, as well as clearly stating how the dataset must be cited when used in further research, giving appropriate credit to the authors.

The Collection is available in three formats. First, there is the canonical TEI encoding, distinguishing the “plain text” variant from the linguistically analysed one. The reason for distinguishing between the two is that not everybody will be interested in the linguistic analyses, and the linguistically analysed variant is much larger than the plain-text one. The TEI encoded Collection comprises the following files:

- ♦ *pregovori-viri.xml*: the file with the bibliographic sources;
- ♦ *pregovori.xml*: the root file (i.e. the *TEI* element) of the plain-text version with the TEI header, which, using the XInclude mechanism, includes the files *pregovori-viri.xml* and *pregovori-text.xml*;
- ♦ *pregovori-text.xml*: the file with all the paremiological units (i.e. the *text* element) in their plain-text version;

- ♦ *pregovori.ana.xml*: the root file (i.e. the *TEI* element) of the linguistically annotated version with the *TEI* header; through the *XInclude* mechanism, it includes the files *pregovori-viri.xml* and *pregovori-text.ana.xml*;
- ♦ *pregovori-text.ana.xml*: the file with all the paremiological units (i.e. the *text* element), including their linguistically annotated version.

Using the *TEI*-encoded variant requires some programming skills and familiarity with *XML* and *TEI*. Therefore, we have converted the *TEI* into a simpler format, which is also made available in the repository entry. This variant is similar to the one used as the source for the *TEI* and consists of two files, both formatted as *TSV* tables with a header row and *TAB*-separated columns:

- ♦ *pregovori-viri.tsv*: the file with the bibliographic sources, giving the ID for each source, the ID as used in the source collection, the category of the unit, its year(s) of publication, and the name of the bibliographical unit;
- ♦ *pregovori.tsv*: the file with the proverbs, giving each ID, the source ID, the diplomatic, critical, and normalised transcription, as well as flags specifying if there is a difference between the diplomatic and critical transcription (flag *DC*) and/or between the critical and normalised transcription (flag *CN*).

Finally, the third format in which the collection is available is the so-called vertical format, automatically converted from the linguistically annotated *TEI* version of the corpus. Vertical files are used by many concordancers, particularly those offered by *CLARIN.SI*, as detailed in the next subsection. A vertical file contains *XML*-like tags (structural annotations), possibly with attributes, and one line for each token, giving the token itself in the first column followed by an arbitrary number of columns with further information on the token, the so-called positional attributes. It should be noted that the vertical file does not contain all the information from the source *TEI*, and that it uses different element and attribute names from the *TEI* source in order to make it more intuitive when used in concordancers. This variant in the repository entry consists of two files:

- ♦ *pregovori.vert*: the vertical file, which has the structure *text* delimiting one paremiological unit and contains attributes for the bibliographical sources, their types and years of publications, as well as an attribute for the diplomatic transcription of the proverb. The subordinate structure is the sentence (*s*), which in turn encompasses the tokens. These tokens correspond to those found in the critical transcription, featuring positional attributes such as the normalised form of the word, its lemma, the *MULTEXT-East MSD* in Slovenian and English, the Universal Dependencies part-of-speech, the Universal Dependencies morphological features, the ID of the token, and its syntactic dependency label to the head. The syntactic head of the token is provided with the same attributes

as the token. This facilitates searching for the various properties of the syntactic head associated with each token in a concordancer.

- ♦ *pregovori.regi*: the so-called registry file, which gives the names and properties of the structural annotations and of the positional attributes, and is used by the concordancers to compile and present the corpus. This registry file is identical to the one used for the CLARIN.SI concordancers, meaning that installing the corpus on some other concordancer necessitates changing some of its values, e.g. the location of the files on the system.

3.2 The Collection in concordancers

To make it possible for the Collection to be searched and analysed, we have also made it available on the CLARIN.SI concordancers, which are online tools that enable users to search text corpora and display the results in various ways. CLARIN.SI hosts two concordancers, namely noSketch Engine (Rychlý 2007; Kilgarriff et al. 2014) and KonText (Machálek 2020). While they differ in their front-end interfaces and login options, they use the same back-end – the part of the software responsible for conducting corpus searches and defining the query language for searching. This query language, called CQL (Corpus Query Language), is very powerful, as it allows searches across combinations of positional attributes and limiting searches to subcorpora defined by their structures. Furthermore, it supports searches not only by literals but also by regular expressions. For example, users can search for words (or lemmas, part-of-speech tags, etc.) that start, end, or contain a certain substring.

To give an impression of a search, Figure 8 illustrates the results of a simple search for the word “volk” (“wolf”) in the Collection. The results show 320 occurrences of “volk” in the corpus, providing the IDs of the proverbs where it appears and the proverbs themselves.

Although the object of the search was “volk”, the displayed proverbs feature this word not only in its various inflectional forms but also with archaic spelling, such as “vovka”, now written as “volka”. This occurs because simple searches not only scan the word tokens but also the normalised forms and especially the word lemmas. Such a search demonstrates the advantages of the added linguistic annotation. It should also be noted that clicking on the displayed metadata of a hit (i.e. the proverb ID) triggers a window to pop up, giving all the available metadata for a proverb.

The concordancers also support other types of analyses, with the main ones being:

- ♦ frequency lists of the chosen positional attribute;
- ♦ keywords computed against a chosen (usually reference) corpus or of a subcorpus against the whole corpus;
- ♦ collocations of the searched-for word or phrase;
- ♦ text-type analysis for the whole corpus.

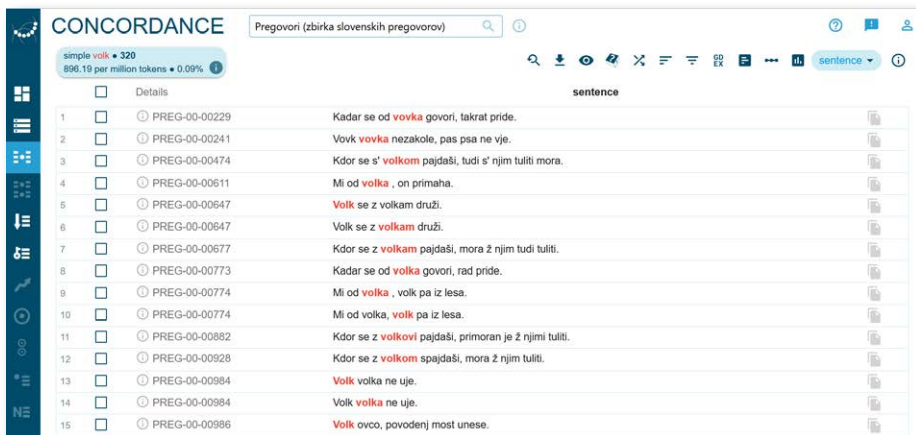


Figure 8: Example of a search in the noSketch Engine.

The interface to the concordancers is RESTful, so when a query is submitted to a concordancer, the URL of the result contains all the information necessary to re-run the query. In other words, it is possible to share the URLs for particular queries with others, and the concordancers can also be queried by programs if appropriate URLs are constructed.

4 Quantitative overview

In this section, we briefly review some of the Collection's more distinctive quantitative characteristics. We first consider its size and scope, and then the use of its paremiological units in a large Slovenian language corpus.

4.1 The size and scope of the Collection

The main distinguishing feature of the Collection is its large size: it contains 37,390 paremiological units, or 286,798 words in the critical transcription. The average length of a unit is therefore 7.67 words. Of all the paremiological units, 4,556 units (12.2%) have a critical transcription different from the diplomatic one, while 8,689 units (23.2%) have an automatically normalised transcription different from the critical one. In terms of the vocabulary used, there are 33,184 different word forms that appear in the critical transcriptions of the units (29,494 if we disregard the casing), 29,615 normalised forms (26,274 disregarding casing), and 15,390 different lemmas (15,170 disregarding casing).

The list of bibliographical sources comprises 2,631 units, although only 2,599 appear as sources for the paremiological units. In other words, 32 units are listed but not used. Out of all the sources, 56 (2%) of them do not have a date of publication. The oldest source was published in 1592, three are from the 17th century, 24 (1%) from

the 18th century, 667 (25%) from the 19th, and 1,880 (72%) from the 20th century. The most recent one is from 2019.

Table 1 gives the use of sources by the paremiological units categorised by their type. The first column gives the number of mentions of the source type in the Collection, the second column indicates the number of bibliographical sources corresponding to a particular type, and the third column specifies the type.

<i>Sources</i>	<i>Units</i>	<i>Type of bibliographical unit</i>
14,962	10	Collections of proverbs
8,067	1,124	Periodicals – newspapers, magazines and newsletters
4,940	93	Professional sources
4,080	27	Occasional sources – private collections
4,008	81	Occasional sources – collection campaigns
3,678	52	Manuscript legacies
2,989	79	Grammars and dictionaries
2,812	166	Periodicals – calendars
2,551	343	Oral sources
2,421	478	Fiction
1,168	10	Seminar and diploma theses
716	51	Museum collections
647	24	Published literary folklore
527	33	Older book sources
370	23	Random sources
214	15	Occasional sources – elementary school work
123	10	Periodicals – chronicles
11	11	Radio and television
54,284	2,630	18

Table 1: Distribution of bibliographical sources by type.

By far the largest number of mentions is from the ten printed collections of paremiological units. In this group, the largest number of mentions (4,884) is associated with the book by Etbin Bojc (1974) “Pregovori in reki na Slovenskem” (“Proverbs and Sayings in Slovenia”), which is also the overall highest-mentioned bibliographical source. As discussed in Babič and Erjavec (2022), it should be noted that this book took many of its proverbs from existing collections and older grammars and dictionaries. However, this collection is the first contemporary one.

4.2 Usage of the paremiological units in metaFida

Investigations into the use of paremiological units in corpora have a long tradition (e.g. Steyer 2017), including in the Slovenian language (e.g. Meterc 2019, 2021; Babič and Erjavec 2022). However, the method used to investigate their behaviour has typically been to concentrate on particular paremiological units (including variations). In contrast, our aim was to investigate how many of the paremiological units from the Collection are used in Slovenian and how frequently. In order to achieve this, we adopted a corpus-based approach, but rather than hand-constructing queries we wrote a script to query all the units on the CLARIN.SI noSketch Engine concordancer and simply counted the number of hits returned by the concordancer.

The corpus we queried is called metaFida (Erjavec 2023) and with 4.7 billion words, it is currently the largest corpus of Slovenian texts. It is composed of 34 individual corpora and is therefore a second-order corpus composed of previously compiled Slovenian corpora. The included corpora contain a wide variety of text types, from historical Slovene texts (the oldest one dating from 1584) to tweets and other user-generated content (the most recent texts being from 2022). The corpus was deduplicated on the paragraph level, which meant 11% of the paragraphs and 7% of the texts were removed. The annotations comprise the structures for individual texts (with attributes for the text metadata), paragraphs, sentences and gaps (for removed paragraphs). Tokens have positional attributes for the normalised form of the word (typically identical to the word, except for historical and user-generated corpora, where the words have been automatically modernised/standardised), the lemma and the MULTEXT-East MSD in Slovenian and English. The word, the normalised word and lemma attributes also have equivalent attributes for querying the lower-cased version of the attribute, e.g. *norm_lc* for a lower-cased normalised word. This makes it possible to query a particular attribute, regardless of its capitalisation in the corpus.

The script that queries the corpora takes each token of the normalised paremiological unit and converts it into a query of that token. As the paremiological units in the corpus often lack or have slightly different punctuation from those in the Collection, the punctuation tokens are queried with the MSD for punctuation (*Z* in the English MSDs), and are given as optional, i.e. with a question mark after the token. For example, if the normalised form of a unit is “*Dolgi lasje, kratka pamet.*” (“Long hair, short mind.”), the constructed query is `[norm_lc="Dolgi"] [norm_lc="lasje"] [tag_en="Z"]? [norm_lc="kratka"] [norm_lc="pamet"]`. Note also that the final full-stop is omitted, as making it optional would give two hits for each attestation where the full-stop is actually present in the corpus: one hit for the unit with the full-stop and one without.

The querying of metaFida for all the paremiological units in the Collection returned 5,733 hits. This means that out of the 37,390 units, 6.5% were actually found in the corpus, indicating how many of the units in the Collection are actually used. As would

be expected, the distribution is Zipfian, i.e. a few units have relatively high frequency, which then drops rapidly with roughly half being hapax legomena. More precisely:

- ♦ 2,953 (47%) of the matched units have a frequency of 1, i.e. only a single instance was found in the corpus;
- ♦ 896 (16%) of units have a frequency of 10 or more, which is at the lower limit of when certain conclusions on their use can be drawn from the context of the found units;
- ♦ 69 (1.2%) have a frequency of over 100, which is frequent enough to enable statistical analysis of the results;
- ♦ only 3 units have a frequency of over 1,000, namely:
 - 1,644: “Enkrat za vselej.” (“Once and for all.”)
 - 1,739: “Vaja dela mojstra.” (“Practice makes perfect.”)
 - 3,091: “Vsak po svoje.” (“To each his own.”)

It should also be noted that a fair number of hits in the metaFida corpus are of paremiological units that are mentioned rather than used in the texts. A large part of the metaFida corpus comes from the “Corpus of scientific texts from the Open Science Slovenia portal OSS” (Žagar et al. 2023), which contains 2.6 billion words from texts such as senior, masters and doctoral theses, some of which deal with proverbs and other paremiological units. Here the proverbs are not used in free text, but rather mentioned as examples, making studies of the use of such cases less than useful.

5 Conclusions

Digitising folklore materials facilitates their analysis and enables new, more exact and statistics-based methods for their study. Open access to such materials and using proven and widely used ways to encode them also means that other researchers can perform their own studies and even correct errors or add further annotations, thus increasing their value. This paper has discussed the encoding, publication and size of a large collection of Slovenian paremiological units, as well as their usage in a large Slovenian language corpus. Hopefully, all this will lead to new research into these types of short texts and enable humanities researchers in Slovenia to make further progress in the new field of Digital Humanities.

Regarding further work, the Collection almost certainly contains some typos or inconsistencies, which should be fixed. Another labour-intensive upgrade would be the manual correction of the currently automatically assigned normalised variant of the units. Alternatively, the accuracy of the automatic methods could be increased with a relatively small manually normalised sample from the Collection, and cSMTiser could also be retrained using this newly available training set.

In future, it would be interesting to explore the variation of “the same” paremiological units, adding the annotation that would selectively group them. Some units only differ in terms of punctuation or spelling, and determining which units are variants could be relatively simple to achieve with mostly automatic methods. However, a more ambitious undertaking would also consider variants that differ more substantially, e.g. using (near) synonyms, adding or removing phrases, etc.

Acknowledgements

The author would like to thank the anonymous reviewer for their very valuable comments and suggestions. The work presented in this paper was supported by the basic research project *Traditional paremiological units in dialogue with contemporary use* (ARRS J6-2579) and by the CLARIN.SI research infrastructure.

References

- Babič, Saša; Erjavec, Tomaž, 2022: Izdelava in analiza digitalizirane zbirke paremioloških enot [The compilation and analysis of the digitised collection of paremiological units]. In: *Proceedings of the Conference on Language Technologies & Digital Humanities 2022*. September 15-16 2022, Ljubljana. Online: https://nl.ijs.si/jtdh22/pdf/JTDH2022_Babic_Erjavec_Izdelava-in-analiza-digitalizirane-zbirke-paremioloskih-enot.pdf.
- Babič, Saša; Peče, Miha; Erjavec, Tomaž; Ivančič Kutin, Barbara; Šrimpf Vendramin, Katarina; Kropelj Telban, Monika; Jakop, Nataša; Stanonik, Marija, 2022: *Collection of Slovenian paremiological units Pregovori 1.0*. Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1455>.
- Babič, Saša; Peče, Miha; Erjavec, Tomaž; Ivančič Kutin, Barbara; Šrimpf Vendramin, Katarina; Kropelj Telban, Monika; Jakop, Nataša; Stanonik, Marija, 2023: *Collection of Slovenian paremiological units Pregovori 1.1*. Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1853>.
- Dobrovoljc, Kaja; Erjavec, Tomaž; Krek, Simon, 2017: The Universal Dependencies Treebank for Slovenian. In: *Proceedings of the 6th Workshop on Balto-Slavic Natural Language Processing*. Association for Computational Linguistics, 33–38. DOI: <http://dx.doi.org/10.18653/v1/W17-1406>.
- Erjavec, Tomaž, 2012; MULTEXT-East: Morphosyntactic Resources for Central and Eastern European Languages. *Language Resources and Evaluation* 46/1, 35–57.
- Erjavec, Tomaž, 2015; *Reference corpus of historical Slovene goo300k 1.2*. Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1025>.
- Erjavec, Tomaž, 2023: *Corpus of combined Slovenian corpora metaFida 1.0*. Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1775>.
- Kilgarriff, Adam; Baisa, Vít; Bušta, Jan; Jakubiček, Miloš; Kovář, Vojtěch; Michelfeit, Jan; Rychlý, Pavel; Suchomel, Vít, 2014: The Sketch Engine: ten years on. *Lexicography* 1/1, 7–36. DOI: <https://doi.org/10.1007/s40607-014-0009-9>.
- Koehn, Philipp, 2010: *Statistical Machine Translation*. Cambridge University Press.
- Ljubešić, Nikola; Dobrovoljc, Kaja, 2019: What Does Neural Bring? Analysing Improvements in Morphosyntactic Annotation and Lemmatisation of Slovenian, Croatian and Serbian. In:

- Proceedings of the 7th Workshop on Balto-Slavic Natural Language Processing, Association for Computational Linguistics*, 29–34, DOI: <http://dx.doi.org/10.18653/v1/W19-3704>.
- Machálek, Tomáš, 2020: KonText: Advanced and Flexible Corpus Query Interface. In: *Proceedings of the Twelfth Language Resources and Evaluation Conference*, 7003–7008. Online: <https://aclanthology.org/2020.lrec-1.865.pdf>.
- de Marneffe, Marie-Catherine; Manning, Christopher; Nivre, Joakim; Zeman, Daniel, 2021: Universal Dependencies. *Computational Linguistics* 47/2, 255–308.
- Meterc, Matej, 2019: Analiza rabe paremioloških uvajalnih sredstev v parlamentarnih razpravah iz korpusa siParl. In: Tivadar, Hotimir (ed.). *Slovenski javni govor in jezikovno-kulturna (samo)zavest*. Ljubljana: Znanstvena založba Filozofske fakultete (Obdobja, 38). 135–141.
- Meterc, Matej, 2021: Aktualna raba in pomenska določljivost 200 pregovorov in sorodnih paremioloških izrazov. *Jezišlovni zapiski* 27/1, 45–61.
- Rychlý, Pavel, 2007: Manatee/Bonito - A Modular Corpus Manager. In: *Proceeding of the Conference "Recent Advances in Slavonic Natural Language Processing" (RASLAN)*, 65–70.
- Scherrer, Yves; Ljubešić, Nikola, 2016: Automatic Normalisation of the Swiss German Archi-Mob Corpus Using Character-Level Machine Translation. In: *Proceedings of the 13th Conference on Natural Language Processing (KONVENS 2016)*, 248–55.
- Schöch, Christoph; Patraş, Roxana; Erjavec, Tomaž; Santos, Diana, 2021: Creating the European Literary Text Collection (ELTeC). *Modern languages open*. DOI: <https://doi.org/10.3828/mlo.v0i0.364>.
- Stanonik, Marija, 1996: *Slovenski pregovori in rekla* [Slovenian proverbs and sayings]. Project proposal.
- Stanonik, Marija, 2004: *Informatizacija neoprijemljive dediščine za etnologijo in folkloristiko* [The informatisation of intangible cultural heritage for ethnology and folkloristics]. Project proposal.
- Stanonik, Marija, 2009: *Slovenski pregovori kot kulturna dediščina: klasifikacija in redakcija korpusa* [Slovenian proverbs as cultural heritage: classification and redaction of the corpus]. Project proposal.
- Stanonik, Marija, 2015: Slovenski pregovori kot kulturna dediščina. Klasifikacija in redakcija korpusa [Slovenian proverbs as cultural heritage: Classification and redaction of the corpus]. *Traditiones* 44/3, 171–214.
- Steyer, Kathrin, 2017: Corpus Linguistic Exploration of Modern Proverb Use and Proverb Patterns. In: Mitkov, Ruslan (ed.), *EUROPHRAS 2017. Computational and corpus-based phraseology: Recent advances and interdisciplinary approaches. Proceedings of the Conference Volume II (short papers, posters and student workshop papers)*. London, Geneva: Editions Tradulex, 45–52.
- TEI Consortium, 2022: *TEI P5: Guidelines for Electronic Text Encoding and Interchange*. Online: <https://tei-c.org/guidelines/P5>.
- Terčon, Luka; Ljubešić, Nikola, 2023: *CLASSLA-Stanza: The Next Step for Linguistic Processing of South Slavic Languages*. arXiv 2308.04255. DOI: <https://doi.org/10.48550/arXiv.2308.04255>.
- Žagar, Kristjan; Ferme, Marko; Ojsteršek, Milan; Jemec Tomazin, Mateja; Erjavec, Tomaž, 2023: *Corpus of scientific texts from the Open Science Slovenia portal OSS 1.0*. Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1774>.

Typologies and digital databases of proverbs as research tools

Tipologije in digitalne zbirke podatkov pregovorov kot raziskovalna orodja

The author of this presentation has had the opportunity to help in saving and bringing into use a large database of proverbs and their literal sources. Professor of folkloristics at Helsinki University, Matti Kuusi (1914–1998) wanted to test his idea of creating an international typology of proverbs. The idea was to create a flexible system of proverb types with a versatile typology based on different cultural meanings accrued within these traditional expressions. Alternative interpretations of proverbs connected to their changing forms and additions in relation to typologies should be discussed. Developing cross-references between proverb types within the typology to help with their interpretation has proven difficult. The variety of changing contexts of proverb use is a challenge for any flexible digital network of proverb interpretations.

• **Keywords:** Matti Kuusi, proverbs, paremiology, typology, semantic classification, database

Avtorica članka je imela priložnost pomagati pri reševanju in uporabi velike zbirke pregovorov in njihovih dobesednih virov. Profesor folkloristike na Univerzi v Helsinkih, Matti Kuusi (1914–1998) je želel preizkusiti svojo zamisel o oblikovanju mednarodne tipologije pregovorov. Zamisel je bila ustvariti prilagodljiv sistem tipov pregovorov z vsestransko tipologijo, ki bi temeljila na različnih kulturnih pomenih, nabranih v teh tradicionalnih izrazih. Obravnavane so bile tudi alternativne razlage pregovorov, povezane z njihovimi spreminjajočimi se oblikami in dopolnitvami v povezavi s tipologijami. Razvijanje navzkrižnih povezav med tipi pregovorov znotraj tipologije, ki bi pomagale pri njihovi razlagi, se je izkazalo za težavno. Raznolikost spreminjajočih se kontekstov uporabe pregovorov je izziv za vsako prožno digitalno mrežo razlag pregovorov.

• **Ključne besede:** Matti Kuusi, pregovori, paremiologija, tipologija, semantična klasifikacija, podatkovna baza

Classification and conceptualizing are central tools for scientific research. In order to find regularities, similarities and dissimilarities between phenomena you need some criteria for structuring the material. In practice, groups and hierarchy are more or less artificial forms of classifying the material at hand. Problems arise only when we have to use pre-established groups and find slots between them. Sometimes you can even refuse to accept a given classification.

Thus, a better approach than static classification would be a dynamic network and its nodes. The aim of this article is to study typologies and digital databases of proverbs. Most of them are focused on one language area or serve the contrastive paremiology of two or three cultures.

Construct an all-embracing and global proverb typology seems like an irrational challenge. Who would realistically consider the possibility of capturing the multidimensionality of proverbial semantics, and who would embark on the task of searching

for global proverb types equipped with a general meaning? As yet, there have only been several such attempts. One of them, the Finnish Matti Kuusi International Type System of Proverbs¹ is the main focus of this article. Another general typology of proverbs is elaborated by the Perm-born Soviet paremiologist Grigory Permyakov² (1919–1983).

Proverbs are understood here to be a global folklore genre, which is unique in each culture and still alive in folklore and vernacular speech.³ The internationality of the proverb genre as a whole is unanswerable. The global or near global distribution of numerous proverb ideas and generative patterns is well proven.⁴

Professor emerita Galit Hasan-Rokem collates well the multifaceted challenge concerning any proverb typology.

The study of proverbs has long been a subject of research by scholars from different disciplines, such as folklorists, philologists, logicians, psychologists and anthropologists.

Consequently, various research motivations and methods have produced diverse approaches towards the organization and classification of proverbs into convenient indices for further search and retrieval. These approaches mostly differ in what should be considered key proverbial characteristics upon which an index or classification scheme should be devised, for example: text, logical or syntactic structure, context of usage, literary source, time and country of origin, or semantic field (Hasan-Rokem and Kats 2009).

Folkloristic, linguistic and social psychological approaches are all useful for appraising how different meanings accumulate for each proverb type. Literature and historical research also have valid methods for paremiology. The history of paremiology, the traditional study of proverbs, is strongly based on printed sources, dictionaries and collections. Its foundation lies in lexicography, safely rooted in linguistics. Constructing a proverb typology that would meet all the challenges that different scholars might pose is therefore a complex task.

1 The background of the Matti Kuusi international database and typology

Matti Kuusi (1914–1998), professor of folkloristics at Helsinki University, wanted to test his ideas for an international typology of proverbs. The idea was to create a flexible system of proverb types with a versatile typology based on different cultural meanings

¹ <https://www.mattikuusiproverbtypology.fi/> (visited 18.12.2023). The author of this article is acting as one of the administrators.

² More on the significance of Permyakov in: http://peter-grzybek.eu/science/publications/1985/grzybek_1985_permyakov.pdf (visited 18.12.2023).

³ The best recent source for definitions of the proverb is Hrisztova-Gotthardt's and Aleksa Varga's comprehensive guide to proverb studies *Introduction to Paremiology* (2014).

⁴ More about proverb patterns in Lauhakangas (2021).

accumulated within these traditional expressions. How did he end up realising his idea of an international typology? Let us shed light on his background in folkloristics and his qualifications to construct a cross-cultural proverb typology.

Kuusi's dissertation in 1949 was a study of Finnish Karelian Kalevala epics and their typology. He had noticed that proverbs were potential instruments for dealing with extensive material and making conclusions about language variations and structures (Kuusi 1978: 11). Kuusi's interest was in essential questions of human mentality and he soon became attentive to comparative paremiology. The idea of creating an international typology of proverb types was latent in his early research methodology.

Kuusi's first challenge in paremiology was a source-critical study of the oldest collections and unpublished manuscripts of Finnish proverbs. His scientific approach did not prevent him from converting the results into a popular Finnish collection (1954).

During the 1960s, Kuusi instructed students to find proverbs in the manuscripts of the folklore archives of the Finnish Literature Society. He continued his project by leading his assistants to construct a card-index of c. 300,000 archival variants of proverbs from Finnish parishes. A guiding principle was to gather all the variants for each proverb type. Meanwhile, in his free time Kuusi began collecting foreign-language equivalents of Finnish proverb types, noting down references to proverb collections and bilingual dictionaries. He had access to the main European collections, multilingual collections and special editions, which he bought privately or which were sent to him during his time as editor-in-chief of the *Proverbium*⁵ journal (1965–1974). He continued with this activity until the end of the 1980s, constantly searching for the latest archival materials to be saved in the folklore archives in Helsinki (Kuusi 1988).

Kuusi wrote his first outline for constructing a 'type-index of international proverbs' in Proverbium 15 (1970a: 473–476), and a detailed experiment of categorizing number proverbs as a test bed for further research in Proverbium 19 (1972: 699–735). A much-needed challenge for typological work came from the Soviet Union, when Grigory Permyakov's book At pogovorki do skazki was published in 1970. It was partly translated into English: From proverb to folk-tale. Notes on the general theory of cliché (1979). Kuusi had already commented the classification in Permyakov's important collection Poslovitsy i pogovorki narodov vostoka [Proverbs and sayings of the peoples of the East] (1968 [1961]) in Proverbium 15.

Kuusi took elements of basic oppositions from Permyakov's theory and typology for his own classification. Kuusi wanted to 'construct a universal comparison of deep

⁵ According to Kuusi's own notes, the scientific journal was established in Helsinki through the persistent encouragement of prof. Archer Taylor. It began as *Proverbium*, later evolving into four volumes of *Proverbium Paratum* by prof. Vilmos Voight et al. (1980–89), followed by prof. Wolfgang Mieder's *Proverbium. the Yearbook of International Proverb Scholarship* (1984–2021), now an open access peer-reviewed journal helmed by editors-in-chief Dr. Melita Aleksa Varga and Dr. Hrisztalina Hrisztova-Gotthardt (2022–) <https://naklada.ffos.hr/casopisi/index.php/proverbium/about> (retrieved 18.12.2023)

structures of proverbs'.⁶ As early as the 1960s, before he became actively involved in constructing the universal typology, Kuusi began cooperating with Estonian scholars, especially with Arvo Krikmann who was also interested in analysing proverb types. Krikmann's leading idea was to find an optimal level for comparing proverb types, their history and distribution in different cultures. Krikmann's work continued after Kuusi's death. He published a classification of Estonian proverbs in English (Krikmann 1989) and in 2013 he carried out an unpublished experiment⁷ on his typology using an English corpus. Krikmann's typology might interest those who prefer studying the linguistic structure of proverbs.

While waiting for the beginning of the first Baltic-Finnish proverb project, which was delayed until the 1980s due to political reasons (Kuusi et al. 1985), Kuusi focused on other cultures. For example, he accidentally stumbled upon a Finnish missionary's unpublished collection of African Ndonga proverbs and began detailed work on equivalents in African languages (Kuusi 1970b).

We can conduct a thought experiment and compare the current AI compilation and search facilities to Kuusi's 'personal computer', namely his ability to process the material he had in his memory and personal card index with literature references. He possessed an exceptional visual memory and had creative ways of connecting comparable network nodes.

The main focus of Kuusi's project was accumulating references for each proverb type. By connecting proverbs from different sources he was able to hypothesise either on the similarity or simultaneity of these cultural items. The Matti Kuusi international proverb typology made connections between Polish proverbs, Chinese or Finnish dialectal utterances, traditional wisdom in the Ndonga dialect, etc. The connections are suggestions to be evaluated, ideally on an individual basis.

Kuusi (1966: 97–98) had three basic concepts concerning any proverb. Each one has a core element, a form or schema can define a proverb's type, but the most critical thing is to evaluate the idea (especially to recognize synonymous proverbs).

That was how he arranged and searched for prototypical proverb types and their exceptions. We can imagine that he had his own ontology,⁸ in the modern sense of the term, in his mind. This term is now used in digital humanities to refer to the process of composing search engines.

⁶ An unpublished research plan for the second stage written immediately after the *Proverbia Septentrionale* was published in 1985.

⁷ The author of this article received a draft file of the applied typology from A. Krikmann before his serious illness and death in 2017.

⁸ Ontology is a set of concepts and categories in a subject area or domain that shows their properties and the relations between them. E.g. 'what's new about our ontology is that it is created automatically from large datasets' (source: languages.oup.com).

2 The uniqueness of proverbial thinking brought up in Matti Kuusi's proverb typology

The construction of the Matti Kuusi international database and typology has been described in a detailed report (Lauhakangas 2001), and the proverb types of Kuusi's system have also been used and studied from the standpoint of the social functions of proverb use (Lauhakangas 2004, 2023). The material and its classification open the possibility of gauging tensions in the human mind on a broader scale. They touch all possible human situations, preconditions, options and means of survival.

For an over-all picture of Kuusi's whole type system, the main themes cover both concrete human conditions such as 'Practical knowledge of nature' and abstract concepts such as 'Time and sense of time'. The focus is always on human experience. It may be experience of the world and life or how to cope in life and learn. The most elaborate and central paremic questions deal with human interaction, communication and hierarchy.

There are themes that are specific to proverbial thinking such as 'basic observations and socio-logic' (Kuusi's main theme C). The term 'socio-logic' was actually chosen by the author of this article when constructing Kuusi's relational database. It derives from the work of Paul Goodwin and Joseph Wenzel (1981) and is a socially generated practical way to reason. David Cram (1983) explains proverbs as quasi-tautologic axioms. They are expressed like self-evident conclusions, but they have an idiosyncratic form that allows inconsistency. Finally, the social context determines the effect. Any metaphor is accepted when it is needed.

The main theme C has six classes with four to nine subgroups each. Here are some examples.

- (1) *Boys will be boys.* C1b 11
- (2) *You can't skin a flint.* C3a 10
- (3) *Beware of a silent dog and still water.* C5e 10

The socio-logic way of thinking characteristic of proverbial reasoning can also be approached from a structural point of view as a special form of syntactic architecture. Wolfgang Mieder has shown that traditional proverbial markers in American English have been transformed into simplified forms, especially syntactic and phonological devices (Mieder 2012: 144–147; Mac Coinnigh 2014: 118–119.) We can still ask if this linguistic evolution has substantially changed the function of proverbs as practical tools of reasoning.

Kuusi introduced another important theme in proverbial thinking: the notion of 'sense of proportion' (Kuusi's main theme E). It deals with the ways people can explain situations better when events unfold unexpectedly, but a proverb aptly defines the on-going situation. The following three proverbs demonstrate this way of thinking.

- (4) *Crooked logs make straight fires.* E1h 11
- (5) *Behind the cross there is the devil.* E1i 13
- (6) *There is good in everything, even in misfortune.* E1k 31

3 Current digital sources and databases of proverbs

Nowadays, there are more or less entertaining tools available on the internet such as Gary Martin's web-based archive of phrases, sayings, quotes and clichés. He is the author of the English Meanings and Origins section of the Phrase Finder website,⁹ which was founded in 1995. The website includes 680 English proverbs, half of them equipped with external links to explanations and historical information. Martin has been working on a computational linguistics research project. His aim is to offer historical knowledge and explanations of phrases, so his approach to classifying phrases involves typical keywords, which we could find in any collection of quotations. Martin does not work for any institution, but he is one of those valuable "content creators" who help connect old sources to modern databases.

One of the most scientific tools for studying special language corpora is an Austrian Academy Corpus.¹⁰ For example, it includes digitized magazines. One of them is "Die Fackel"¹¹ 1899-1936, originally published and almost entirely written by the satirist and language critic Karl Kraus in Vienna. The electronic text of the entire journal has more than six million tokens. It was already available in 2007 with indexes and search and navigation tools in a functionally designed interface (Biber 2015).

The Sketch Engine,¹² which makes automatic derived summaries of the grammatical and collocational behaviour of words, is a huge resource for studying words in their context in different languages. Another large system that has opened new possibilities for lexicographers and phraseologists is GATE – General Architecture for Text Engineering,¹³ which has been developed at the University of Sheffield from 1995. It is a global tool for natural language processing tasks.

International databases specialising in proverbs have also been developed as part of academic projects. A good example is the Sprichwort-Plattform: Sprichwort-Datenbank.¹⁴ In addition to the database documents, there is material on the current use of proverbs in five interlinked languages (German, Slovenian, Slovak, Czech and Hungarian). This means you can access a large number of authentic texts in which the proverbs appear, and obtain information about meanings, abnormalities in use, etc.

An empirical study has demonstrated that even widely-used online corpora and marketing tools, such as Google Trends, make it possible to map the dissemination of modern proverbs (Villers 2022). The possibility of obtaining samples of real life

⁹ <https://www.phrases.org.uk/meanings/proverbs.html> (retrieved 18.12.2023).

¹⁰ https://arche.acdh.oeaw.ac.at/browser/oeaw_detail/23885 (retrieved 18.12.2023).

¹¹ <https://fackel.oeaw.ac.at/> (retrieved 18.12.2023).

¹² <https://www.sketchengine.eu/> (retrieved 18.12.2023).

¹³ <https://gate.ac.uk/> (retrieved 18.12.2023).

¹⁴ <https://www.germanistik-im-netz.de/informieren/detail/ressource/sprichwort-plattform-sprichwort-datenbank/> (retrieved 18.12.2023).

situations was practically impossible before the appearance of internet discussion forums. We still have time to collect reliable material before algorithms begin to remove whatever mimics human conversation.

4 The point of view of information technology

There has been a recent attempt to classify proverbs semantically (Zhitomirsky-Geffe, Prebor, and Bloch 2017). This is a small-scale study called “Improving proverb search and retrieval with a generic multidimensional ontology”.

The authors write about the goal of their research in the abstract. Their aim seems to fit with the purpose of Kuusi’s classification system. However, another criterion for success emerges, which is entirely connected with information retrieval:

[...] to develop a generic ontological model for proverbs that unifies potential classification criteria and the various characteristics of proverbs to enable their effective retrieval and large-scale analysis.

[...] to improve the search recall, precision, user satisfaction, and efficiency, and to minimize user effort during the search process.

A small proverb corpus has been utilised to study to what extent given semantic relations speed up the search for proverbs. This is the point at which folkloristic and linguistic aims differ from the objective of information technology. Although the Israeli team indicates their target to build ‘a multidimensional ontology suitable for proverb classification’, the criterion for the ‘effectiveness of the constructed ontology’ seems to be more about user satisfaction than achieving increased understanding of the proverb genre.

The Israeli research group introduces a definition of ontology as a semantic scheme, which comprises the main classes (concepts) of a given domain of knowledge, their properties, inter-relationships and instances (Noy & McGuinness, 2001).

The Israeli group used some of Kuusi’s main classes and similar terms such as antonyms and causal relations. When studying their solutions for proverb ontology, there were familiar analyses, such as ‘cause-of’, ‘antonym’, ‘part of’. The authors write critically about Matti Kuusi’s typology:

[...] the structure of the thematic hierarchy is based on free associations that might in some cases be quite controversial. In addition, some of the theme classes represent the explicit terminology of the proverbs related to them (e.g. “natural elements”), while others convey their implicit meaning (e.g. “social position”).

The Israeli group’s claim that Kuusi ‘had often used free association’ can be viewed in light of Kuusi’s expertise in folklore and paremiology prior to constructing his typology. The second part of the criticism bypasses the cross-reference system in Kuusi’s typology.

They differentiate between ET as ‘explicit terminology’ and MT as ‘metaphoric terminology’. For metaphoric use they have chosen ‘Kuusi’s generic categories: social elements, cognitive elements, elements in human life, elements of time and elements of religion.’

The Israeli team accepted many concepts common to Kuusi’s main groups: ‘conception of morality’, ‘communication’, ‘dynamic of needs’, ‘human life’, ‘time’, ‘sense of proportionality (relativity)’, ‘social interaction’, ‘social life’. But they have used some concepts that belong to the subgroup level in Kuusi’s classification, on the same level: ‘social status’, ‘family ties’, ‘success’, ‘work’. Their terms, such as ‘women’, ‘studying’ and ‘work’ seem quite non-analytical, and some terms like ‘criticism’, ‘grading’, ‘behaviour’, ‘element in human life’ and ‘person’ seem miscellanea. The small scale of their corpus naturally explains the difficulties to construct any hierarchy of typology. Even a more difficult challenge is “to bind the metaphoric concepts to their semantic interpretations”, if we think about multifaceted nature of proverbs.

5 The importance of internal connections of the Matti Kuusi type-system

It has been difficult to find ways to create such cross references between proverb types in the typology that would help their interpretation. The variety of changing contexts of proverb use is a challenge for any flexible digital network of proverb interpretations.

When the language of folklore is concerned, it is pointless to expect that concepts would be nicely distributed to categories of concrete, abstract, metaphoric and semantically interpretable items. For example, ‘bread’ as a word in a proverb may be translatable to almost all languages, but the meanings of ‘bread’ in different contexts and in all the layers of each culture, in which the word ‘bread’ is used may considerably differ from culture to culture.

Professor emerita Galit Hasan-Rokem in Hebrew University of Jerusalem has commented the difficult question about proverb semantics: “The most important part of the meaning of the proverb lies outside of it” (1982: 15). This brings the need of contextual knowledge to qualify the classification of proverbs. But we can also take Hasan-Rokem’s argument to remind that there are general characteristics of proverbial speech that we must be aware. They are present in other genres of oral tradition, too. For example, the stylistic device of hyperbole, exaggeration is realized, certainly understood and expected as part of narrative as well as proverbial speech. For example

(7) *Fear has a thousand eyes* or

(8) *Buyers want a hundred eyes, sellers none.*

This is what we must always be prepared for when we try to make consistent classifications of proverbs – especially international proverb lore. The material as such is actually impossible to delimit to exclusive classification. Overlapping is more like

strength of proverbial effectiveness. That is why we need references both to explicit terminology and to implicit meanings.

Kuusi's database and typology is full of this kind of cross-references.

Connections between the next three proverb types in Kuusi's typology add possibilities to interpret the chosen proverb type in comparison not only with the other proverbs of its subgroup, but with those in other categories. The first proverb type A1a16 is an example, how proverbs can be taken as observations and knowledge of human conditions. The second proverb type T1j 20 is a cross-reference to the previous connecting it with the idea of the proverb image "Coming and going with water". It refers to occasional or accidental gain. The third type T2a 36 parallels the metaphors of 'water' and 'wind'.

(9) *What has come with water, goes with water.* (A1a16, a translation from Maltese)

(10) *By river come, by water go.* (T1j 20, a translation from Chinese)

Was mit der Flut kommt, geht mit der Ebbe wieder weg.

Quel che vien di ruffa in raffa, se ne va di buffa in baffa.

(11) *What comes with the wind, leaves whistling* (T2a 36, English)

The next proverb type (12) is primarily bound or connected to a proverb subgroup concerning questions about age. The cross-references to other subgroups give at least four possibilities for studying its connotations. If we had the social contexts where this proverb was used at our disposal we could better interpret its impression.

(12) *Lazy youth makes (for a) lousy (old) age.* G6c 13

- | | |
|-------------------------------------|---|
| 1. the element of age | G6c If not x when young, then not y, when old |
| 2. X will produce a corresponding y | C2a a reward thanks to action |
| 3. x= not work – y= no food | M7a diligence and efforts will be rewarded |
| 4. honour and shame | F1b honesty will be rewarded |
| 5. the element of time | T3c preparing for the future |

6 Applications of Kuusi's typology

Some scholars have used Matti Kuusi's typology or at least reacted to it. George Goce Mitrevski writes about the classification of Macedonian Proverbs in an electronic database and notices (correctly) that Kuusi's typology is an extendable system. He appreciates the possibility of 'adding additional themes and classes, as well as of expanding the depth of the coding system'. Mitrevski emphasizes their 'need to index a variety of relevant data about each proverb, and to be able to construct complex queries when selecting proverbs with specific characteristics'. That is why their 'data in the database is defined so that it can be reorganized and accessed in a number of different ways depending on specific needs (Mitrevski 2007: 19)'. Their aim has been to accomplish projects of contrastive paremiology in the Balkans.

The most recent study in which Kuusi's typology has been applied is an analysis of the forms and functions of overstatement in Kurdish proverbs (Muhedeen and Ahmed 2023: 61). Scholars found five out of thirteen categories suitable for a model of their study. These were the main themes: 'Basic observations and socio-logic', 'Concepts of morality', 'Agreements and norms', 'Coping and Learning', 'Time and Sense of time'. There are two examples of proverbs connected to the 'Coping and Learning' group. One of them is

(13) *'He has forded a thousand waters, his ankle has not become wet.'*¹⁵

The authors comment that an overstatement is used for the sake of emphasis to show how experienced the person is. A linguist would recognize only the stylistic device of hyperbole.

The next proverb example has been situated in the 'Time and Sense of time' group:

(14) *'In (all) seven heavens he has no star.'*

'The implied meaning of the proverb is that a person is so unlucky that they enjoy no luck in anything.' Authors interpret the use of overstatement so that the listener will neither take the meaning in a literal sense nor as a lie. They continue explaining: 'If this proverb is uttered in a situation where a person faces lots of opportunities and loses all of them in turn, humour will be the function of the used trope' (Ibid. 64–65).

They seem to apply Kuusi's main themes to Kurdish proverbs from the point of view of social interaction between the speaker and the listener.

The general problem in classifying proverbs lies in determining to what extent it is justified to typify proverbs from different cultures on common semantic criteria. Even if we could agree on defining proverbs as impersonal yet impactful utterances or speech in which one appeals to the authority of a third person (see Mukařovský 1983 [1971]), what criteria can we use to categorise these expressions under common key terms? Do we know which expressions belong only to ritual use, which are euphemisms or which are completely out of use?

Accessing Kuusi's material and typology was the basis on which the author of this article worked on another project twenty years ago. We had the possibility to get closer to the roots of our own culture. Scholars in Finland and Estonia continued working on a large project of contrastive paremiology from the 1980s concerning the proverbs of Fenno-Ugric people around the Baltic Sea (Kuusi et al. 1985). *Proverbia Septentrionalia* II (not yet published) is a very detailed corpus. The database is the result of comparing archival material collected in museums of literature and archives of tradition both in Helsinki and in Tartu.

When the distribution of Fenno-Ugric proverb types in the subgroups of the themes of Kuusi's typology were compared to the percentages in the international database,

¹⁵ The Kurdish proverbs are from the collection: MacKenzie, N. 1970: Some Kurdish Proverbs. *Iran*, Vol. 8, 105–113. <http://www.jstor.org/stable/4299636> (retrieved 18.12.2023).

the amounts proved to be almost the same in both corpora. Only the main ‘practical knowledge of nature’ group was more represented in the Fenno-Ugric corpus than in the global proverb corpus. The explanation may lie in the abundance of weather proverbs and practical advice of local peasant cultures.

The author of this article primarily adopts a social psychological approach, but the key terms and theories concerning paremiology come from the fields of folkloristics and linguistics. The practical need to use proverbial databases often arises from translation studies and cognitive linguistics. Whatever the motivation for studying proverbs, there is much work to be done to digitise this genre of oral tradition and its interpretation. Both archival and field research are important for identifying the characteristics of the proverb genre and finding out how people still create novel compositions or anti-proverbs on the basis of proverbial patterns.

7 Summary

Matti Kuusi’s proverb typology is grounded in the basic manners of human perception. It depends on the aim of specific research how the typology is used. The subject matter of the database – the proverb types with their variants – are like hypotheses concerning the similarity and simultaneity of proverbs in different cultures.

An important quality of typologies and the databases connected with them should be their flexible application to the needs of users. That is why any system of cross-references between proverbs, anti-proverbs and their classification should be developed into a dynamic network. There is much to do in contrastive paremiology, cultural and translation studies and folkloristics to equip proverbial databases with contextual evidence.

We should seek the best ways of obtaining reliable translations, authentic experience and contextual data from different cultures in order to find out which proverb ideas can be called ‘global’.

That is why the typology must be used as a grid of relations between different levels of interpretations. Kuusi did not receive much feed-back about his system in the 1970s, although he wrote in 1970:

An index of international proverb types cannot be successful if it is the work of a single man.

It presupposes international cooperation, consultation and the consolidation of divergent scholarly interests (Kuusi 1970a).

Nowadays, it is widely recognised that the synergy of research capacities is necessary for producing successful scholarly work. Depending on the research requirements, there are now better possibilities to combine or supplement different typologies. Krikmann’s typology (1998) is useful for those who wish to survey the linguistic structure of any proverb corpus based on cultural meanings. There could also be special cultural nuances influencing the definition of new semantic classes in the Matti Kuusi typology.

Kuusi and Permyakov were system builders who drew on their exceptional personal memory capacity and we can try to further their work using new technology.

References

- Biber, Hanno, 2015: AAC-Fackel. Das Beispiel einer digitalen Musteredition. In: Baum, Constanze; Stäcker, Thomas (eds.), *Grenzen und Möglichkeiten der Digital Humanities*. DOI: http://dx.doi.org/10.17175/sb001_019.
- Cram, D., 1994 [1983]: The linguistic status of the proverb. In: Mieder, Wolfgang (ed.), *Wise words: Essays on the proverb*. New York: Garland, 73–97.
- Goodwin, Paul; Wenzel, Joseph, 1981: Proverbs and Practical Reasoning: A Study in Socio-Logic. In: Mieder, Wolfgang; Dundes, Alan (eds.), *The Wisdom of Many: Essays on the Proverb*. Garland Publishing, 140–160.
- Hasan-Rokem, Galit, 1982: *Proverbs in Israeli Folk Narratives: A Structural Semantic Analysis*. Helsinki: Academia Scientiarum Fennica (Folklore Fellows Communications, 232).
- Hasan-Rokem, Galit; Kats, Pavel, 2009: From a Structural Semantic Analysis towards a Computational Proverb Classification Framework. In: McKenna, Kevin (ed.), *The Proverbial “Pied Piper”. Festschrift Volume of essays in Honor of Wolfgang Mieder on the Occasion of his Sixty-fifth Birthday*. New York: Peter Lang, 111–125.
- Hrisztova-Gotthardt, Hrisztalina; Aleksa Varga, Melita (eds.), 2014: *Introduction to Paremiology. A comprehensive guide to proverbs studies*. Warsaw, Berlin: De Gruyter. DOI: <https://doi.org/10.2478/9783110410167>.
- Krikmann, Arvo, 1998: On the relationships of the rhetorical, modal, logical, and syntactic planes in Estonian proverbs. Part 2. *Folklore. Electronic Journal of Folklore* 8, 51–99. DOI: <http://dx.doi.org/10.7592/FEJF1998.08.syntmod2>.
- Kuusi, Matti (ed.), 1988: *Rapatessa roiskuu. Nykysuomen sananparsikirja* [When you splatter it splashes. A collection of proverbial phrases of current Finnish]. Helsinki: SKS.
- Kuusi, Matti et al. (eds.), 1985: *Proverbia Septentrionalia*. Helsinki: Academia Scientiarum Fennica (Folklore Fellows Communications, 236), 29–36.
- Kuusi, Matti, 1978: *Suomalaista, karjalaista vai savokarjalaista?* [Finnish, Karelian or Savo-Karelian]. Helsinki: SKS.
- Kuusi, Matti, 1972: Towards an International Type-System of Proverbs. *Proverbium* 19.
- Kuusi, Matti, 1970a: How can a Type-Index of International Proverbs be outlined? An Experiment and Five Questions. *Proverbium* 15, 473–476.
- Kuusi, Matti, 1970b: *Ovambo proverbs with African parallels*. Helsinki: Finnish Academy of Science and Letters (Folklore Fellows Communications, 208).
- Kuusi, Matti, 1966: Ein Vorschlag für Terminologie der parömiologischen Strukturanalyse. *Proverbium* 5, 97–104.
- Kuusi, Matti, 1953: *Vanhan Kansan Sananlaskuviisaukset* [Proverb wisdom of the old folks]. Helsinki: WSOY.
- Kuusi, Matti; Krikmann, Arvo; Joalaid, Marje; Kokare, Elsa; Laukkanen, Kari; Leino, Pentti; Mälk, Vaina; Sarv, Ingrid, 1985: *Proverbia Septentrionalia. 900 Balto-Finnic Proverb-Types with Russian, Baltic, German and Scandinavian Parallels*. Helsinki: Academia Scientiarum Fennica (Folklore Fellows Communications, 236).

- Lauhakangas, Outi, 2023: Matti Kuusi's Typology in the Light of Contemporary Use of Proverbs. In: Kotova, Marina Y.; Lauhakangas, Outi (eds.), *Proverbs Are Never Neutral*. Palgrave Macmillan, Cham. DOI: https://doi.org/10.1007/978-3-031-32646-2_3.
- Lauhakangas, Outi, 2021: Better Familiar Bad than Unfamiliar Good. Re-evaluation of Common values by Proverb Use. In: Szerszunowicz, Joanna (ed.), *Reproducibility of multiword expressions in paremiological and linguo-cultural studies*. University of Białystok Publishing House (Intercontinental Dialogue on Phraseology, 10), 159–174.
- Lauhakangas, Outi, 2004: *Puheesta ihminen tunnetaan. Sananlaskujen funktiot sosiaalisessa vuorovaikutuksessa* [A man is known by his words. The functions of proverbs in social interaction]. Helsinki: The Finnish Literature Society.
- Lauhakangas, Outi, 2001: *The Matti Kuusi International Type System of Proverbs*. Helsinki: Academia Scientiarum Fennica (Folklore Fellows Communications, 275).
- Mac Coinnigh, Marcas, 2014: Structural aspects of proverbs. In: Hrisztova-Gotthardt, Hrisztalina; Aleksa Varga, Melita (eds.), *Introduction to Paremiology. A comprehensive guide to proverbs studies*. Warsaw, Berlin: De Gruyter, 112–132.
- Mieder, Wolfgang, 2012: "Think outside the box": Origin, Nature, and Meaning of Modern Anglo-American Proverbs. *Proverbium: Yearbook of International Proverb Scholarship* 29, 137–196.
- Mitrevski, George, 2007: *On the Classification of Macedonian Proverbs in an Electronic Database*. XIV International Congress of Slavists.
- Muhedeen, Bekhal Latif; Ahmed, Shayma Mahmud, 2023: Pragmatic Study of Overstatement in Kurdish Proverbs. *Journal of University of Raparin*, 45–71. DOI: [http://dx.doi.org/10.26750/Vol\(10\).No\(3\).Paper3](http://dx.doi.org/10.26750/Vol(10).No(3).Paper3).
- Mukařovský, Jan, 1983 [1971]: The Proverb as Part of Context. Selected passages from Mukařovský – Penfield, Joyce, *Communicating with quotes: The Igbo case*. Appendix B. Westport CT: Greenwood Press, 96–104.
- Noy, Natasha F.; McGuinness, Deborah L., 2001: *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory. Online: http://protege.stanford.edu/publications/ontology_development/ontology101.pdf.
- Permyakov, Grigory, 1979 [1970]: *From proverb to folk-tale*. Moscow.
- Permyakov, Grigory, 1968 [1961]: *Izbrannye posloviy i pogovorki narodov vostoka*. Moskva.
- Villers, Damien, 2022: What makes a good proverb? On the birth and propagation of proverbs. *Lexis. Journal in English Lexicology* 19. DOI: <https://doi.org/10.4000/lexis.6383>.
- Zhitomirsky-Geffe, Maayan; Prebor, Gila; Bloch, Orna, 2017: Improving proverb search and retrieval with a generic multidimensional ontology. *Jasist (Journal of Association for Information Science and Technology)* 68/1, 141–153.



Dataset of stylistic features of Croatian folklore genres

Nabor podatkov o slogovnih značilnostih hrvaških folklornih žanrov

This paper aims to present the creation of a small dataset derived from the computational analysis of Croatian folklore genres. The data was extracted from a collection of charms, counting-out rhymes, tongue twisters, blessings, curses and proverbs using script for automated syllabification and n-gram extraction. The dataset consists of syllable and n-gram ratios and the numerical representation of other stylistic features.

• **Keywords:** feature extraction, machine learning, computational stylistics, Croatian folklore genres

Namen prispevka je predstaviti oblikovanje majhnega nabora podatkov, ki izhaja iz računalniške analize hrvaških folklornih žanrov. Podatki so bili pridobljeni iz zbirke zagovorov, izštevank, lomilcev jezika, blagoslovov, kletvic in pregovorov z uporabo skript za samodejno silabizacijo in ekstrakcijo n-gramov. Nabor podatkov sestavljajo razmerja zlogov in n-gramov ter številčna predstavitev drugih slogovnih značilnosti.

• **Ključne besede:** ekstrakcija značilnosti, strojno učenje, računalniška stilistika, hrvaški folklorni žanri

1 Introduction

As part of an ongoing research effort which combines the stylistic analysis of Croatian oral literature and natural language processing methods, the authors have constructed a dataset based on a collection of Croatian folklore genre texts. Previous research (Nikolić and Bakarić 2016) has shown that the automated classification of these types of text is possible when observing the text at a sub-word level, the level of the syllable and the phoneme n-gram. Computational stylistic (Rybicki et al. 2016) analysis allows for further examination of other stylistic features such as rhyme (Hoorn et al. 1999) and other figures of style.

The creation of the dataset is part of research for a doctoral thesis, which aims to design a model for classifying Croatian folklore genres based on stylistic features and the use of machine learning classification algorithms.

2 Description of the corpus

The collection of charms, counting-out rhymes, tongue twisters, blessings and curses are part of the Archive of manuscript collections of the Chair for Croatian oral literature at the University of Zagreb's Faculty of Humanities and Social Sciences. An additional collection of modern Croatian language texts containing prose and newspaper articles

was added to the collection as a test baseline for classification purposes.¹ The collection of texts was then processed and normalized to create a unified corpus. Since the focus of the research is on short rhetorical genres², which can range in length from 15 to 200 characters, other text forms were adjusted to match them in size in order to make them comparable.

The corpus comprises 66 charms, 321 counting-out rhymes, 76 tongue twisters, 125 proverbs, 72 blessings and 342 curses, totalling 8,700 word tokens. Additionally, it includes 25,000 tokens of modern language text.

3 Selection of stylistic features

Even a casual glance at folklore genres detects a certain order and structure in their sound and syllable organization, while a more focused analysis reveals that some genres are more structured than others, especially in terms of sound and syllable repetition. The original purpose of this dataset and the tools used to create it was to measure the levels of sound repetition with the purpose of determining euphony in the poetic expression of poets such as A. G. Matoš (Nikolić and Bakarić 2018).

The first observed stylistic feature group was the repetition coefficient of sounds. The method used to measure the coefficient requires that the segments of the analysed corpora be of similar size for the comparison of results to be meaningful. It is impossible and thankless to talk about constants because the repetition coefficient only makes sense when observed in relation to other elements (due to the mathematical model). It can be further divided into two approaches for measuring repetition – n-gram repetition, which observes sounds (or the approximation of sounds through characters), and syllable repetition. N-gram repetition is independent of any kind of prosodic knowledge and therefore much simpler to measure. Syllable repetition requires that the text is separated into syllables, which is not a trivial task for the Croatian language. The problem of syllabification has several levels but can be described by a rule-based algorithm, which first solves specific and rare syllable border problems and then moves on to more general and common syllable types. Issues that affect the syllabification results are related to the difference between spoken and written text and, even though

¹ Although the folklore texts were recorded in all of Croatia's regions (with relatively high dialectal diversity), the phonological encoding of the texts was comparable to the encoding for the texts written in standard Croatian language. In most cases the original folklore records were not transcribed using narrow phonological transcription because the variations in vowel quality were not relevant for the meaning.

² *Rhetorical genres* is a term used by many Croatian folklorists, especially those with a philological background, to refer to a group of relatively small genres, which are characterized by elements of persuasion and verbal playfulness (Čubelić 1970; Kekez 1996, 1998; Botica 1995, 2013). Verbal charms, toasts, counting-out rhymes, tongue-twisters, blessings and curses are considered to be typical rhetorical genres, which use their aesthetic qualities to produce specific communicative effects (Nikolić 2019: 70).

Croatian spelling is mostly phonetic, this should be taken into consideration when preparing the input text. Repetition of sounds contributes to euphony, in the form of alliteration, assonance and rhyme (Bishop 1985: 345), and is a distinguishing feature when comparing and classifying text.

An additional stylistic feature used to detect levels of euphony is the pronunciation difficulty coefficient. This feature looks at two phenomena, which contribute to the difficulty of pronunciation within a given text. The authors focused on the syllabic r and consonant clusters within syllables. Reporting the occurrence of these phenomena also requires syllabification and the additional analysis of syllable quality, which is an integral part of the syllabification algorithm. Although essentially quantitative, the pronunciation difficulty coefficient still includes some qualitative elements. This concerns the counting of specific elements previously recognized as non-euphonic (syllabic r and consonant groups), but still neglects the vocal composition within the consonant groups. The coefficient does not differentiate between sets of voiceless and voiced consonants. A thorough analysis of tongue twisters confirmed a high proportion of sets with voiceless consonants (Nikolić 2013: 119–120, 2019: 200), which is an indisputable indicator of pronunciation difficulty and affects the overall reduced impression of euphony.

Computational stylistics is described as a data-driven field, which promises new insights into texts through digital methods (Herrmann et al. 2021). However, to make its contribution to analysing euphony more interesting and useful to literary historians and prosodists, it is necessary to link the results of objective measurements at the phonetic level of the text with the content level. The results of such research could confirm (or perhaps refute) euphony as a distinguishing feature between genres, or even between individual authors of written literature. Using a dataset such as the one presented here as a distant reading tool can assist in detecting phenomena, which can then be focused on and further analysed.

4 Feature extraction

Corpus preparation included changing characters to lowercase and transcribing text with the goal of having one character represent one phoneme. This is important for accurate syllabification, one of the steps of feature extraction. Simple regular expressions were used to replace the digraphs lj, nj and dž with ǝ, ń and ž respectively. A similar method was used to mark the syllabic r as any r between two consonants. An additional attempt was made to solve the difference in spelling and pronunciation of accentuated and non-accentuated words by removing the spaces between the words in question. However, it is still unclear to what extent this impacts the final results.

The corpus was then processed in order to extract the stylistic features of each entry. Syllabification was carried out using the custom syllabification algorithm written in Python (Bakarić and Nikolić 2017), mentioned in the previous paragraph. Using the algorithm, the authors were able to create a statistical overview of the genres and extract features from individual entries. They were able to observe the frequency of syllable types for each of the genres, as can be seen in Table 1. C represents consonants and V vowels. The syllabification algorithm differentiates between consonant groups and uses them to define some of the syllabification rules.

	Charms	Counting-out rhymes	Curses	Blessings	Proverbs	Tongue twisters
CV	0.60	0.65	0.65	0.61	0.67	0.61
V	0.15	0.06	0.10	0.09	0.07	0.05
CCV	0.14	0.11	0.17	0.18	0.16	0.15
CVC	0.08	0.13	0.06	0.09	0.06	0.13
CCVC	0.01	0.02	0.01	0.01	0.01	0.04
VC	0.01	0.02	0.01	0.01	0.01	0.00

Table 1: Frequency of syllable types across genres.

The syllabification script was extended to include an algorithm, which calculates the type/token ratio for unigrams, bigrams, trigrams, syllables, meta-syllables, consonant clusters and syllabic /r/ coefficients and rhyming coefficients. The script can be further developed to include other features, such as syllable quality (e. g. open/closed syllable ratio) or their frequency, which can be seen in Table 2.

	Charms	Counting-out rhymes	Curses	Blessings	Proverbs	Tongue twisters
1.	i	ka	da	ti	je	po
2.	u	je	o	i	ne	pe
3.	ko	ti	ti	bo	na	je
4.	o	li	la	go	i	ko
5.	su	ci	bo	o	u	ka
6.	ne	i	te	vi	ni	o
7.	na	ko	gda	da	ti	pa
8.	ka	na	i	ne	ma	na
9.	a	se	u	bla	ka	pi
10.	si	po	se	bog	se	la

Table 2: Top 10 syllables across genres.

Phonostylistics, a sub-discipline of stylistics, delves into stylistic effects at the phonetic and phonological levels (Trubetzkoy 1971). It analyses literary texts with a primary focus on the repetition of sound patterns on multiple levels: the repetition

of metrical patterns, syntactic parallelism, the repetition of motifs or key words and the repetition of certain sounds or sound sequences (alliteration, assonance, poetic homophones), as can be seen in Figure 1. As mentioned earlier, n-gram and syllable repetition represent two approaches to detecting stylistic features caused by sound sequences. While n-gram analysis does not require language knowledge, it is unable to observe metrical patterns and sound sequences on the same level as the syllable analysis approach. It allows a comprehensive qualitative exploration of syllables, including analysis of consonant clusters, the syllabic r, open and closed syllable ratios, to mention a few.

The design of the syllabification algorithm allows for further qualitative analysis with minimal intervention. The syllabification rules are based on dividing consonants into groups according to their production and sonority. This approach offers a finely-tuned qualitative method that might yield further insights into the nature of phonostylistic phenomena.

<p><i>Ni u moru mjere, ni u mački vjere.</i> (alliteration) <i>Pusti koku na policu, ona će i na stolicu.</i> (assonance) <i>Od zboru do tvora ima prostora.</i> (poetic homophones)</p>
--

Figure 1: Example of sound sequences in proverbs.

5 Dataset structure

The main idea behind the creation of this dataset was to describe stylistic features as numerical values, which can then be analysed, compared and used as a testing ground for data driven analysis at the sub-word level.

The dataset is formatted as a table with the following fields (from left to right):

Text entry – original text entry prior to processing (type: text)

The input text entry is prepared for processing in the previous step. Preparation includes transcription of digraphs, switching to lowercase and marking the syllabic r. The input text is fed to the processing algorithm as a plain text unicode UTF-8 standard file. Individual inputs are separated by new lines and each genre is contained in a separate file.

Unigram – type/token ratio of unigrams or single characters subtracted from 1 (type: number, interval 0-1)

The algorithm determines the frequency of single characters, which, at this point, represent sounds. The output is the type/token ratio which is then subtracted from 1 in order to represent the repetition coefficient.

$$K_{Unigram} = 1 - \frac{Type_{Unigram}}{Token_{Unigram}}$$

Bigram - type/token ratio of bigrams, two consecutive characters in a word, subtracted from 1 (type: number, interval 0-1).

Here we determine the type/token ratio of pairs of sounds in the input text. The pairs are created by looking at consecutive pairs of characters with overlap. For example, an input text consisting of six characters (without spaces) would produce five distinct bigrams. The number of bigram tokens is therefore n-1, where n is the number of characters in the input text. Subtracting the type/token ratio from 1 yields the bigram repetition coefficient.

$$K_{Bigram} = 1 - \frac{Type_{Bigram}}{Token_{Bigram}}$$

Trigram – type/token ratio of trigrams or three consecutive characters in a word (type: number, interval 0-1).

The trigram determines the type/token ratio of sequences of three sounds in the input text. Character trigrams are created by observing consecutive triplets of characters with overlap. Therefore, the number of trigram tokens is n-2, where n is the number of characters in the input text.

$$K_{Trigram} = 1 - \frac{Type_{Trigram}}{Token_{Trigram}}$$

Syllable – type/token ratio of all syllables in an entry subtracted from 1 (type: number, interval 0-1)

The syllable repetition coefficient represents the type/token ratio of syllables subtracted from 1. The algorithm calculates the type/token ratio after the syllabification procedure.

$$K_{Syllable} = 1 - \frac{Type_{Syllable}}{Token_{Syllable}}$$

Meta-syllable – type/token ratio of all syllables transcribed as only vowels and consonants (type: number, interval 0-1)

The meta-syllable coefficient is calculated after the syllabification and transcription of syllable elements (consonants are transcribed as C, vowels as V). Here all syllables are reduced to categories such as CV, CCV, V, etc. After that, the type/token ratio is calculated and subtracted from 1.

$$K_{Meta-syllable} = 1 - \frac{Type_{Meta-syllable}}{Token_{Meta-syllable}}$$

CC-R – ratio of syllables containing consonant clusters or syllabic /r/ (type: number, interval 0-1)

The pronunciation difficulty coefficient is calculated by counting the syllables, which contain consonant clusters (a group of more than one consecutive consonant in a syllable, such as CCV) and syllables which contain the syllabic r. We then calculate the type/token ratio, where type is the total number of syllables containing either consonant clusters or the syllabic r and the token is the total number of syllables.

$$CCR = \frac{Type_{CCR}}{Token_{CCR}}$$

Rhyme coefficient – ratio between all the words and the rhyming words in an entry (type: number, interval 0-1)

The rhyme coefficient was added to the algorithm later and is not part of the original script (available on Github (Bakarić and Nikolić 2017)). It calculates the ratio between words that end with the same syllable and the total number of words in a text entry.

Length – length of text entry under 1 (type: number, positive integer)

The length of text entry was calculated after processing using Excel's len() function. It includes interpunction and spaces as contained in the original entry. This auxiliary field is used for filtering entries of similar or identical lengths, which is important for separating the dataset into comparable subsets.

Class – genre of text entry (type: text, list of names [charm, counting-out rhyme, ..., modern language text])

The last field contains the entry class or genre. Ultimately, using the data described in the previous fields, we will train a machine learning model that will be able to classify the entry by genre.

The structure of the dataset allows for statistical analysis and preparation for further processing. It should be taken into consideration that machine learning classification algorithms are very sensitive to data distribution (Hastie et al. 2009) and most of the values in this dataset are skewed in some way. The analysis of the dataset can show which parameters of machine learning algorithms should be considered and how the data should be approached.

Table 3 is a representation of the algorithm output for a single entry, in this case a tongue twister. The algorithm uses this output to create the dataset, but it can be modified to produce several different outputs beyond the dataset described in this paper.

petar petru plete pet u	
unigram	e:5 p:4 t:4 r:2 u:2 a:1 l:1 :1
bigram	et:4 pe:3 ta:1 ar:1 tr:1 ru:1 pl:1 le:1 te:1 t :1 u:1
trigram	pet:3 eta:1 tar:1 etr:1 tru:1 ple:1 let:1 ete:1 et :1 t u:1
syllable	pe:3 tar:1 tru:1 ple:1 te:1 t u:1
meta syllable	kv:4 kkv:3 kvk:1

Table 3: Example of the extracted data for a tongue twister.

The authors have made the dataset publicly available for download in CSV and XLSX formats using a data repository (PUH 2022) (Bakarić and Nikolić 2023). The syllabification Python script is publicly available on GitHub (Bakarić and Nikolić 2017). There are plans to extend the dataset with lyrical and epic poetry, as well as poetry written by named authors.

5 Conclusion

The dataset, as seen in Table 4, was primarily designed for the machine classification of Croatian folklore texts. However, the described methods and the dataset itself can be used as a basis for drawing general conclusions on stylistic features in other texts, a direction the authors plan to take.

The presented process depicts the steps taken to build a dataset of numerically described stylistic features of selected Croatian folklore texts. Building a dataset is the crucial first step of many natural language processing research activities. Its design and structure can have a great impact on research results. The dataset presented here is limited in scope but can serve as a starting point for future efforts in the computational stylistic analysis of Croatian language texts.

Entry	Uni-gram	Bigram	Trigram	Sylla-ble	Meta-syll.	CC and R	Length	Class
Enci benci nakamenci šija bija kompanija van!	0,66	0.35	0.25	0.24	0.82	0.00	45	count- ing-out rhyme
Četrstočetrdesetčetiri čavke čuče načamcu ičopaju čičke.	0,65	0.22	0.05	0.09	0.83	0.22	57	tongue twister
Vrake drpo išćipo cjelu noć.	0,25	0.05	0.00	0.10	0.60	0.50	29	curse
Višeti Bog dao negunjega želila.	0,42	0.05	0.00	0.00	0.77	0.00	32	blessing

Table 4: Example of the dataset.

References

- Bakarić, Nikola ; Nikolić, Davor, 2017 : Croatian syllabification script. Online: <https://github.com/nbakaric/Croatian-syllabification> (retrieved 3.12.2023).
- Bakarić, Nikola; Nikolić, Davor, 2023: Stylistic features of Croatian folklore genres. Online: <https://puh.srce.hr/s/ygbceraCJCnprPtT> (retrieved 3. 12. 2023).
- Bishop, Lloyd, 1985: Euphony: A New Method of Analysis. *Language and Style* 18/4, 343–362.
- Botica, Stipe, 1995: *Hrvatska usmenoknjiževna čitanka*. Zagreb: Školska knjiga.
- Botica, Stipe, 2013: *Povijest hrvatske usmene književnosti*. Zagreb: Školska knjiga.
- Čubelić, Tvrtko, 1970: *Usmena narodna retorika i teatrologija*. Zagreb: s. n.
- Hastie, Trevor; Tibshirani, Robert; Friedman, Jerome, 2009: Model Assessment and Selection. In: *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. New York: Springer. 219–257.
- Herrmann, J. Berenike; Jacobs, Arthur M.; Piper, Andrew, 2021: Computational Stylistics. In: Kuiken, Don; Jacobs, Arthur M. (eds.), *Handbook of Empirical Literary Studies*. Berlin, Boston: De Gruyter. 451–486.
- Hoorn, J. F.; Frank, S. L.; Kowalczyk, W.; van der Ham, F., 1999: Neural network identification of poets using letter sequences. *Literary and Linguistic Computing* 14/3, 311–338. DOI: <https://doi.org/10.1093/lc/14.3.311>.
- Kekez, Josip (ed.), 1996: *Poslovice, zagonetke i govornički oblici*. Zagreb: Matica hrvatska.
- Kekez, Josip, 1998: Usmena književnost. In: Škreb, Zdenko; Stamać, Ante (eds.), *Uvod u književnost: Teorija, metodologija*. Zagreb: Nakladni zavod Globus. 133–192.
- Nikolić, Davor, 2013: *Fonostilistički opis hrvatske usmenoknjiževne retorike*. Doctoral thesis. Zagreb: University of Zagreb.
- Nikolić, Davor, 2019: *Između zvuka i značenja: fonostilistički pristup hrvatskim usmenoretoričkim žanrovima*. Zagreb: Disput.
- Nikolić, Davor; Bakarić, Nikola, 2016: What Makes Our Tongues Twist?: Computational Analysis of Croatian Tongue-Twisters. *Journal of American Folklore* 129/511, 43–54. DOI: <https://doi.org/10.5406/jamerfolk.129.511.0043>.
- Nikolić, Davor; Bakarić, Nikola, 2018: Korelati eufonije u Matoševim sonetima. In: Botica, Stipe et al. (eds.), *Šesti hrvatski slavistički kongres. Zbornik radova. Vol. 2*. Zagreb: Hrvatsko filološko društvo, 801–811.
- PUH data storage and management, 2022. University computing centre-SRCE. Online: <https://www.srce.unizg.hr/en/puh> (3.12.2023).
- Rybicki, Jan; Hoover, David L.; Eder, Maciej, 2016: Computational Stylistics and Text Analysis. In: Crompton, Constance; Lane, Richard J.; Siemens, Ray (eds.), *Doing Digital Humanities: practice, training, research*. London, New York: Routledge, 123–144.
- Trubetzkoy, Nikolay Sergejevich, 1971: *Principles of phonology*. Berkley, Los Angeles: University of California Press.





II

Folkloristična medžanrskost
Intergeneric relations in folklore



Žanrske, pomenske in zgradbene značilnosti bovških pregovorov in (primerjalnih) frazemov z leksemsko sestavino *rit*

Genre, semantic, and structural characteristics of Bovec proverbs and (comparative) idioms with the lexeme component *rit* (*ass, butt*)

V prispevku so prikazani žanrska raznolikost, pomen, funkcije in jezikovno-strukturne značilnosti bovškega frazeološkega gradiva, in sicer na primerih somatskih paremij (pregovorov) in frazemov s sestavino *rit*, ki je poleg leksemov *bog*, *hudič* (*zlodej*), *nos* in *drek* med najpogostejšimi v zbirki. *Rit* je v narečnem govoru nevtralen leksem za ta del telesa, toda ko skupaj s sobesedilom tvori paremiološki izraz, prestopi iz nevtralnega v metaforični jezik. Zaradi lege, oblike, funkcije in nena zadnje simbolike ima ta del telesa bogato semantiko in je pripraven za vključevanje v paremiološke strukture, kar izpričuje bogato gradivo, 107 enot pregovorov, primerjalnih in drugih frazemov, ki so bili povečini zapisani iz diskurza samega, v gramatični obliki, kot so bili izrečeni v pragmatičnem kontekstu.

♦ **Ključne besede:** bovški narečni govor, paremiologija, primerjalni frazemi, pregovori, leksem *rit*, folklor, besedno izročilo

The paper presents the genre diversity, meaning, functions, and linguistic-structural characteristics of the Bovec paremiological material, using examples of somatic paremia and idioms with the constituent *rit* (Eng. *ass, butt*) which is among the most frequent in the collection, along with the lexemes *bog* (*god*) *hudič/zlodej* (*devil*), *nos* (*nose*) and *shit*. *Rit* (*ass*) is a neutral lexeme for this part of the body in dialectal speech. Still, when it forms a paremiological expression together with its context, it passes from the neutral to the category of stylistic language. Due to its position, shape, function, and, last but not least, symbolism, this body part has a rich semantics and is ready to be integrated into paremiological structures, as evidenced by the rich material, 107 units of proverbs, comparative and other phrases, most of which were written down from the discourse itself, in grammatical form, as they were uttered in a pragmatic context.

♦ **Keywords:** Bovec dialect, paremiology, phrases, comparative idioms, proverbs, lexeme *ass / butt*, folklore, oral tradition

1 Opis zbirke, metode zbiranja in zapisa ter žanrska opredelitev

V rastoči zbirki bovškega¹ frazeološkega in folklornega gradiva so pregovori, reki, primerjalni obrazci, zdravstvena in kmetijska napotila, velerizmi, nekonvencionalni odgovori, uganke, zdravice, blagoslovi, pozdravi, kletvice, grožnje, raznovrstni vzkliki, pa tudi vraže, kratke šaljive rime, izštevanke, zbadljivke, krajše folklorne molitve ter

¹ Bovško je mišljeno kot celotno območje, kjer se govori bovški narečni govor, kar ustreza tudi mejam Občine Bovec.

druge staln(ejš)e metaforične/ekspresivne besedne zveze ipd. Zbirka nastaja od srede devetdesetih let 20. stoletja, in sicer priložnostno, največkrat kot stranski produkt terenskih raziskav (pripovednega) folklornega izročila; narečje je namreč jezik, v katerem informatorji ubesedujejo lokalno slovstveno folkloro. Poleg poklicnega interesa me k zapisovanju posebnih/značilnih narečnih besed, besednih zvez, stavčnih struktur spodbuja osebna naklonjenost do svojega domačega narečja in s tem povezana skrb, da bi se za zanamce ohranila bogata in v marsičem posebna in edinstvena jezikovna in folklorna dediščina. Ta se namreč zaradi spremenjenega načina življenja, napredka, migracij, vpliva šolstva, medijev, tehnologije itd. naglo spreminja in pozablja. Tudi zato sta bila bovški narečni govor in bovška slovstvena folkloro vpisana v nacionalni register nesnovne dediščine (gl. SV1, SV2). Zbirka v delovni različici vsebuje nekaj čez 2000 enot, in čas je že, da se jo kritično pregleda, dopolni in uredi po principih znanstvene akribije, da bo gradivo uporaben vir za jezikoslovne, folkloristične, etnološke in druge raziskave (narečne) kulture.

Večina gradiva je prišlo v rokopisno zbirko na naslednje načine: A) z zapisom iz spontanega narečnega diskurza; v taki situaciji se najbolje obnese takojšen zapis na listek, v telefon, terenski zvezek. Pri zapisu iz diskurza so zavzeto sodelovali tudi člani moje ožje družine iz Bovca² B) iz folklornih, predvsem pripovednih besedil, ki sem jih dokumentirala med terenskim delom na Bovškem sama ali iz zapisov drugih raziskovalcev/zapisovalcev; C) iz beleženja vzporednic gradiva, na katerega sem naletela v drugih, predvsem primorskih virih (npr. v Podberšič 2007; Dolenc: RV1, objave v Trinkovem koledarju); Č) z vključitvijo gradiva, ki so mi ga posredovali domačini, posamezniki, ki se zanimajo za ohranjanje lokalnega narečnega govora; D) v zadnjih nekaj letih je bogat vir socialno omrežje Facebook, kjer je deluje skupina »Pu buško« (Po bovško) in »Buške pšade an kej wes žuli« (Bovške čenče in kaj vas žuli).³ V teh skupinah domačini komentirajo aktualne in pretekle lokalne dogodke, stare fotografije itd.⁴

Gradivo je zapisano z različnimi metodološkimi pristopi in načini grafičnega zapisa: v fonetičnem dialektološkem ter v poenostavljenem narečnem zapisu; ta dva zapisa prevladujeta na gradivu, ki sem ga dokumentirala sama ali so ga zapisali drugi raziskovalci. Enote, ki so jih zapisali domačini, ljubitelji domačega narečja, so v amaterskem (pogosto dialektološko neustreznem) zapisu, ki pa zaradi te pomanjkljivosti niso nič manj dragocene za zbirko; ker gre za gradivo, ki je povečini

² Iz srca hvala sestri, očetu in pokojni mami.

³ Pu buško: https://www.facebook.com/groups/315961109848?locale=sl_SI
Buške pšade an kej wes žuli: <https://www.facebook.com/groups/4918620484851153/>

⁴ Ljudje v tej skupini pišejo v narečju, zapis seveda ne sledi dialektološkim pravilom, pač pa posluhu posameznika; zapisi so z običajnimi znaki abecede, včasih je dodan znak w, kjer se u ali v izgovarja dvoustnično. Večina članov polglasnik označuje s številom 3, saj je podoben obrnjenemu e-ju.

del širšega lokalnega jezikovnega znanja oz. rabe, takih enot ni težko preveriti in »popraviti« v ustrezen grafični zapis.⁵

Frazeološke enote in folklorne obrazce z Bovškega sem do nedavnega dodajala na seznam v Wordovih dokumentih, jedrni del je obsegal cca 60 strani, kar je bilo precej nepregledno. Zdaj je gradivo v procesu urejanja v Excelovo tabelo, in sicer najprej v narečnem zapisu in v gramatični obliki, v kateri je bila enota izrečena, tj. v različnih osebnih glagolskih oblikah ali nedoločniku, z različnimi slovničnimi števili in v različnih slovničnih časih, naklonih in načinih. Sledita prevod v slovenski knjižni jezik in pomensko pojasnilo; vsaki enoti je dodana žanrska opredelitev (žanri so naštetih v prvem odstavku uvodnega razdelka tega članka). Del gradiva sodi v več žanrskih skupin hkrati, ti imajo oznako »večžanrski primeri«, nekatere enote pa v nobeno, zato imajo oznako »nerazporejeno«; pri slednjih bo potreben premislek, kam jih uvrstiti. Predviden je tudi stolpec za navedbo vira in druge morebitne opombe (npr. če je enota znana samo v posamezni vasi, specifičen kontekst ipd.). Tako urejeno gradivo bo zaradi prevoda v knjižni jezik omogočalo tako iskanje po leksemskih sestavinah (koren besede ali različni skloni izbrane besede) kot po žanrih, deloma pa tudi po drugih metapodatkovnih oznakah.

Saša Poklač ugotavlja (2004: 138), da največjo težavo pri raziskovanju narečnih frazeoloških enot predstavlja pomankanje večjega nabora gradiva. Namen je, da bi iz bovškega gradiva nastala znanstvenokritična slovarska objava, vendar brez diplomatskega prepisa, ki predvideva dosleden prepis z znaki, kot so bili uporabljeni v izvorniku: del gradiva so, kot že rečeno, zapisali laiki brez premišljenega koncepta zapisa in poznavanja dialektoloških zakonitosti,⁶ zato diplomatski prepis ne bi prinašal dodane vrednosti in relevantnih informacij. V slovarskem naboru frazeološkega gradiva in folklornih obrazcev, bo zato izhodišče kritični zapis z ustreznimi (usklajenimi) znaki za isti fonem. Gradivo bo razporejeno v žanrske skupine, vsaki enoti bo dodan knjižni prevod in po potrebi tudi pomensko pojasnilo ob primerih, ki nimajo ustreznice v knjižnem jeziku, ter dostopni metapodatki in različni (jezikoslovni, folkloristični,

⁵ Več o najpogostejših napakah ljubiteljskega zapisa gl. v opombi 6.

⁶ Naj izpostavim eno ali dve najpogostejši veliki napaki laičnega zapisa »na posluh«: domačini v diftongih *ie uo* pogosto »slišijo« glas *j* in ga s tem fonemom tudi zapisujejo npr. *cjsta, mljko* namesto *ciesta, mlieko* za besedo *cesta* in *mleko*, če uporabimo poenostavljen narečni zapis z običajnimi znaki slovenske abecede. V narečnem zapisu bi zapis vključeval še znak za zelo ozki izgovor (piko pod drugim delom diftonga). Iz tega sledi, da na mestu z dolgim (dvojnimi) in vedno naglašanim vokalom stoji kar konzontant *j*, vokala pa sploh ni, pri čemer hkrati »izpade« še naglasno mesto. Dialektološki zapis bi vključeval tudi oznako za tonem (akut ali cirkumfleks), ki ni odvisen le od leksikalne, pač pa tudi od sintaktične intonacije. Drug zelo razširjen problem laičnih zapisov je velika raznolikost označevanja polglasnika, npr. z znakom 3 (gl. opombo 4) ali drugimi znaki, ki jih slovenska abeceda ne vsebuje (npr. *x, y*), z apostrofom (') ali pa kar izpustom polglasnika, tudi če je le-ta artikularan; še posebej je »pester« zapis polglasnika pred fonemom *r*, pa naj je ta zlogotvoren ali ne (bovščina ima namreč veliko redukcij nenaglašanih vokalov, prav tako so v polglasnik reducirani tudi nekateri naglašeni vokali, ti pa imajo lahko dinamični ali tonemski naglas (več o značilnostih bovškega govora gl. v Ivančič Kutin 2007 in Kenda-Jež 2011: 143–155).

pojasnjevalni) komentarji. Fonetični zapis nameravam nadomestiti s kodo QR, ki bo vsebovala zvočni posnetek izgovora v narečju.

Pri nastajanju narečnega slovarja, pa naj bo to slovar narečnih leksemov, frazemov, paremij in drugih vrst narečnega gradiva, je izredno pomembno sodelovanje domačinov – nosilcev narečnega govora: v vlogi informatorjev (kot udeleženci pri spontanem govornem aktu, pripovedovalci folklornih in spominskih pripovedi ...), pri zbiranju in preverjanju gradiva. Nenazadnje so domačini (in naslednje generacije) tudi naslovniki, ki jim je mdr. objava tovrstnega gradiva namenjena (Ivančič Kutin 2009: 271).

2 Nekaj dosedanjih ugotovitev o bovškem frazeološkem gradivu in fokus tukajšnje obravnave

Nekaj gradiva je bilo že vključenega v ponazarjalne razdelke posamičnih gesel v *Slovarju bovškega govora* (Ivančič Kutin 2007). Usmerjene analize so bile na bovškem gradivu doslej opravljene le na nekaterih kulturnospecifičnih enotah, ki se eksplicitno navezujejo na segmente lokalnosti in jih je mogoče razumeti le ob poznavanju kulturnega konteksta; to so modelne podobe, ki neposredno kažejo odnos do življenjskih okoliščin, v katerih se nahajajo rojeni govorniki (Vrbinc in Vrbinc 2019: 70). V bovškem gradivu so bili na ta način obravnavani primeri, ki vsebujejo sestavino iz lokalnih antroponimov, kot so imena in kolektivni vzdevki za prebivalce določene vasi ali individualni antroponimi posameznih prebivalcev: ime / priimek / hišno ime / vzdevek, slednji je pogosto kar hišno ime (Ivančič Kutin 2022); gradivo je bilo ob tem razvrščeno v naslednje žanrske sklope: 1. pregovori; 2. primerjalni frazemi, znotraj teh pa še a) primerjave s prebivalci cele vasi, b) primerjave s konkretno osebo oz. njeno lastnostjo, ter c) primerjava s konkretno, v prostoru (nekoč) obstoječo živaljo; 3. zmerljivke, šaljive in druge preproste rime z lokalnimi antroponimi (n. d. 339–345). Za kulturološko interpretacijo so še posebej uporabni primerjalni frazemi, ki izražajo fizične ali psihične lastnosti človeka v primerjavi z živalmi, rastlinami, naravnimi pojavi itd. (Kržišnik 2008: 39); bovško gradivo preko enot z vsebnostjo lokalnih antroponimov to potrjuje. Frazemi, ki so vezani na točno določen lokalni kulturni kontekst, so izrazito lokalni in nimajo ustreznice ne v knjižni frazeologiji niti v drugih narečjih (Poklač 2004: 139).

Tudi v tukajšnji obravnavi bo prikazano gradivo, ki je v nekaterih pogledih kulturnospecifično. *Bog, hudič oz. zlodej, glava, rit, nos, drek* so med najpogostejšimi (samostalniškimi) leksemi, ki sestavljajo frazeološke enote v bovškem gradivu. Za ponazoritev žanrske raznolikosti ter glavnih jezikovno-strukturnih in pomenskih značilnosti so izbrani primeri s sestavino *rit*; bogat nabor tovrstnih frazemov in paremij imajo tudi druga slovenska narečja (gl. Smole 2007). V konotativnem pogledu so primeri s to sestavino nevtralnejši od ustreznice v knjižnem jeziku, kjer so v

(frazoloških) slovarjih označeni s kvalifikatorjem *vulgarno* (Smole n. d.). Poleg tega je namen tukajšnjega prispevka objaviti relativno obsežen in po mojem mnenju precej raznolik fond konkretnega gradiva s to leksemsko sestavino, in sicer v obliki, zapisani iz diskurza samega. Empirični podatki o aktualni rabi frazeoloških enot namreč omogočajo analizo in interpretacijo s sociolingvističnih vidikov npr. jezikovne zvrstnosti in analize glede na izbrane zunajjezikovne parametre (Jesenšek 2013: 19), prav tako pa omogočajo preučevanje semantike frazemov, kot jih npr. predvideva etnolingvistični ali lingvokulturološki pristop (Kržišnik 2004: 67).

3 Frazeološke enote in paremije kot folklorni trenutek in/ali kot del teksture folklornega dogodka

Osnovne lastnosti frazeoloških enot so večbesednost, stalnost, nezamenljivost sestavin oz. neizpeljivost pomena iz sestavin in ekspresivnost (Kržišnik 1994: 92). Frazemi vseh vrst, tudi metaforične besedne zveze in primerjalni frazemi so nestavčne strukture, paremije pa so pregovori in njim podobne frazeološke (relativno) stalne stavčne jezikovne strukture, ki jih Paremijakov opredelil kot minimalne teksteme (Permjakov, gl. v Meterc 2017: 26–27). V slovenščini paremije predstavljajo raznoliko in številčno množico izrazov, ki v strnjeni in lahko zapomnljivi obliki posredujejo široko paleto sporočil, naukov in učinkovitih opisov zelo tipičnih pa tudi zelo specifičnih življenjskih situacij (Meterc 2017: 9). S strukturo, izvorom, funkcijo teh izrazov se ukvarja paremiologija, to je jezikoslovna in folkloristična veda, Saša Babič jo opredeljuje kot vejo slovstvene folkloristike (Babič 2011: 27), *Sprotni slovar slovenskega knjižnega jezika* (Krvina 2014–) paremiologijo opiše kot jezikoslovno in folkloristično disciplino, ki proučuje pregovore in sorodne žanre, *Enciklopedijski riječnik lingvističkih naziva* pa kot vedo, ki se ukvarja z razlaganjem pregovorov glede na to, kako izražajo duha (Rikard 1969). Matej Meterc definicije združi v ugotovitev, da pregovori, reki in njim sorodni izrazi spadajo v okvir paremiologije, ki z jezikoslovnega vidika sodi v frazeologijo, z vidika žanrskih značilnosti teh frazemov pa v slovstveno folkloristiko (Meterc 2017: 9).

K folklorističnemu in jezikoslovnemu (frazologija, paremiologija) vidiku je – z ozirom, da so frazemi in paremije ndr. tudi stilne strukture, saj je njihov pomen prenesen, metaforičen – treba upoštevati tudi literarno-teoretični vidik. Tovrstno gradivo je torej lahko predmet in presečišče interdisciplinarnih raziskav. Ne glede na glavno perspektivo discipline, ki ga obravnava, pa je gradivo vselej močno vpeto v kontekst, ki obsega ne le sobesedilo diskurza / pripovedi, pač pa celoten sistem (kulturnega, jezikovnega ...) védenja / izkušenj tistega, ki frazem / paremijo (ali folklorni obrazec) izreka, in onega, na katerega je ta izraz naslovljen. Še posebej, če jih opazujemo kot sestavni del vsakdanjega (neumetnostnega) diskurza, te strukture niso zgolj zamenjava

za nek nevtralen izraz, pač obe sestavini (figuralna in nevtralna) stopata v interakcijo in s svojimi pomenskimi prvini učinkujeta druga na drugo; funkcija frazema / paremije (kot metafore) torej ni zgolj dekorativna, pač pa strnjen sistem asociiranih občih mest, stereotipnih, klišejskih predstav (Juvan 1992).

Slovstvena folklorja je besedna umetnost *govorjenega jezika* oz. narečij (Stanonik 2001: 143–146); slovstvena folkloristika folklorna besedila obravnava kot preplet treh neločljivo povezanih in prepletenih ravnin: *teksta* (kar lahko grafično zapišemo), *konteksta* (okolščine, v katerih je besedilo posredovano) in *teksture* (načina, kako je folklorno besedilo oblikovano). Tekstura je lahko: *besedna*, ta se nanaša na oblikovanje z izborom besed, besednega reda, premege govora, stilnih figur, diskurzivnih signalov; ter *nebesedna* (*glasovna*, *neglasovna*), ki folklorno besedilo oblikuje z glasom (glasnost, način izgovora, intonacija, imitacija ...) in govoricu telesa (gibi, mimiko) (podrobneje gl. Ivančič Kutin 2011: 39–49). Vse tri ravnine se dopolnjujejo, podpirajo in vplivajo druga na drugo ter skupaj sestavljajo t. i. folklorni dogodek (Dundes 1980). Teorija o folklornem dogodku temelji na dejstvu, da se slovstvena folklorja realizira z naravnim (govornim) tipom komunikacije, v živi interakciji med pripovedovalcem / govorcem / pevcem ter naslovnikom (sprejemalcem) (gl. Stanonik 2001: 141–142). Tudi če gre za dokaj stalno besedilo, sta tekstura in kontekst lahko bistvena za pomen/razumevanje besedila (Dundes 1965).

Na ravni vsakdanjega govornega posredovanja, npr. diskurza, ki ne sodi v umetnostno zvrst govornega besedila, zaseda izreka frazema/paremije/kratkega folklorne obrazca droben a učinkovit estetski/dramski/sporočilni insert. To so besedilne enote z velikim semantičnim potencialom (Grotjahn in Grzybek 2000). Upoštevaajoč kontekst in teksturo je izreka take zveze v slovstveni folkloristiki definirana kot samostojni *folklorni trenutek*, lahko pa je sestavina drugih zvrsti folklornih besedil – folklorne pesmi in folklorne proze (Babič 2015: 14–16; 2019: 22). Ko je zveza del drugih folklornih besedil, soustvarja pripovedni slog, zato je sestavni del besedne teksture. Te oblike so še posebej v pripovednih besedilih zelo pogoste, saj se lahko nizajo druga za drugo. Govorec z njimi poskuša svoje sporočilo (govor/pripoved/trditev ...) podati tako, da čim bolj pritegne/animira/vzbudi zanimanje in pozornost naslovnika.⁷

⁷ Veliko frekvenco teh izrazov nazorno prikaže kratek odlomek ene izmed bovških variant pravljice *Pu sbiata, u Čaljouca mīsto* [Po svetu, v Celovcu mesto] (Bremenski godci, ATU 130), paremije podčrtane: /.../ Tukiš guonjo usak sojo štimo na use grlo, (da) se čajo kar pertrgato. So nardilo takbiš romouš, da so se rauborjo ustrašle, pstil fčierjo an lätiel, kajk so moglo gor pruoť boško. /.../ [/.../ Tako gonijo vsak svoj glas na vse grlo, da se hočejo kar pretrgati. So naredili tak hrup, da so se jih ravbarji ustrašili in sl lezeli, kolikor so mogli navzgor proti gozdu /.../] (RV2).

4 Etimologija in značilnosti leksema *rit* v slovenskem jeziku in v bovškem narečnem govoru

Žanrsko raznolikost ter glavne jezikovno-strukturne in pomenske značilnosti frazemov in paremij bom poskušala prikazati na gradivu z leksemsko sestavino *rit*, saj je bilo v bovškem narečnem govoru evidentiranih več kot 100 takih enot. Etimološka razlaga besede *rit* (Snoj 2003: 623) je naslednja:

rit ríti ž lat. 'podex, anus' (16. stol.), rítka, ríten, rítnica, rítenski. Enako je cslovan. rítъ 'rit', hrv. rít v enakem pomenu, strus. rítъ z domnevnim pomenom 'kopito' ali 'bedro', pom. řac 'rit', češ. řít v enakem pomenu. Pslovan. *rítъ je dalje sorodno z lit. rietas 'stegno, krača, stegnenica', let. riéta 'noga' in arm. eri, rod. ervoy 'pleče pri živalih', vse iz ide. *rēiti-, *rēito- 'pleče' (M. S. pri Be III, 183).

Enote s sestavino *rit*, *zadnjica*, *zadnja plat* je v luči kvalifikatorjev analizirala dialektologinja Vera Smole. Na podlagi gradiva, tj. 49 primerov iz zahodnodolenjskega govora Šentruperta, ki ga je primerjala s knjižnimi ustreznici, je ugotovila razhajanja:

»Do bistvenega razhoda med knjižnim jezikom in neknjižnimi različki je prišlo v času, ko je prvotno nevtralni leksem *rit* v knjižnem jeziku postal vulgaren. Nadomestni nevtralni leksem *zadnjica* že po svoji fonetično-morfološki zgradbi v večini frazemov leksema *rit* ni mogel zamenjati, posledica tega je bila, da se je raba nekaterih frazemov s to sestavino omejila, večina »preživelih« pa je dobila predznak *vulgarno* ali *nizko*. (Smole 2007: 299)

To ugotovitev podpira tudi bovško gradivo. Kot je razvidno iz etimološke razlage, ima beseda *rit* v slovenščini dva pomena 1) *zadnjica*, in 2) debelejši del nečesa, npr. snopa, jabolka, drevesa, kar se (lahko) ujema tudi s pomenom najdebelejši del noge (Smole 2007: 301). Beseda *rit* je bila prav tako kot beseda *zadnjica* v slovenščini prvič zapisana že v 16. stoletju v Dalmatinovem Registru (1584). Vitalnost besede se kaže v tvorbi njene obsežne besedne družine (Smole n. d. 301). V Slovarju slovenskega knjižnega jezika je beseda v obeh pomenih označena s kvalifikatorjem *vulgarno*:

rit ríti ž (ĩ ĩ)

1. vulg. *zadnjica*: imel je strgane hlače, da se mu je videla rit; voda mu je segala do riti; položiti si blazino pod rit; udariti po riti; vedno sta bila (skupaj) kot rit in srajca / po riti se je spustil po stopnicah
2. vulg. neznačajan, bojazljiv človek: ta rit je vse izdal; kaj se dogovarjaš z njim, ki je čisto navadna rit / kot psovka: molči, ti rit; ti rit domišljava; prekleta rit. (SSKJ)

Beseda je v bovškem narečnem govoru nevtralna beseda, ki imenuje ta del telesa; sopomenke, kot so npr. *zadnja plat*, *zadnjica*, *ta zadnja* v bovškem narečnem govoru sploh ne obstajajo; ekspresivno (ljubkovalno) konotacijo ima beseda *rita* in manjšalna oblika *ritca*: *Nə jej čukulade, rita débela! Ma ritce lih zə no piest.*

rit || ə:rt ri:tə; ri:tə ž – 1. rit: iti: ukò:u ri:tə tu wà:ržet > iti po daljši poti / miéte pùhno ə:rt usó:ya; 2. ekspr. zadnja, včasih tudi spodnja stran česa: wən_s tó:ya ló:nca nə muó:řəš nəli:t tu flà:ško, kər té:če pud ə:rt > pri prelivanju tekočina teče po zunanji strani posode. (Ivančič Kutin 2007: 85–86)

Kljub nevtralnosti poimenovanja je ta del telesa pripraven za tvorjenje najrazličnejših frazeoloških struktur zaradi njegove lege, funkcije, oblike, podobnosti V kontekstu frazeološke enote pa skupaj s celotnim sobesedilom prestopi iz nevtralnega v metaforični jezik!

Rit kot nevtralen leksem ali kot del frazema ali paremije se pojavlja v različnih žanrih slovstvene folklore, poleg folklornih obrazcev tudi v folklorni pesmi⁸ in v pripovedni folklori, npr. v zaključnih formulah pravljic.

Precejšen delež enot, ki jih pozna knjižni jezik, pozna tudi narečje (seveda z narečno fonetiko in leksiko), kar pokaže primerjava z gradivom, ki je vključena v *Kebrov Slovar slovenskih frazemov* (Keber 2003; Meterc 2020–); primeri (v kategoriji *frazemi* in *sestavina izrazov*), ki jih bovški govor ne pozna, so podčrtani:

Frazemi in pregovori s sestavino *rit*:

bíti kómu za rítjo, bítí kot rit in srájca, bítí tího kot rit, bítí v rítí, bfcniti kóga v rit, dáti [jáh] kómu po rítí, dáti kómu brco [v rit], dáti kómu [êno] po rítí, dobíti brco [v rit], dobiti jih po rítí, dobíti žúlje na rítí, hodíti okóli rítí v váržet, hodíti okóli rítí v žêp, iméti mévlje v rítí, iméti sršene v rítí, íti [hudiču] v rit, íti okóli rítí v žêp, izprašíti kómu rit, kúzla skáče v rit kómu, lahkó me kúšnete na rit, lésti kómu v rit, lézenje v rit, naložíti [jáh] kómu po rítí, okóli rítí v váržet, okóli rítí v žêp, píši me v rit, píšmevrit, po rítí bi biló tréba kóga, pójdi [hudiču, vrágu] v rit, práskati se po rítí, prítí okóli rítí v žêp, sprána rit, státi kómu za rítjo, súniti kóga v rit, svetíti se kot púmpeževa rit, temà kákor v rítí, vsè k rítí prinèsti kómu.

Baba pijana, rit prodana, Glava sivi, rit nori, Na rit, Od jeze rit raste, S polno ritjo je lahko sрати, Vsaka rit pride na svoj sekret, Za vsako rit raste palica, Za vsako rit se šiba najde.

Poleg nepodčrtanih bovški govor pozna še številne druge primere.

5 Kratka analiza gradiva

Bovške frazeološke enote s sestavino *rit*, so razvrščene v tri tabele glede na žanrsko skupino (glej prilogo v zadnjem delu članka): 1. pregovori, 2. frazemi; 3. primerjalni frazemi.

⁸ Npr. v folklorni pesmi Terzinka zgodaj vstala /.../ vsó črno rit je imela /.../ (Blažević idr. 2022: 50, 51), zagovorih: *Za te hude mozou, za ščirauc: Ti te hudi mozou, si un alboj ona? Teci ta na to dulinico, ki je den muž, ki an ma uaginj, an serje, an nima riti* (Palleti 1984: 140). Zaključna formula pravljice: *Meni so dali iz fngreta jesti, z rešeta piti pa z loparjem po riti, imam še zdaj graben na riti. Pol sem pa šel po svetu, sem pa pravil, kako se mi je godilo. Kdor govori, je norec, kdor posluša, je pa še bolj. Zdaj je pa konec* (Križnik 2019: 194; gl. tudi Kropelj Telban 2019: 20). O formulativnem zaključevanju slovenskih pravljic gl. več v Stanonik 2023: 306–319.

T. i. somatični frazemi in paremije, ki vključujejo poimenovanje dela telesa, torej tudi *rit*, so v večini jezikov in njihovih narečjih zelo številni, saj so poimenovanja za telo del jedrnega besedja (Smole 2023: 139). Pomeni teh besednih zvez so motivirani s strani kognitivnih mehanizmov, kot so splošno sprejeto znanje ter konceptualne metafore in metonimije (Polanc 2020). Čeprav je leksem *rit* v rabi tudi za poimenovanje dela živalskega telesa, pa se v bovških enotah skoraj izključno pojavlja le kot del človeškega telesa, izjemi sta: *dielo mo grie ud rukie koj smola p̄aso ta ud rit̄* 'težko'; *Ma šuobe ku kurja ert* 'rdeče ustnice (od šminke)'. V bovškem govoru je bilo evidentiranih 107 enot, od tega 18 pregovorov ter 18 primerjalnih frazemov, kar kaže na izjemno bogato semantiko leksema *rit* v tem narečnem okolju.

Pregovori v prispodobi na kratko povzemajo neko splošno izkušnjo, posebna podvrsta so vremenska predvidevanja, kmetijska in zdravstvena napotila in so relativno stalne stavčne zveze (Babič 2015: 61–69). V bovškem gradivu gre največkrat za dvo- ali štiristavčne povedi, nekateri pregovori pa so zgolj enostavčni, izjemoma z elipso glagola. *Tejst_k zguodo ustaja, mo kruha ustaja, tejst_k dugo laži, se mo slama rit drži; Kar grlo p̄asa, rit n̄ fali.; Tie sit̄ se lohko z ritjo kriaga.; Jaga – r̄at naga.* Njihova pragmatična raba je v tesni povezavi z besedilnim in/ali situacijskim kontekstom, ki ga pregovor povzema, ocenjuje/parafrazira, in sicer z zelo raznolikimi funkcijami, npr. kot šaljiv/nevtralen komentar, svarilo, poduk, opozorilo, napotilo, predvidevanje ... Iz načina, kako je pregovor izrečen, lahko razberemo humor, očitek, svarilo, neodobravanje, napotek ... tj. govorčev odnos do situacije, stanja, dejanja, lastnosti, ki je predmet pregovora. Enako velja tudi za vse oblike frazemov. Pregovor je že sam po sebi metafora, torej z visoko stilno vrednostjo (več gl. Babič n. d.). Od stilnih sredstev so v bovških pregovorih s sestavino *rit* največkrat zastopani rime, ritem, včasih tudi personifikacija žariščnega leksema, vse skupaj npr. v enoti *Dokl̄ar je d̄amača muoka, ert dobro puoka; k_pride n̄ kpilo, glieda r̄at milo.*

Če so pregovori stalne stavčne strukture, ki so v besedilo umeščene vedno v (relativno!) enaki slovnični obliki, pa so frazemi vseh vrst (tudi primerjalni), povsem prilagojeni besednemu okolju (Meterc 2021: 59). Zato se lahko pojavijo v najrazličnejših (osebnih) glagolskih oblikah, z različnimi slovničnimi števili in v različnih slovničnih časih, naklonih in načinih. Zapis enote v slovnični obliki, v kateri je bila enota uporabljena v komunikaciji, ohranja pomemben del konteksta, ki ga iz prepisa v slovarski obliki ne bi bilo mogoče razbrati. V govorjenem diskurzu bovškega narečja le redko zasledimo rabo frazemov in paremij v nevtralni (nedoločni) obliki. Izjema so tisti primeri, ki se pojavljajo v kombinaciji z modalnim glagolom: *Tu muore iti prej ko slej zluodiurt* 'ni druge poti, kot da propade'. Iz osebne glagolske oblike lahko razberemo, na koga je zveza usmerjena: na (prvo- ali drugoosebnega) udeleženca diskurza ali na tretjo osebo ali tretje osebe, ki niso prisotne. Npr. pri frazeoloških enotah, ki hkrati vsebujejo sestavljene *iti, hudič* in *rit* lahko opazujemo različne slovnične strukture in pomenske motivacije: *Je šlo zluodiv̄r̄at* kot metaforični opis v pomenu 'nekaj je šlo

narobe, se je razblinilo, propadlo'; *Biež zluodivṛət!* kot psovka, kletvica; *Eh, zluodivṛət!* kot vzklík v jezi; *Biāš z_buogəm zluodivṛt* kot pozdrav in kletvica hkrati v pomenu odslovitve, zavrnitve.

Primerjalni frazemi⁹ izražajo podobnost med podstavo in podobo, ki sta povezani z veznikom *kot, kakor, ko* (bovško *ku, kuj, kajk*) (Dolinar idr. 1987: 114.), v bovškem gradivu se pojavlja tudi primerjalni veznik (*tako*) *da*, npr. *Gueri, dā niāma ne rit an ne glaue*. Na bovškem dokumentirani primerjalni frazemi imajo različne pomenske motivacije. Po opredelitvi Mihaele Koletnik (2020) izražajo: 1) osebnostne lastnosti človeka, npr. *Uket, dā b drek ta spuod rit pujādu*; 2) telesne lastnosti, značilnosti, videz, hibe, npr. *Ma rət zā /kajk ən kəznienk* 3) čustva, občutke, občutenja *Səm se ustrašu, dā m_je kar gače tu rət putāgnilo*; 4) vedenjske vzorce *Je ko dā_b mov sršiene tu ritā /gačəh*; 5) dejanja *Gueri, dā niāma ne rit an ne glaue*; 6) stanja, lastnosti *Zmuškəno ku wən_s kraulje ritā*.

Primerjalni frazemi kažejo odnos do življenjskih okoliščin, v katerih se nahajajo rojeni govorniki (Vrbinc in Vrbinc 2019: 70). V njih gre za pripisovanje lastnosti ene stvari kaki drugi stvari, pri čemer so nekateri etaloni skupni mnogim (nacionalnim itd.) skupnostim, obstajajo pa tudi taki, v katerih se kažejo posebnosti določenega naroda, jezika in/ali kulture (Kržišnik 2008: 39). V bovškem gradivu je mogoče najti tudi primere, ki niso evidentirani v knjižnem jeziku in drugih slovenskih narečjih.

6 Ugotovitve

Med najpogostejšimi frazemi in paremijami v bovškem gradivu so tiste z besedno sestavino *bog, hudič oz. zlodej, rit, nos in drek, glava*. Za prikaz žanrske raznolikosti ter glavnih jezikovno-strukturnih in pomenskih značilnosti so izbrani primeri s sestavino *rit*; leksem *rit* je v knjižnem jeziku z razvojem skozi čas pridobile ekspresivno konotacijo (v sodobnih slovarjih s kvalifikatorjem vulgarno), v bovškem narečnem govoru pa je to osnovna, stilno nezaznamovana beseda za poimenovanje tega dela telesa. Ta je glede na lego, funkcijo, obliko, podobnost, simboliko in še kaj zelo pripraven za tvorjenje najrazličnejših frazeoloških struktur; ko leksem skupaj s sobesedilom tvori frazem ali paremijo, prestopi iz nevtralnega v metaforični jezik. Precejšen delež enot, ki jih pozna knjižni jezik, pozna tudi bovški narečni govor (seveda z narečno fonetiko in leksiko). Somatične frazeološke enote so v večini jezikov in njihovih narečjih zelo številni, saj so tudi poimenovanja za telo del jedrnega besedja (Smole 2023: 139). Čeprav je leksem *rit* v rabi tudi kot del živalskega telesa, pa se v bovškem gradivu skoraj izključno pojavlja le kot del človeškega telesa.

⁹ Literarna teorija, ki primerjalne frazeme imenuje primere/komparacije pa tudi *podobje*, razlikuje več vrste (literarne) figure (gl. Pavlič 2010).

Velik nabor 107. frazeoloških enot s sestavino *rit* kaže na izjemno bogato semantiko tega leksema v bovškem narečnem govoru (gl. tabele v prilogi). Tako kot v splošnem tudi za bovške pregovore velja, da so stalne stavčne strukture, ki so v besedilo umeščene v (relativno) enaki slovnični obliki, medtem ko se frazemi vseh vrst (tudi primerjalni), povsem prilagajajo besednemu okolju (Meterc 2021: 59). Zato se lahko pojavijo v najrazličnejših (osebnih) glagolskih oblikah, z različnimi slovničnimi števili in v različnih slovničnih časih, naklonih in načinih. Besedilo/situacija, ki ga paremija povzema, ocenjuje oz. parafrazira, lahko razberemo raznolike pragmatične funkcije: šaljiv/nevtralen komentar, svarilo, poduk, opozorilo, napotilo, predvidevanje, očitek, svarilo, neodobravanje ... tj. govorčev odnos do povedanega/situacije/dejanja; k razumevanju pomena pripomore tudi načina izreke (npr. ton glasu, ponazoritev s kretnjo). Zapis enote v slovnični obliki, v kateri je bila enota uporabljena v komunikaciji, ohranja del konteksta, ki ga iz prepisa v slovarski obliki ne bi bilo mogoče razbrati, zato je tak prikaz gradiva lahko koristna podlaga za nadaljnje jezikoslovne, folkloristične in druge raziskave, ki se ukvarjajo z besedno (narečno) kulturo.

Zahvala

Članek je nastal v okviru raziskovalnega programa *Etnološke, antropološke in folkloristične raziskave vsakdanjika* (P6-0088) ter projekta *Tradicionalne paremiološke enote v dialogu s sodobno rabo* (J6-2579).

Za zavzeto zapisovanje gradiva se zahvaljujem svoji ožji družini iz Bovca – sestri, očetu in pokojni mami; prav tako pa tudi teti in vsem drugim domačinom, ki so prispevali gradivo.

Viri in literatura

- Babič, Saša, 2011: Paremiologija na križišču. V: Simona Kranjc (ur.), *Meddisciplinarnost v slovenistiki*. Ljubljana: Znanstvena založba Filozofske fakultete Univerze v Ljubljani (Obdobja, 30), 27–31.
- Babič, Saša, 2015: *Beseda ni konj. Estetska struktura slovenskih folklornih obrazcev*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612547660>.
- Babič, Saša, 2019: Obravnava folklornih obrazcev v osnovnošolskih berilih in revijah. *Jezik in slovstvo* 64/2, 19–30.
- Blažević, Marko; Ferlan, Darko; Haler, Miha; Šivic, Urša (ur.): *Kapelski pubje še tak popevamo*. Kapele: Kulturno društvo Kapele – ljudski pevci kapelski pubje.
- Dolinar, Darko; Koren, Evald; Kos, Janko; Kralj, Lado; Orešnik, Janez; Žnideršič, Martin, 1987: *Literatura*. V: Kos, Janko; Dolinar, Ksenija (ur.), *Leksikoni Cankarjeve založbe*. Ljubljana: Cankarjeva založba.
- Dundes, Alan, 1965: *The Study of Folklore*. Englewood Cliffs, NJ: Prentice-Hall.
- Dundes, Alan, 1980: Texture, Text, and Context. *Interpreting Folklore*. Indiana: University Press, 20–32.
- Grotjahn, Rüdiger; Grzybek, Peter, 2000: Methodological remarks on statistical analyses in empirical paremiology. *Proverbium: Yearbook of International Proverb Scholarship* 17, 121–132.
- Ivančič Kutin, Barbara, 2007: *Slovar bovškega govora*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610503194>.

- Ivančič Kutin, Barbara, 2009: Sodelovanje domačinov pri pripravi narečnega slovarja (na primeru Slovarja bovškega govora). V: Smole, Vera (ur.), *Slovenska narečja med sistemom in rabo*. Ljubljana: Znanstvena založba Filozofske fakultete Univerze v Ljubljani (Obdobja, 26), 269–276.
- Ivančič Kutin, Barbara, 2011: *Živa pripoved v zapisu. Kontekst, tekstura in prekodiranje pripovedi Tine Kravanja iz Bavšice*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612543518>.
- Ivančič Kutin, Barbara, 2022: Paremiološko gradivo iz Zgornjega Posočja v dveh rokopisnih zbirkah. Elementi lokalnosti s poudarkom na kulturnospecifičnih folklornih obrazcih z lokalnimi antroponimi. *Slavistična revija* 70/3, 333–348.
- Jesenšek, Vida, 2013: Frazeologija nemškega jezika: slovensko-nemški kontrastivni in medkulturni vidiki. Predstavitev projekta. V: Jesenšek, Vida; Babič, Saša (ur.), *Več glav več ve. Frazeologija in paremiologija v slovarju in vsakdanji rabi*. Maribor, Ljubljana: Oddelek za germanistiko, Filozofska fakulteta univerze v Mariboru in Inštitut za slovensko narodopisje ZRC SAZU. 10–25.
- Juvan, Marko, 1992: Pesmi v stereofoniji: skica za medbesedilne figure. *Slavistična revija* 40, 43–59.
- Keber, Janez, 2015: *Slovar slovenskih frazemov*. Na spletu: <https://fran.si/192> (dostop 12. 1. 2024).
- Kenda-Jež, Karmen, 2011: Opis narečnega govora. V: Ivančič Kutin, Barbara, *Živa pripoved v zapisu. Kontekst, tekstura in prekodiranje pripovedi Tine Kravanja iz Bavšice*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612543518>.
- Koletnik, Mihaela, Nikolovski, Gojko, 2020: Primerjalni frazemi z zoomimnimi sestavinami v prekmurskem narečju in makedonščini. *Slavia Centralis* 13/2, 83–99. DOI: <https://doi.org/10.18690/scn.13.2.83%E2%80%9399.2020>.
- Križnik, Gašper, 2019: *Sveti Coprijan. Pravljice, povedke in drugo pripovedno izročilo iz Motnika in okolice*, ur. Podbrežnik Vukmir, Breda; Kropelj Telban, Monika. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610504498>.
- Kropelj Telban, Monika, 2019: Križnik – Folklorist – samouk. V: Podbrežnik Vukmir, Breda; Kropelj Telban, Monika (ur.), *Sveti Coprijan. Pravljice, povedke in drugo pripovedno izročilo iz Motnika in okolice*. Ljubljana: Založba ZRC, ZRC SAZU, 13–21. DOI: <https://doi.org/10.3986/9789610504498>
- Krvina, Domen (ur.), 2014–: *Sprotni slovar slovenskega jezika*. Na spletu: <https://fran.si/132> (dostop 4. 1. 2024).
- Kržišnik, Erika 1994: Frazeologija kot izražanje v »podobah«. V: Križaj-Ortar, Martina; Bešter, Marja; Kržišnik, Erika (ur.), *Pouk slovenščine malo drugače*. Trzin: Different, 91–140.
- Kržišnik, Erika, 2008: Viri za kulturološko interpretacijo frazeoloških enot. *Jezik in slovstvo* 53/1, 33–47.
- Meterc, Matej, 2017: *Paremiološki optimum: Najbolj poznani in pogosti pregovori ter sorodne paremije v slovenščini*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610504153>.
- Meterc, Matej, 2020–: *Slovar pregovorov in sorodnih paremioloških izrazov*. Na spletu: <https://fran.si/216> (dostop 11. 1. 2024).
- Meterc, Matej, 2021: Paremiologija v bazi znanja: o pregovorih in sorodnih izrazih. V: Ulčnik, Natalija (ur.), *Slovenščina na dlani* 3. Maribor: Univerza v Mariboru, Univerzitetna založba, 57–62.
- Palleti, Silvana, 1984: Samoobrambne besede, navade in verovanja v Reziji. *Traditiones* 13, 139–143.

- Pavlič, Darja, 2010: Pesniške, mentalne in druge podobe. *Primerjalna književnost* 33/3, 235–249.
- Pleteršnik, Maks, 2014: *Slovensko-nemški slovar*. Na spletu: <https://fran.si/136> (dostop 5. 2. 2024).
- Podbersič, Renato, 2007: *Čičke prekvantice z Goca. Folklorne obrabice z Golca v Slovenski Čičariji*. Koper: Libris.
- Poklač, Saša, 2004: Nekaj teoretičnih izhodišč o narečni frazeologiji. *Jezikoslovni zapiski* 10/1, 137–145.
- Polanc, Tea, 2020: *Ne vlečem vas za nos: kontrastivna analiza idiomov, osnovanih na delih telesa, z vidika kognitivnega jezikoslovja*. Magistrsko delo. Na spletu: <https://repozitorij.uni-lj.si/IzpisGradiva.php?lang=slv&id=121459> (dostop 2. 2. 2024).
- Rikard, Simeon, 1969: *Enciklopedijski rječnik lingvističkih naziva na 8 jezika: hrvatsko-srpski, latinski, ruski, njemački, engleski, francuski, talijanski, španjolski*. Zagreb: Matica hrvatska.
- Smole, Vera, 2007: Vprašanje kvalifikatorjev v slovenskih (frazeoloških) slovarjih knjižnega jezika in narečij (na primeru frazemov s sestavino rit, zadnjica, zadnja plat). V: Kržišnik, Erika; Eismann, Wolfgang (ur.), *Frazeologija v jezikoslovju in drugih vedah: Europhras Slovenija 2005*. Ljubljana: Filozofska fakulteta Univerze v Ljubljani, Oddelek za slovenistiko, 299–312.
- Smole, Vera, 2023: Pomen leksema oko kot sestavine frazema v slovenskem vzhodnodolenjskem govoru Šentruperta. *Jezik in slovstvo* 68/2, 137–158.
- Snoj, Marko, 2016–: *Slovenski etimološki slovar* 3. <https://fran.si/193> (dostop 12. 1. 2024).
- SSKJ: *Slovar slovenskega knjižnega jezika*, druga, dopolnjena in deloma prenovljena izdaja. Na spletu: <https://fran.si/133> (dostop 12. 1. 2014).
- Stanonik, Marija, 2001: *Teoretični oris slovstvene folklore*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9616358316>.
- Stanonik Marija, 2023: *Poetika slovenske slovstvene folklore. Besedna umetnost v koreninah*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610507659>.
- Vrbinc, Alenka; Vrbinc, Marjeta, 2019: Prepletenost kulture in jezika na zgledih angleških in slovenskih izlastnoimenskih frazemov. *Slavistična revija* 67/1, 69–85.

Rokopisni viri (RV)

- RV1: Dolenc, Janez, Dijaški arhiv Janeza Dolenca. Kopije hrani Arhiv Inštituta za slovensko narodopisje ZRC SAZU.
- RV2: *Pu sbiətə, u Cəljoucə miasto*. Zapisala Dana Ivančič po pripovedi mame Terezije Komac (1924–1980). Osebni arhiv Barbare Ivančič Kutin.

Spletni viri (SV)

- SV1: Register slovenske nesnovne dediščine, Bovški govor: <http://www.nesovnadediscina.si/sl/register/bovski-govor> (dostop 12. 1. 2024)
- SV2: Register slovenske nesnovne dediščine, Bovška slovstvena folklor: <http://www.nesovnadediscina.si/sl/register/bovska-slovstvena-folklor> (dostop 12. 1. 2024)

Ustni viri:

Domačini z Bovškega.

Priloga

Nabor paremij in (primerjalnih) frazemov s sestavino *rit* v bovškem govoru

Bovške frazeološke enote s sestavino *rit* so razvrščene v tri tabele glede na žanrske skupine: 1. pregovori, 2. frazemi; 3. primerjalni frazemi.

Gradivska priloga k članku

	Bovško v diskurzivni obliki¹⁰	Dobesedni knjižni prevod	Pomen / frazemska sopomenka
1.	<i>Dokl̥ar je d̥amača muoka, ert dobro puoka; k pride na kpilo, gleda r̥at milo.</i>	Dokler je domača moka, rit dobro poka, ko pride na <i>kupilo</i> (= ko je treba moko kupiti), gleda rit milo.	Metaforičen komentar o času obilja in pomanjkanja (poosebitev).
2.	<i>Tejst k zguodo ustaja, mo kruha ustaja, tejs_t k_dugo laži, se mo slama rit drži.</i>	Kdor zgodaj vstaja, mu kruha ostaja, kdor dolgo leži, se mu slama riti drži.	<i>Brez dela ni jela.</i>
3.	<i>Kar grlo p̥asa, rit na fali.</i> Tudi varianta Č̥̃ te puj̃, te že posirje.	Kar grlo preide, riti ne zgreši. Če te poje, te že <i>poserje</i> (izloči z blatom).	Kar se zaužije, se tudi izloči. Tudi šaljiva »tolažba« otroku, ki se boji npr. teme.
4.	<i>Tejst k leti, na zdravi rit̃ sedi.</i>	Kdor teče (hitro hodi), na zdravi riti sedi.	(Hitro) gibanje je / odraža zdravje.
5.	<i>Tie sit̃ se lohno z ritjo kriaga.</i>	Siti se lahko z ritjo krega.	Kdor je premožen, si lahko privošči.
6.	<i>Klubasa ert n̥abasa, plienta jo ubrianta.</i>	Klobasa rit <i>nabasa</i> (napolni), polenta jo pa <i>obrenta</i> (odebeli).	Klobasa naredi človeka sitega (dá moč), poleta pa samo redi.
7.	<i>Č̃̃ stuop̃aš pl̃adr̃a ta ñ jez̃ak, pa z ritjo upliata.</i>	Če stopiš klepetulji na jezik, pa z ritjo opleta.	Nemogoče je prevzgojiti osebo; vedno je način, da se izrazi.
8.	<i>Buoh ma takob̃iš šibo tu ruoc̃oh, d̃a ñ usako ert paša.</i>	Bog ima tako šibo v rokah, da se na vsako rit prilega.	Vsakdo dobi ustrezno kazni; svarilo.
9.	<i>Z̃a usako r̃at rase paljca.</i>	Za vsako rit raste palica.	Vsakdo dobi ustrezno kazni, svarilo, poduk.
10.	<i>Ubrni se kuj č̃̃š, ert je zmierti tem zad.</i>	Obrni se kot hočeš, rit je vedno zadaj.	Zadeve so take, kakršne pač so, ne glede na interpretacijo / zorni kot.
11.	<i>Č̃̃ je ñ sbĩat Lieñart l̃apu, lohno še usaka lĩna ert steljo dubuo.</i>	Če je na svetega Lenarta lepo, še vsaka lena rit steljo dobi.	Vremenski pregovor z besedno igro.
12. ab	<i>Svieta Nieža kur̃m ert opre.</i> <i>Ña svieta Niežo se kur̃m ert opre.</i>	Sveta Neža kuram rit odpre. Na sveto Nežo se kuram rit odpre.	O godu sv. Neže (21. 1.) kokoši začno nesti jajc po zimskem premoru.
13. ab	<i>Jaga – ert naga.</i> <i>Jaga – ert naga, puša ña striel, ñjiwa u pl̃biev.</i>	Jaga – rit naga. Jaga – rit naga, puška na strel, njiva v plevel.	Strasten lovec je slab gospodar.
14.	<i>Soude an ert je trieba skrita.</i>	Denar in rit je treba skriti.	O tem se ne govori / tega se ne razkazuje.

¹⁰ Narečni zapis je v poenostavljeni strokovni obliki z običajnimi znaki slovenske abecede z dodanim polglasnikom. Izgovor dvoustičnega zaokroženega *u* iz *v* in *l* je zapisan z *v*, dvoustični nezaokroženi pred zadnjimi samoglasniki, največkrat *a*, pa z *w*. Spajanje besed je označeno z znakom *_*. Priporniški *ȳ* je zapisan z znakom *g*.

15.	<i>Lies pukuonca, prsie tu rilc an baba tu ritə.</i>	Les (ko stoji) pokonci, prašič v rilcu in baba (ženska) v riti.	Pojasnilo, kje ima kdo/kaj največjo moč.
16.	<i>Ret je rət an rət ustane.</i>	Rit je rit in rit ostane.	Hinavec, človek s slabim značajem se nikoli ne spremeni.
17.	<i>Sniwanja so puhna ət spanja.</i>	Sanjanja so polna rit spanja.	Utujen človek spi, ne pa sanja. Če sanjaš, je spanja dovolj.
18.*	<i>Toja welja, tuoj je sviet, zə trompeto maš pa rət!</i>	Tvoja volja, tvoj je svet, za troben-to imaš pa rit.	Naredi pa po svoje, če že ne želiš upoštevati mojega nasveta; s cinično pripombo, kako lahko svoje (trmoglavo) mnenje ali ravnanje razglasiš drugim. * Paremija, ki po strukturi spominja na pregovor.

Tabela 1: Pregovori.

	Bovško v diskurzivni obliki	Dobesedni knjižni prevod	pomen
1.	<i>Je mislu, də buga zə rət drži.</i>	Je mislil, da boga za rit drži.	biti aroganten, vzvišen
2.	<i>A b red dwa konja z no ritjo jahu?</i>	Ali bi rad dva konja z eno ritjo jahal?	početi dve nezdružljivi stvari
3.	<i>Čə mə gluwe uzaməš, rt tə šenjkəm.</i>	Če mi glavo vzameš, ti rit podarim.	vzemi vse, če že vzameš večji /boljši /glavni del
4.	<i>Ne sed nə mrzlem, dubuoš uka tu rt!</i>	Ne sedi na mrzlem, dobiš volka v rit!	<i>Dobiti /imeti volka v riti</i> je narečni medicinski izraz za skeleč občutek v danki po sedenju na mrzlem.
5.	<i>Se je bau ritə.</i>	Se je bal riti.	bati se kazni, posledic za svoje dejanje
6.	<i>A te muorm zmiəri miet zə ritjo?</i>	Te moram vedno imeti za ritjo?	imeti nazor / zasledovalca
7.	<i>Kar z rukamə nəriedi, pa z ritjo pudere.</i>	Kar z rokami naredi, pa z ritjo podre.	biti neroden
8.	<i>Je čist nə ritə.</i>	Je povsem na riti.	biti v finančno v slabem položaju
9.	<i>Pu rit an pu glav dəwəlj ma usəga.</i>	Po riti in po glavi dovolj ima vsega.	imeti na pretek, več kot dovolj
10.	<i>Ga je muoru cbnit tu rt. Səm se muogla prov tu rət cəbnits, də səm se učila.</i>	Moral ga je brcniti v rit. Sem se morala prav v rit brcniti, da sem se učila.	prisiliti /dati komu grobo spodbudo, da kaj stori
11.	<i>Usak ga cabne tu rt.</i>	Vsak ga brcne v rit.	imeti nekoga zadosti, naveličati se koga / obmiti hrbet komu
12.	<i>So ga cəbil tu rət.</i>	So ga brcnili v rit.	odsloviti /dati odpoved komu
13.	<i>Je nucu cəbalo tu rət.</i>	Potreboval je brco v rit.	potrebovati (grobo) spodbudo k začetku dela /aktivnosti
14.	<i>Je dubiv cəbalo tu rət.</i>	Dobil je brco v rit.	biti odpuščen
15.	<i>Ga puozna tri pədilje tu rt.</i>	Ga pozna ga tri pedi v notranjost riti.	natanko vedeti kaj /kako nekdo razmišlja
16.	<i>Jo je kar nə rət wrгло!</i>	Jo je kar na rit vrglo.	biti zelo presenečen /šokiran
17.	<i>Šlo je vse zluodivrt.</i>	Je šlo vse vragu /zlodeju v rit.	propasti
18.	<i>Ma vsəga puhno rt.</i>	Ima polno rit.	biti razvajen; tudi biti premožen
19.	<i>Ga je mov puhno rt. So ga žie malo mielə.</i>	Ga je imel polno rit. So ga že malo imeli.	biti pijan biti že nekoliko okajeni
20.	<i>Mo grie səmu zə sojo rt.</i>	Mu gre samo za lastno rit.	biti sebičen, egoist

21.	<i>A t kure nesəjo? Ja, majo zlato rt.</i>	Ti kokoši nesejo? Ja, imajo zlato rit.	biti zadovoljen s kokošmi (konkretualen frazem, vezan na ozek kontekst)
22.	<i>Smo žlahta ud rit dā lahta.</i>	Smo žlahta od riti do kopolca.	sorodniki brez tesnejših odnosov, ali sorodniki, ki so skregani med sabo
23.	<i>K sadəš riepo, se murəš pugladta ta pu ritə, dā buo dbila.</i>	Ko sadiš repo, se moraš pogladiti po (svoji) riti, da bo debela.	vraža
24.	<i>Sta šla ukov rit tu wəət.</i>	Sta šla okoli riti v žep.	iti po daljši poti; zakomplicirati pot do rešitve
25.	<i>Nima ritə (zə učenje).</i>	Nima riti (za učenje).	ne imeti potrpljenja za enolično sedeče delo, npr. učenje
26.	<i>Ga je gladu pu ritə.</i>	Ga je gladil po riti.	hvaliti koga, pritrjevati komu z namenom lastne koristi
27.	<i>Zmieri lieze tu rət tə psuokm.</i>	Vedno leze v rit ta visokim (pomembnejšem).	se prilizovati, biti narejeno prijazen z namenom koristi
28.	<i>Biež zluədi _v_ rət !</i>	Pojdi hudiči v rit.	kletvica, psovka, poslati koga v izničenje, propad
29.	<i>Je šlo zluodivrt use dielo.</i>	Je šlo zlodeju v rit vse (dosedanje) delo.	propasti
30.	<i>Tu je nĵəmo pišmevrt.</i>	To je njemu pišmevrit.	ne zanimati se za, ne ceniti koga/česa
31.	<i>A čəš dubit pu ritə?</i>	Hočeš dobiti po riti?	biti tepen, groznja
32.	<i>A buste žie uzignlə rət?</i>	Boste že dvignili rit?	1. oditi; 2. začeti delati
33.	<i>S təm pupirĵəjə lohno ədino rət ubriše.</i>	S tem papirjem lahko edino rit obriše.	dokaz brez vrednosti
34.	<i>Mat mo je use ta h rit prnesla.</i>	Mati mu je vse k riti prinesla.	imeti (preveliko) pomoč, biti nesamostojen
35.	<i>Se mo je rət triesla.</i>	Se mu je rit tresla.	bati se
36.	<i>Pazi, d _te kej nə zegrab zə rt!</i>	Pazi, da te kaj ne zagrabi za rit.	Šaljivo svarilo, če je koga strah, npr. iti sam nekam ponoči
37.	<i>Biež u rət s təm jamrənjəm.</i>	Pojdi v rit s tem tarnanjem.	zavrnitev
38.	<i>Pši _me u_ rət.</i>	Piši me v rit (v rit me piši).	zavrnitev, psovka
39.	<i>Naj tu rət utakne!</i>	Naj (si nekaj) v rit vtakne.	zavrnitev
40.	<i>Ma autio rət an tətə – (=avtoritetə).</i>	Ima avto, rit in teto.	imeti avtoriteto, besedna igra
41.	<i>Ojej, zdej smo pa tu ritə.</i>	Ojej, zdaj smo pa v riti.	biti v težkem položaju
42.	<i>On je pa ries (na) rət.</i>	On je res (ena) rit.	biti hinavec, zmerljivka
43.	<i>Mo je ubrnu rət.</i>	Obrnil mu je rit.	ignorirati koga, obrniti se stran, pokazati hrbet
44.	<i>Mo muorəš use wən z rit putəgnitə.</i>	Mu moraš vse iz riti potegniti.	težko izvedeti /dobiti kaj od nekoga
45.	<i>Še rit ni ubrnila.</i>	Še riti ni obrnila.	ne pokazati nikakršnega zanimanja
46.	<i>Mo grie təžku ta ud ritə.</i>	Mu gre težko od riti.	biti skop
47.	<i>Niso dielələ druzga ku ret praskələ.</i>	Niso delali drugega, kot rit praskali.	lenariti
48.	<i>Klade vse pud (sojo) rət.</i>	Daje vse pod (svojo) rit.	varovati, skrivati svojo lastnino
49.	<i>Pazi, tie lonc teče pud rət.</i>	Pazi, ta lonc teče pod rit.	curek pri izlivanju tekočine teče po celotni zunanji strani posode

50.	<i>Tem nā ših̄tā greje rāt.</i>	V službi greje rit.	nič narediti, lenariti
51.	<i>Huodjo rāt zmrzawēt gore nā Kanin.</i>	Hodijo rit zmrzovat na Kanin.	po nepotrebnem se izpostavljati mrazu
52.	<i>A ne_b šla malo ritā slpindrāt wān?</i>	Ne bi šla riti zmigat ven?	iti na sprehod /na zrak se razgibat
53.	<i>Ne bā škudilo, čā_tuo rāt malo splindrāla.</i>	Ne bi škodovalo, če bi to rit malo zmigala (v pomenu shujšala)	shujšati
54.	<i>Tu je na liena rāt.</i>	To je ena lena rit.	biti len človek, lena oseba
55.	<i>Ah kabi, bžita, punucita tie bone, u rit jih busta miela.</i>	Ah, seveda, pojdita in porabita te bone, v riti jih bosta imela.	izrabiti ponujeno priložnost
56.	<i>Buzeranto an lejtrām ne smieš nkol ritā kōzātā.</i>	Buzerantu in lestvi ne smeš nikoli riti kazati.	biti obrnjen s hrbtom
57.	<i>Nāč mo ni trieba, ma rāt n pospraviti nā gorkem.</i>	Nič mu ni treba, ima rit na gorkem (toplem).	biti preskrbljen/ biti v ugodnih razmerah
58.	<i>A nā muorāš ānkol zā sojo ritjo pusprabta?</i>	Ali ne moreš nikoli za svojo ritjo pospraviti?	pospraviti za seboj
59.	<i>Eh, naj u rāt naj grie.</i>	Eh, v rit naj gre.	vzklik; = <i>Naj se gre solit.</i>
60.	<i>Nā jej čukulade, rita dēbela!</i>	Ne jej čokolade, rita debela!	poimenovanje z (ljubkovalnim) vzdevkom
61.	<i>On je z_miesta, je fina rāt.</i>	On je iz mesta, je fina rit.	biti vaje višjega standarda (glede na govorca)
62.	<i>B_j dau no čāz rāt, mulā, dā pride h_pamātā.</i>	Bi ji dal eno čez rit, smrkliji (najstnici), da pride k pameti.	udariti, spametovati koga
63.	<i>Wās cajt mā žie wezi zā ritjo.</i>	Ves čas mi že vozi za ritjo.	ne upoštevati varnostne razdalje pri vožnji
64.	<i>Biāš z_buogām zluodivārt!</i>	Pojdi zбоgom zlodeju v rit!	psovka, zavrnitev in poslovilni pozdrav hkrati
65.	<i>Prāklieta rāt usrana!</i>	Prekleta rit usrana!	psovka
66.	<i>Sāmu še pavōvo pere tur rit ji manjka.</i>	Samo še pavovo pero v riti ji manjka.	biti pretirano okrašen, načičkan
67.	<i>Ne buom zdej kofe pila, mam žie tri tu ritā denās.</i>	Ne bom zdaj kave pila, imam že tri v riti danes.	zaužiti, imeti v sebi
68.	<i>Pazi de nā pojdaš pu ritā!</i>	Pazi, da ne pojdeš po riti!	pasti
69.	<i>Je blu strmu, je blu trieba iti pu ritā douskuos.</i>	Je bilo strmo, je bilo treba iti po riti navzdol.	premikati se z zadnjico po tleh
70.	<i>Je stisnu rāt, k_je blu trieba nāglas puviedātā.</i>	Je stisnil rit, ko je bilo treba naglas povedati.	umakniti se / ne nadaljevati aktivnosti zaradi strahu
71.	<i>Je ubljubu ārt an tričērt.</i>	Je obljubil rit in tričetrt.	preveč obljubiti / prazna obljuba

Tabela 2: Frazemi.

	Bovško v diskurzivni formi	Dobesedni knjižni prevod	Pomen
1.	<i>Diālo mo griā ud ruok, koj pāso smola ta ud ritā.</i>	Delo mu gre od rok kot psu smola od riti.	biti nezmožen, nesposoben za delo
2.	<i>Ma kuožo koj utruāška rt.</i>	Ima kožo kot otroška rit.	biti gladek
3.	<i>Ma guopc koj kraulja rāt.</i>	Ima gobec kot kravja rit.	govoriti grde, umazane stvari; jezicati, nevljudno odgovarjati
4.	<i>Ma rāt ku (krušna) pejč.</i>	Ima rit kot (krušna) peč.	velika zadnjica

5.	<i>Ma rət zə /kajk ən kəznienk</i>	Ima rit za / kolikor en kaznenik. ¹¹	zelo velika zadnjica
6.	<i>Sta koj rət an grbiaše.</i>	Sta kot rit in hlače.	vedno biti skupaj
7.	<i>Gueri, də niəma ne rit an ne glaue. Tu nima ne rit ne glawe.</i>	Govori (tako), da nima ne riti in ne glave. Ti nima ne riti ne glave.	ne imeti smisla/biti nerazumljivo, nepovezano
8.	<i>Je təma ku tu ritə.</i>	Je tema kot v riti.	biti zelo temno
9.	<i>Je uket, də b drek ta spuod rit pujədu.</i>	Požrešen (tako), da bi še drek izpd riti (komu/sebi) pojedel.	biti zelo požrešen, ješč
10.	<i>Ta srajca je, ku ta s_s kravlje ritə.</i>	Ta srajca je kot iz kravje riti.	zelo zmečkan
11.	<i>Je bla tiho ku rət.</i>	Je bila tiho kot rit.	nič reči, molčati
12.	<i>Je ko də_b mov sršiene tu ritə / gacəh.</i>	Je kot bi imel sršene v riti / spodnjicah.	biti nemiren, ne imeti obstanka
13.	<i>Je šu, ku d_b mo guriela rət.</i>	Je šel, kot da bi mu gorela rit.	zelo hitro iti/teči
14.	<i>Se je ubnašu, ku de buga ze rət / jajca drži.</i>	Se je obnašal, kot bi boga za rit / jajca držal.	biti vzvišen
15.	<i>Səm se ustrašu, də m_je kar gacə tu rət putəgnilo.</i>	Sem se (tako) ustrašil, da mi je kar gate v rit potegnilo.	zelo se prestrašiti
16.	<i>Ma ritce lih zə no piest.</i>	Ima ritke le za eno pest.	zelo majna zadnjica
17.	<i>Je snəga du ritə. Je snəga du ritə, če seiedəš nuotr.</i>	Snega je do riti. Snega je do riti, če sedeš vanj.	mera, kako visoko sega sneg
18.	<i>Ma šuobe ku kurja ərt.</i>	Ima ustnice kot kurja rit.	rdeče ustnice

Tabela 3: Primerjalni frazemi (primere/komparacije).

¹¹ V Pleteršnikovem slovarju je *kaznenik* opisan kot prostorninska mera (neka žitna mera: pol mernika ali četrt vagana). Na bovškem površinska mera, npr. njiva, velika en kaznenik. Informatorji so navajali različno površino (640 m², 250 m²); morda iz prvotnega pomena, koliko pridelka je dala ta površina.

The fox in Slovenian proverbs and sayings

Lisica v slovenskih pregovorih in rekih

The article is based on 103 proverbs about foxes collected in the electronic database of Slovenian paremiological units Proverbs 1.0 on Clarin.si. Slovenian proverbs and sayings about foxes usually depict them as tricksters, and they cover different experiences people encountered with foxes in everyday life. This paper focuses on fox-related proverbs from animal tales and fables. For the most part, the narratives about the fox are mostly international and not merely a reflection of Slovenian creativity. Their origins reflect the influence of foreign literature and art. Yet the proverbs, they don't possess the same international quality, as they use the character of a fox not only to criticise social and ethical conditions, but also reflect common experiences with this predatory animal. Proverbs considered in this paper fall under six international types of folktales.

♦ **Keywords:** proverb, fox, folkloristics, paremiology, fables

V prispevku so analizirani 103 pregovori o lisici, zbrani v elektronski bazi slovenskih paremioloških enot Pregovori 1.0 na Clarin.si. Slovenski pregovori in frazemi pogosto govorijo o zviijačnosti lisice, vendar pa tudi o različnih izkušnjah, ki so jih doživeli ljudje z lisico v vsakdanjem življenju. V tem prispevku bom predstavila pregovore, ki so povezani z nekaterimi živalskimi pravljicami in basnimi o lisici, ki so mednarodno razširjene in niso le odsev ustvarjalnosti v slovenskem etničnem prostoru, pač pa so nastajali pod vplivom tuje literature in umetnosti. Vendar pa pregovori niso tako mednarodni, kajti preko lika lisice niso izražali le kritike družbenih ter etičnih razmer, ampak tudi izkušnje, ki so jih imeli ljudje s to pogosto plenilsko živaljo v vsakdanjem življenju. Obravnavani bodo pregovori, povezani s šestimi mednarodnimi pravljicnimi tipi.

♦ **Ključne besede:** pregovor, lisica, folkloristika, paremiologija, basni

1 Introduction

Proverbs and sayings reflect communicative, metaphoric, semantic and cultural qualities of a specific language or dialect, which naturally dictates certain differences among proverbs of different nations. This is also reflected in proverbs and sayings about animals, or rather in paremiological units in which animals are mentioned. They speak of people's understanding of the animal world, their experiences with different animal species, as well as animal related stereotypes. Proverbs and wellerisms about animals also typically originate in epic and narrative tradition, particularly in fables.

Fox, raynard or vixen, is – it's safe to say – the most common animal hero in the European narrative tradition, particularly in fables, animal tales and parables. And yet – although proverbs can pose as elements of other folklore genres, where they sometimes function as their synopsis – fox related proverbs in Slovenian tradition are fewer than one would expect. In paremiological units, other animals are actually found more frequently than foxes.

Slovenian spoken language has preserved several short sayings about foxes, which are still being used today, including: *Lisici na rep obesiti* / To hang on a fox's tail, *Lisici na rep sestiti* / To sit on a fox's tail, or *Lisici na rep stopiti* / To step on a fox's tail. (Keber 1998: 106, 109). We often say metaphorically: *Potuhnjen kot lisjak* / Sneaky as a fox, *Zvit kot lisica* / Cunning as a fox. (Kropej 2007: 119), also *Lisica – tatica* / Fox – thief, etc.

2 Analysis of Slovenian paremiological units about foxes

The electronic database of Slovenian paremiological units on Clarin.si includes 103 proverbs, wellerisms and sayings about the fox or vixen. In Slovenian language, the word *lisica* is both gender-neutral and gender specific – indicating the feminine gender.¹ This term appears in most proverbs, there are just a handful of examples featuring the male counterpart – *lisjak*.²

Many proverbs portray foxes as cunning animals, sometimes even as tricksters:

[A fox sheds her hair, but does not get rid of tricks]:³

Lesica se dlak zlevi, zvijač se ne znebi.

Lisica dlako obrije, ne pa lisice.

Lisica izgubi dlako, a vicje ne.

Lesica minja dlaku, ma čud ne.

[A fox only steps on ice once]:

Lesica gre samo enkrat na led.

[A fox escaped from a trap, will not fall into a trap again]:⁴

Lisica iz pasti ušla, v past zopet ne bode šla.

Lisica iz pasti ušla, v past zopet ne pojde.

Lesica, ki je skopcu utekla, se ne ulovi več.

Lisica is pasti ušla, v' past zopet ne bode prišla.

A similar type of this proverb is recorded in 'The Matti Kuusi international type system and database of proverbs': *A fox is not taken twice in the same snare*. And also: *There is no catching the fox twice*. (Source 2).

[An old fox can't be fooled]:

Star lisjak se ne da ukaniti.

[A fox which only knows one hole is ignorant]:

¹ For clarity, proverbs using the gender-free and feminine form will be used female pronouns, and masculine forms will be used with male pronouns, thus reflecting the use of pronouns in the Slovenian language.

² This paper doesn't present all of the 130 Slovenian paremiological units about foxes, but instead focuses on different variants or types of Slovenian proverbs about foxes.

³ Similar to a Greek proverb: 'The fox changes its fur but not its habits.' (Source 1). Gaius Svetonius Tranquillus published a Latin proverb describing a roman emperor Vespasianus: *Vulpes pilum mutat, non mores* (Svetonius, Vespasianus, 16). Published also in the Slovenian collection of Latin proverbs (Maver 2016, 74–75),

⁴ There is a similar French proverb: 'The fox is not caught twice in the same snare.' (Source 1).

Tista lisica, ki pozna samo eno luknjo, je neuka.

[A fox even pretends to sleep, in order to catch a hen]:

Lisica tudi na videz spi, da kure ulovi.

Sometimes, however, proverbs point to the fact that even a sleeping fox won't catch the hen:

[A sleeping fox does not catch a hen]:

Lisica, ki spi, kokoši ne ujame.

Lisica, ki spi, ne ujame kokoši.

Speča lisica ni dobra lovica.

Speča lisica kokoši ne ujame.

Lesica, kadar spi, kokoši ne vjame.

Tudi lisici, če spi, kokoši ne pridejo pod zobe.

Če lisica dolgo spi, nobene kokoši ne ulovi.

Similar proverb is recorded in 'The Matti Kuusi international type system and database of proverbs': *When the fox sleeps no grapes fall in his mouth.* (Source 2).

Some proverbs indicate that a fox can still be caught, despite her slyness or cleverness:

[No matter how cunning the fox, it will get caught in a trap one day]:

Še tako zvita lisica se enkrat ujame v past.

Še največji lisjak se enkrat ujame.

Še tak lisjak se nazadnje ujame.

There are also some proverbs describing characteristic fox behaviour:

[Every fox praises her tail]:⁵

Vsaka lisica svoj rep hvali.

[A fox knows very well where to get food]:

Lesica dobro ve, kje kaj dobi.

Lisica dobro ve, če kaj dobi.

[When does a fox repent, if she eats chicken meat?]:

Kdaj se lesica kesa, ako je kurje meso.

[A fox walks foolishly but arrives wisely]:

Lisica neumno hodi ali pametno pride.

[Even in broad daylight, the fox comes to get the hen]:

Lisica tudi pri belem dnevu po kure pride.

[A fox never steals a hen near her den]:

Lisica nikoli ne vzame kure v bližini svojega brloga.

Lisica ne krade blizu doma.

There is a proverb with a similar meaning recorded in 'The Matti Kuusi international type system and database of proverbs': *The fox seldom preys near home, nor does Satan meddle with his own.* (Source 2).

⁵ There is a similar Italian proverb: 'Let every fox take care of his own tail.' (Source 1).

Some proverbs give advice on how to behave towards a fox and how to catch it:

[What the fox leaves in the trap, one should not take]:

Kar lisica pusti v pasti, za to se ne lasti.

Kar lisica pusti v pasti, tega si ne lasti.

[Do not sell the pelt while the fox is still in the den.]:

Ne prodajaj kože, dokler je lisica v brlogu.

[Whoever wants to catch a fox should hunt it with geese]:

Kdor hoče uloviti lisico, naj jo lovi na gosi.

Some proverbs express long-time experience with observation of foxes in relation to weather phenomena:

[If a fox runs across the ice, the ice will also hold people]:

Če lisica čez led beži, tudi ljudi drži.

Če lisica čez led beži, tudi človeka drži.

Weather proverbs featuring foxes usually rely on fox behaviour:

[If a fox barks a lot in December, a severe winter is to be expected]:

Če prosinca lisica laja, še huda zima prihaja.

Če prosinca zelo lisica laja, huda zima še potem nastaja.

[If the sun rises in a clear sky on Candlemas (2 February), the fox returns to her den]:

Če na svečnico sonce na jasnem vzhaja, lisica še v brlog zahaja.

[Bad weather on St. George's day drives the fox out of her den; but if the weather is nice, she won't come out for a fortnight]:

Sv. Jurja gerdo vreme lisico iz jazbine izžene; če je pa lepo, jo še 14 dni ne bo.

[On St. Gregory's day, the fox steps out the door; if the weather is bad, he is not going back in; if the weather is nice, he goes back home for a fortnight]:

O sv. Gregorju lisjak gre pred duri; če slabo je vreme, nazaj več ne gre; je vreme lepo, gre za štirinajst dni domov.

[After St. Matthias' day, a fox does not go home across ice]:

Po sv. Matiji ne gre lisjak čez led domov.

[If a fox barks, rain is coming]:

Če lisica laja, bode dež.

Some proverbs liken characteristics of foxes to those of women:

[A bridesmaid to a bridesmaid, a fox to a fox]:

Družici družica, lisici lisica.

[Sick women and cunning foxes eat the most chickens]:

Bolne ženice pa zvite lesice narveč kokuš snedo.

[God knows well enough, which fox's-tail to step on]:

Bog že ve, kateri lisici da stopi na rep.

In these cases, characteristics of foxes are compared to characteristics of women, unlike in languages where fox is a masculine noun and 'fox characteristics' aren't

reserved exclusively for women. In ‘The Matti Kuusi international type system and database of proverbs’, for instance, the following proverbs can be found:

The fox may grow grey, but never good.

An old fox needs not to be taught tricks.

You may know the fox by his tail. (Source 2).

3 Proverbs related to folktales

Slovenian proverbs and wellerisms about foxes include several examples, which relate to folk tales and songs. Some of them can even be sourced to fables and fairy tales, which can be classified according to the international tale type index.⁶ These particular proverbs exhibit the strongest connection between animal behaviour and human habits and weaknesses.

The following proverbs are related to fairy-tale types or fables:

[To a fox, grapes that are out of reach are sour]:

Lisici je grozdje, ki ga ne doseže, kislo.

[A fox served a wolf, and lost her fur]:

Služila lisica volku, pa je kožuh izgubila.

[Even clever roosters are eaten by the fox]:

Tudi pametne peteline požre lisica.

[When a wolf mourns, a fox celebrates]:

Kadar vuk žaluje, lisica praznuje.

[Over time, even a fox becomes a monk]:

S časom se tudi lisica pomeniši.

[If a hedgehog visits a fox, the fox must leave]:

Če pride jež gostovat k lisici, mora lisica oditi.

[Usually, foxes cower and hide when an eagle calls]:

Kadar se orel oglasi, se po navadi lisjaki strahopetno skrijejo.

And a wellerism:

[‘There are words,’ said the fox when she found a rosary on the road]:

‘So besede’ rekla je lisica, ko je našla na cesti paternošter.

The proverb *When the fox preaches, take care of your geese* is recorded in ‘The Matti Kuusi international type system and database of proverbs’ (Source 2). It relates to the folktale type ATU 62* Forbidden to Sit in Trees. There is no corresponding Slovenian proverb in Clarin.si database.

⁶ Uther 2004.

In fables and animal tales, which mostly go across borders, there is hardly any animal among the protagonists as popular as the fox. But this is more than just a reflection of the creative Slovenian imagination, as the narratives involving foxes were also influenced by foreign literature and art. It's no secret that the *Panchatantra*, a selection of tales from ancient India, includes a lot of interrelated animal fables in Sanskrit verse and prose, among them many tales and fables about a pair of jackals called Kalila and Dimna. Collections written by Phaedrus, Babrios and Avianus preserved the Aesopian fables, which formed the basis for the European narrative and proverbial traditions. Paremiological heritage and fables were even strongly influenced by the medieval and renaissance epics and subsequent literary sources.

Paremiological units in Slovenian electronic database Clarin.si, related to folktales, start with:

[Grapes out of reach are sour to a fox]:

Lisici je grozdje, ki ga ne doseže, kislo.

A wellerism with a similar meaning is recorded in 'The Matti Kuusi international type system and database of proverbs': '*Sour grapes*' as the fox said when he could not reach them. (Source 2).

This proverb is rooted in the folktale type **ATU 59 The Fox and the Sour Grapes**.

In Greco-Roman mythology, the fox often had negative symbolism, she was caught in a hostile relationship, namely with Dionysus, the god of wine, because foxes were seen as pests which destroy grapevine. Aristophanes and Ovid wrote of her love of grapes (Aristophanes, *Equites* 1077; Ovid, *Fasti* 4, 679–712). The Bible (Vp 2:15) influenced the image of vixens as destroyers of vineyards. In medieval tradition, foxes represent heretic and evil demons, but also a treacherous and unjust ruler, such as Herod. The fable 'The Fox and the Sour Grapes', tells of a fox who cannot reach the grapes and therefore tries to convince others that they are still sour. This tale was already popular in ancient Greece, it is one of the Aesop's fables (Perry 1965, 424, No. 15), and was documented later, around 1200, in *Roman de Renart* (XI, 257–333). In Slovenian folklore, one of its iterations can be found in the fable: 'Lisjak in grozdje' (The Fox and the Grapes), recorded by Anton Kosi in 1894:

One day the fox comes to the grapevine and sees beautiful grapes aplenty. The temptation is great, but how could he reach the fruit? He starts jumping around and climbing the grapevine, but all in vain. The grapes are too high and he can't reach them. Some sparrows are in a nearby elder tree, mocking his futile effort. When he notices this, the fox quickly turns away and says, 'I don't want these grapes, they are too sour and too bitter anyway,' and leaves. When he is gone, the sparrows quickly pick the vine so thoroughly that no fox has to climb the grapevine ever again.⁷

⁷ Slovenian original of this fable is published in Kropej Telban 2015: 154.



Figure 1: The Fox and Sour Grapes, oil on canvas. National Museum of Slovenia.

One of the most ancient fables is classified under the folktale type **ATU *56** The Fox and the Eagle**, and also this fable is another example of proverbs rooted in folktales. The first known version dates back to around 650 BC and is today known as the Archilochus's fable. The plot revolves around a broken agreement between a fox and an eagle. In fox's absence, the eagle steals her cubs, and the fox saves them by threatening to burn the tree with the eagle's nest.

The Slovenian proverb related to this is:

[Foxes usually cower and hide when an eagle calls]:

Kadar se orel oglasi, se po navadi lisjaki strahopetno skrijejo.

This proverb differs from the fable, but highlights the same discordance between foxes and eagles.

Among Slovenian narratives is documented a variant of this fable which was published by Jožef Eržen in 1854 as 'Orel in lisica' (The Eagle and the Fox):

The eagle steals the fox's cubs and takes them to his nest to feed his hatchlings. The fox begs the eagle not to kill her cubs. The eagle on top of a tall oak tree thinks that the fox cannot harm it in any way. So it scoffs at her request. What should the fox do? She hastens and gets some fire, places it near the oak with the eagle's nest, and prepares to set it on fire, nest and all. The eagle, however, fears the fox will destroy its young and returns the cubs unharmed. We have to live in peace, even with those weaker than ourselves, as they could also possibly cause us harm.⁸

Another Slovenian proverb which can be compared to the folktale **ATU 57 Raven with Cheese in his Mouth** is:

[Even clever roosters are eaten by the fox]:

Tudi pametne peteline požre lisica.

A raven (krokar) can be substituted with a rooster or a crow (vran) as protagonist. And the fable goes like this: A fox sees a crow with cheese in its beak. It wants to get the cheese, and asks the crow to sing. Flattered, the crow crows, drops the cheese, and the fox carries it away.

The Fox and the Raven is an Aesop's fable, first documented as early as the 7th century, and is also famous for being an integral part of the Latin epic Ysengrimus (c. 1150). In Slovenian Carinthia, this fable was first recorded by Urban Jarnik as 'Vrank in Lisiza' (The Crow and the Fox) in 1814.

There was a time when a crow took some cheese, hanging out a window, and flew with it to a branch of a tall spruce tree. The fox saw this and wanted some of the cheese. So it bowed and scraped to the crow, saying: Oh, my dear Crow! Who can match you? No bird has feathers like yours. And I don't know of any bird who has such beauty, accompanied by such a lovely voice. I know what I am talking about, your voice is truly the sweetest. The crow, listening to this praise, rejoiced and tried to show off even more, and demonstrate its voice, so it stretched out its beak and cawed. And because it opened its mouth, the cheese fell out. The fast and cunning fox took hold of it and quickly devoured it. The cheated Crow looked down placidly, realising the Fox's words were cunning and disloyal.

The moral: Beware of those who bow and scrape and overly praise, because they want to scam you.⁹

One of other Slovenian variants of this fable was published by bishop, poet and writer, Anton Martin Slomšek, who chose to write it in verse: 'Vrana in lesica' (The Crow and the Fox 1846).

Folktale **ATU 80 The Hedgehog and the Fox (or: The Hedgehog in the Badger's Den)** relates to the following proverb:

⁸ Slovenian original of this fable is published in Kropej Telban 2015: 151.

⁹ Slovenian original of this fable is published in Kropej Telban 2015: 152.

[When a hedgehog comes to the fox's den, the fox must leave]:

Če pride jež gostovat k lisici, mora lisica oditi.

A wellerism with similar meaning is recorded in 'The Matti Kuusi international type system and database of proverbs': *The fox knows many tricks, but the hedgehog one great one.* (Source 2).

This plot is found with Archilochus, as well, in a collection of fables from the 16th century (Laurentius Abstemius, no. 72). In Slovenia, this folktale was first published by Janez Svetokriški in his Sermons (Svetokriški 1691/1991, p. 175).

Anton Martin Slomšek provided another version of this plot in his fable 'Jež in lisica' (The Hedgehog and the Fox). It involves a freezing hedgehog asking a fox for shelter. The fox welcomes him into her lair, but when the hedgehog gets warm, it loosens, and throws the fox out of her lair.

Eight more variants of this folktale have been confirmed in Slovenia.¹⁰

Tales about the **Fox-pilgrim** relate to the following Slovenian proverb:

[Over time, even a fox becomes a monk]:

Sčasom se tudi lisica pomeniši.

And the wellerism,:

['There are words,' said the fox when she found a rosary on the road]:

*'So besede' rekla je lisica, ko je našla na cesti paternošter.*¹¹

This wellerism is unique to the European tradition, and is reminiscent of illustrations and paintings featuring a fox-pilgrim holding a rosary. A fine example of this depiction can be found on the fresco in the church of Saint Ulrich (Sv. Urh) in Maršiči village near Ortnek (in the vicinity of Velike Lašče). It was probably painted by the Croatian master Tomaž from Senj around 1515. A fox as a pilgrim, wearing a cardinal's hat, is the subject of the fresco in the presbytery of this church. There are two pairs of crossed keys on his hat, a sign of the Roman Pope, as well as a shell, a pilgrims' sign. The fox is holding a rosary, and has a pilgrim's bag on his shoulders with a dangling head of an 'accidentally strangled' hen.¹²

Iconography of the fox-pilgrim in Maršiči highlights weakness and expresses criticism of the court and the Catholic Church.

A medieval fable about 'Iisjak Reinhart' (Reynard the Fox) spread to the Slovenian cultural area between the 12th and the 15th century. Reynard represented sententious criticism of society. The German version in particular, the *Reyneke de vos* epic (1498), was

¹⁰ Slovenian variants of the 'Hedgehog and Fox' fable are listed in the Type Index of Slovenian Folktales: Animal Tales and Fables (Kropej Telban 2015: 191–193).

¹¹ This wellerism was collected by Janez Trdina in his manuscript 'Zapiski' 1870–1879 (27 handwritten notebooks). It was published in: Snežana Štabi, Igor Kramberger (eds.), *Podobe prednikov: zapiski Janeza Trdine iz obdobja 1870–1879*. 1–3 Ljubljana: KRT 1987. Book III 'Trezne vinske in pravoverne', p. 715. I am grateful to Rok Mrvič for sharing this information.

¹² For more on this, see Kropej 2015 and Komić Marn 2019.

directed toward the court and the Catholic church, which made Martin Luther describe this epic as a vivid representation of life at court. The French medieval fable *Roman de Renart* was equally scathing.

There were also texts and miniature graphics – leaflets or even playing cards with a fox as the most common central character – circling among people. Most times, the fox preaches to hens, chickens or geese, dressed as a monk and holding a pilgrim’s staff or a rosary. In the time of Gothic and the Renaissance, such scenes were not merely a humorous warning, but also a serious criticism and deterrent against hypocrites holding the highest secular and ecclesiastical positions. During the Protestant Reformation, these sorts of motives were most rampant in places where the reformation movement was at its most powerful.



Figure 2: The fresco of a fox-pilgrim in the church of Saint Ulrich in Maršiči near Ortnek (photo by M. Krojej Telban).

This kind of proverbs about foxes were also recorded in *Speculum exemplorum* (1, 15), and also Janez Svetokriški wrote about a penitent fox in his sermons (*Sacrum promptuarium* 1691–1707), noting that the penitent fox donned not just ‘pilgromski gvant’ (a pilgrim’s attire), but also ‘usnjat kreželjc’ (a leather collar) (Svetokriški 1707: 75).¹³

The influence of the *Reyneke de vos* epic also left an incredibly indelible mark in the central European area, where individual episodes were used as homiletic examples or textbook cases in legal matters,¹⁴ which could be the origin of our own Slovenian wellerism,¹⁵ as the same proverbs could be used in different semantic relations.

Folktales about **The wolf and the fox, stealing the food together (ATU 3, ATU 3*, ATU4)** relate to the following proverbs:

[When a wolf mourns, a fox celebrates]:

Kadar vuk žaluje, lesica praznuje. (ATU 4)

[At wolves’ side, even foxes are sated]:

Vz vuke so i lisice site.

[Shouting at the wolf - they feed the fox]:

Na volka kriče - lisico rede. (ATU 3, 3*, 4).

¹³ Komić Marn 2019: 120.

¹⁴ For more on this, see.: Komić Marn 2019: 126.

¹⁵ For more on this, see: Meterc, Mrvič 2024: in print.

These proverbs are related to folktales which are frequent among Slovenian animal tales, some of them are also included in the folktale index *The Types of International Folktales* (Uther 2004), among them:

ATU 3 **Simulated Injury**: A fox covers its head in blood or mud and scares a beaten wolf, saying that her head has been smashed and her brains are leaking.

ATU 4 **Sick Animal Carries the Healthy One**: A beaten wolf carries a fox, while the fox is singing: ‘The sick carries the healthy!’

ATU 3* **The Wolf Supplies Food for the Fox**: A fox advises a wolf to go into the hen den and sweep out all the hens. When people come, the fox picks up a hen and runs away, while the wolf is caught.¹⁶

Slovenian paremiological tradition also includes proverbs with the fox, not the wolf, getting the short end of the stick:

Hunting with wolves, foxes are always hungry.

A fox served a wolf, and lost her fur.

4 The fox in Slovenian dialectal phraseology

Linguists from the Austrian Carinthia have been researching proverbs and sayings about animals from the Slovenian ethnic territory, as well. Studies by Ludvik Karničar and Tatjana Trampusch focused mainly on sayings, comparative phrases and rhymes among Slovenians living in the Austrian Carinthia. Ludvik Karničar has been researching phraseology of Carinthian Slovenian dialects in Austria, and has only come upon a single humorous rhyme *falerka* (taunt with humorous content) including a fox. It comes from Sele / Zell in Rož / Rosental in Carinthia and also features a fox:

Lisica je vajava, volq j pa tulu, medved j jazbic svam u rit bulu. (Karničar 1994: 120). [Lisica je lajala, volk pa je tulil, medved je jazbecu slamo v rit tlačil.]

[The fox barked, the wolf howled, and the bear stuffed straw in the badger’s butt.]

Similarly, Tatjana Trampusch researched phraseology in spoken language – the Podjuna/Yuantal dialect) among residents of the village Dob near Pliberk in the Austrian Carinthia, and only registered comparatives between foxes and cunning, e.g. a saying about a very cunning person:

[She is cunning like a fox]:

Je z’wi:ta / na’wi:ta k na li’si:ca.

There is also a saying:

Je še ‘to:k lis’jo:k, p’ri:de nə’ba:rt u ‘po:st. [Še tak lisjak se nazadnje ujame.]

[Every fox gets caught eventually]:

There is a popular counting-out rhyme among children:

¹⁶ Slovenian variants of these folktales are presented in Kropelj Telban 2015: 54–73.

[A reynard and a vixen, were drinking mead, there was no mead, they drank water]:

Li'si:ca, lis'jo:k, sta pi:wa me'do:k, me'do:ka ni 'bu:ə, sta 'pi:wa wə'du:ə. [Lisica, lisjak, sta pila medico, medice ni bilo, sta pila vodó.] (Trampusch 1999: 115).

However, as Ludvik Karničar finds, most of these cases can be traced in everyday lives of Slovenians from the Austrian Carinthia but only in dialect. They usually cannot be traced in Slovenian literary lexicography and literature. (Karničar 1994: 126).

5 Conclusion

Many proverbs in use today have been passed down through generations, from century to century. Some proverbs, that were used for many years, are no longer part of today's language and can only be found as records on paper (Babič 2015: 50). Despite the fact that nowadays people rarely include proverbs or sayings about foxes in everyday speech, they have been somewhat preserved, which is corroborated by the case of Slovenian dialects in the Austrian Carinthia, where researchers find dialectal sayings and rhymes about animals with strong similarities to literary units. Nevertheless, associations enable spontaneous creations of variant forms (Trampusch 1999: 125), which may be a further reason why they have persisted to this day.

Proverbs and sayings discussed in this paper, also reflect metaphoric language, as well as semantic and cultural characteristics of the Slovenian language or dialect. According to findings by Saša Babič and Piret Voolaid, paremiological material is international on one hand, but on the other, it can provide not only data on the tangible heritage, but also on specific cultural experiences, views, and concepts. Paremiological units are adapted to each language and society and are used to fulfil the function of (successful idiomatic or metaphorical) communication. (Babič, Voolaid 2023: 388). It seems that the fox has played quite an important role in everyday lives of people living on the Slovenian territory, as relatively many proverbs and sayings have been preserved in comparison to other nations.

While it is true that the meaning and use of specific proverbs can vary in different cultures and environments (Lauhakangas 2013: 218), there is nevertheless a significant level of international integration, especially in the case of proverbs originating from fables and animal tales. However, there are few international proverbs about foxes or even proverbs merely mentioning foxes, when compared with fables, where the fox is a frequent protagonist.

As we can see the fox is a frequent protagonist in folk tales, songs and proverbs as well as in folk art. However, since people used the fox's image to express criticism of social and ethical conditions, their imagination flourished. But they express also their common experience in dealing with this predatory animal in everyday life, and this was often not positive, since foxes often caused damage to the farmers. Narratives as

well as proverbs about foxes were therefore an important part of narrative folklore in previous centuries. But nowadays they are used most often only to express the cunning character of the fox and as the metaphor for the similar behaviour of the people.

Acknowledgements

The research for this paper was supported by the Slovenian Research and Innovation Agency (ARIS): *Traditional Paremiological Units in Dialogue with Contemporary Use* (J6-2579) and *Ethnological, Anthropological and Folklore Studies Research on Everyday Life* (research core funding, P6-0088).

References

- Babič, Saša, 2015: *Beseda ni konj. Estetska struktura slovenskih folklornih obrazcev*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612547660>.
- Babič, Saša; Voolaid, Piret, 2023. A Sociocultural View of Estonian and Slovenian Proverbs on Alcohol and Drinking. *Folklore* 134/3, 370–394. DOI: <https://doi.org/10.1080/0015587X.2023.2175972>.
- Karničar, Ludvik, 1994: Živali v frazeologiji koroških slovenskih narečij. *Koledar Mohorjeve družbe v Celovcu za leto 1994*, 119–126.
- Keber, Janez, 1998: *Živali v prisposodbah 2*. Celje: Mohorjeva družba.
- Komić Marn, Renata, 2019: Upodobitev lisjaka romarja v cerkvi sv. Urha pri Maršičih v evropskem kontekstu. Ikonografski, likovni in zgodovinski viri. In: *Profano v sakralnem. Študije o vizualizaciji posvetnih teženj in motivov v sakralni umetnosti*. Oter Gorenčič, Mija (ed.). Ljubljana: Založba ZRC, ZRC SAZU.
- Kropej, Monika, 2007: Lisica. Njena vloga in sporočilnost v slovenskem in srednjeevropskem izročilu. *Traditiones* 36/2, 115–142. DOI: <https://doi.org/10.3986/Traditio2007360206>.
- Kropej Telban, Monika 2015: *Tipni indeks slovenskih ljudskih pravljic. Živalske pravljice in basni*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612547707>.
- Lauhakangas, Outi, 2013: The Matti Kuusi International Database of Proverbs. Introduction to Paremiology. *Oral Tradition* 28/2, 217–222.
- Maver, Aleš, 2016: *Kapljica izdolbe kamen. Latinski pregovori in reki*. Celje: Mohorjeva družba.
- Meterc, Matej; Mrvič, Rok, 2024: Zgradbene in pomenske značilnosti slovenskih velerizmov iz zbirke *Pregovori*. *Slavistična revija* 72/2, [v tisku]. DOI: <https://doi.org/10.57589/srl.v72i2.4183>.
- Trampusch, Tatjana, 1999: Živalska frazeologija v govoru vasi Dob pri Pliberku na avstrijskem Koroškem. *Slovenski jezik / Slovene linguistic Studies* 2, 109–217.
- Uther, Hans-Jörg, 2004: *The Types of International Folktales: A Classification and Bibliography. Based on the System of Antti Aarne and Stith Thompson I–III*. Helsinki: Academia Scientiarum Fennica (FF Communications, 284–286).

Online sources

Source 1: https://www.uniguide.com/fox-quotes-sayings-proverbs#Fox_Proverbs

Source 2: Proverb types | The Matti Kuusi international type system and database of proverbs (mattikuusiproverbtypology.fi)



Slovensko pregovorno izročilo o besednih zavezah: pregled kontekstualnih dejavnikov v izbranih paremioloških enotah

Slovene proverbial lore on verbal bonds: An overview of contextual factors in selected paremiological units

Razprava prinaša strnjene rezultate poskusne opredelitve zunajjezikovnih kontekstualnih dejavnikov, izpeljanih iz zgradbe paremioloških enot, zlasti rekov in pregovorov, ki tematizirajo obljubljanje, priseganje in nasploh različne oblike sklepanja besednih zavez. Na podlagi izbranih iskalnih pogojev smo iz paremiološke zbirke *Pregovori* izločili več kot 100 paremij, sestavine v njihovi zgradbi, ki vsebujejo podatke o kontekstih rabe zavez, pa interpretirali kot »drobce žanrskih pravil« kratkega folklornega žanra zavez. Takim »drobcom« pripisujemo svojevrstno vlogo v folklorističnih raziskavah, saj so se v različnih skupnostih slovenskega jezikovnega prostora ustalila in ohranila v razmeroma daljših časovnih obdobjih (1684–2006). S tem je opravljen preizkus paremiološkega pristopa k sodobnim folklorističnim raziskavam zgodovine tega žanra, s katerim se odpira nov pogled na podatke, ki jih druge vrste slovstvenofolklornega gradiva ne vsebujejo.

• **Ključne besede:** paremiologija, folkloristika, kratki folklorni žanri, klasifikacija pregovorov, klasifikacija kontekstualnih dejavnikov, paremiološka zbirka *Pregovori*

The article presents condensed results of a trial characterization of extralinguistic contextual factors derived from the structure of paremiological units, particularly sayings and proverbs, thematizing promises, vows, and various forms of verbal bonds. Based on selected search terms, over 100 paremias were extracted from the paremiological collection “*Pregovori*,” and the components in their structure containing data on the contexts of verbal bonding were interpreted as “fragments of genre rules”. Such “fragments” are attributed a unique role in folkloristic research as they have been preserved within a number of Slovene speech communities in the Slovenian language area over relatively long periods of time (1684–2006). This serves as a test of the paremiological approach to contemporary folkloristic research on the history of the genre of verbal bonds, opening up a new perspective on data not contained in other types of verbal folklore material.

• **Keywords:** paremiology, folklore studies, short folklore genres, classification of proverbs, classification of contextual factors, paremiological collection “*Pregovori*”

1 Uvod

Leta 2022 je bila objavljena zbirka *Pregovori*,¹ ki predstavlja najobsežnejšo zbirko slovenskega paremiološkega gradiva. Nove raziskovalne možnosti, ki jih je z objavo zbirke omogočilo hitro korpusno iskanje po več kot 2500 paremioloških virih, so vodile do postopnega oblikovanja raziskovalnega vprašanja, ki ga v okviru te razprave

¹ Zbirka je nastala v okviru projekta *Tradicionalne paremiološke enote v dialogu s sodobno rabo*. O podrobnem historiatu in preteklih stopnjah urejanja te paremiološke zbirke podrobneje v Stanonik 2015 in Babič 2022.

postavljamo v središče našega zanimanja: »Ali je mogoče iz sestavin v zgradbi paremij izločiti podatke o zunajjezikovnih oz. družbenokulturnih kontekstih rabe tekstov drugih kratkih slovstvenofolklornih žanrov?« V okviru širše raziskave besednih zvez, tj. folklornih tekstov, kot so obljube in prisege, ki jih posamezniki v vsakdanjem sporazumevanju uporabljajo za govorno izvedbo zagotovila o svojem trenutnem ali nadaljnjem ravnanju, želimo torej podatke, izločene iz zgradbe posamičnih paremij,² omejiti na proces folklornega sporazumevanja med vsaj dvema govorcema oz. akterjema (Jakobson, Bogatyrev 1971 [1929]; Bauman 1984; Stanonik 2001; Babič 2015). To pomeni, da bi zgolj na podlagi izpričanih paremij delno rekonstruirali rabo tekstov žanra besednih zvez in s tem v omejenem obsegu pridobili vpogled v starejša stanja njihove rabe, o katerih nimamo podatkov oz. so ti izrazito pomanjkljivi. Obenem bi ti podatki olajšali presojo o tem, kako so skupnosti označevale posamezne pragmatične kategorije, ki so vplivale na tovrstno folklorno sporazumevanje in kakšno vlogo so jim pripisovale.

Na tem mestu velja poudariti, da je folkloristično zavedanje o ključnem pomenu konkretnih okoliščin v trenutku uresničevanja oz. izvedbe folklornega teksta nenadomestljivo za natančno interpretacijo preučevanega pojava (Bauman 1984; Bauman, Briggs 1990; Ben-Amos 1993). To spoznanje velja enakovredno upoštevati, tudi ko govorimo o izvajanju »folklornih tekstov o folklornih tekstih«, torej o paremijah, ki tematizirajo sklepanje zvez, kar je svojevrstno medžanrsko razmerje v okviru folklorističnih obravnav paremiološkega gradiva (prim. Mieder 2004: 134; gl. tudi Kirshenblatt-Gimblett 1973). Z ozirom na nekatere dosedanje slovenske slovstvenofolkloristične raziskave se zdi pristop, ki snov navidez črpa le z ravnine folklornega teksta (Dundes 1980: 23), na prvi pogled pomanjkljiv in neskladen z doslej uveljavljenimi težnjami po karseda celoviti obravnavi folklornega dogodka (Stanonik 2001: 117, 257, 293; Ivančič Kutin 2011: 17) ali folklornega trenutka (Babič 2015: 12–16). Odmik od konkretnih uresničitvev folklornih tekstov v smer analize metapodatkovno skromno opremljene pregovorne zbirke je uzaveščen in s semiotične perspektive pomeni korak v smer pridobivanja manjkajočih podatkov o pragmatični razsežnosti znakov folklornega sistema (gl. Grzybek 1994, 2014), za katere je značilna zlasti večja časovna odmaknjenost. Zbirka ne vsebuje dodatnih podatkov o posamičnih paremijah; izjemo predstavljajo le nekateri metapodatki,³ medtem ko v zbirko vključene paremije iz najrazličnejših pisnih in ustnih virov⁴ znotrajjezikovnega konteksta oz. sobesedila ne vsebujejo. Čeprav

² Za strnjen pregled in opis posamičnih paremioloških žanrov gl. Meterc 2017.

³ Temeljne kategorije metapodatkov, s katerimi smo v okviru zbirke razpolagali pri pregledu gradiva, so štiri: a) vrsta vira (npr. »zbirka pregovorov«), b) ime zbiralca oz. zapisovalca (npr. »Bojc, Etbin«), c) naslov vira (npr. »Pregovori in reki na Slovenskem«) in d) letnica vira, ki je večinoma enaka letu objave vira (1974). Slednja je lahko v primerih ustnih in drugih neobjavljenih virov enaka letnici prvega zapisa paremije.

⁴ Korpus v najnovejši različici 1.1 obsega 37.633 paremij iz 2515 virov, ki so nastali oz. bili objavljeni v skoraj 500-letnem obdobju od 1578 do 2010. 20 % tega obsega sestavljajo načrtno sestavljene in vsaj delno urejene zbirke pregovorov; nadaljnjih 15 % predstavljajo različne periodične publikacije (revije, časopisi, glasila); strokovne in znanstvene razprave, v katerih je bilo opisano in klasificirano paremiološko gradivo,

kontekst izjavljanja paremij in drugi podatki o njihovi rabi niso na voljo, to ne pomeni, da različnih folklorističnih pristopov ne moremo podpreti z ugotovitvami in podatki pretežno semantične analize v okviru zgornjega raziskovalnega vprašanja. Pristop, ki smo ga pripravili ob tej priložnosti, je v določeni meri soroden etnolingvističnim in lingvokulturološkim pristopom (prim. s frazeološkim raziskovalnim okvirom v Telija 1996 in Kržišnik 2005), le da v nadaljevanju zbrani podatki predstavljajo kontekstualne dejavnike rabe folklornega žanra in obenem odražajo vrednote, verovanja, omejitve in druge zunajjezikovne dejavnike, ki so pomembni za sistemsko obravnavo in razumevanje folklornega žanra besednih zvez v posamičnih skupnostih.

Razprava temelji na ugotovitvah diahronih paremioloških raziskav, ki so izpostavile, da pregovori kot jezikovni in kulturni znaki kljub stalnim spremembam omogočajo akumulacijo in prenos drugih kulturnih znakov (Mieder 2004: 132, 135, 137; prim. Grzybek 1994) oz. kot je strnil Wolfgang Mieder:

Pregovori zapolnjujejo človeško potrebo po povzemanju izkušenj in opažanj v drobce modrosti, ki nam ponujajo predpripravljene komentarje medosebnih razmerij in družbenih razmer. Za vsak kontekst, ki si ga lahko zamislimo, obstaja pregovor, a kljub temu si pregovori pogosto nasprotujejo. (Mieder 2004: 1; prevedel avtor)

Da so se v zgradbi paremij lahko ohranili kulturni znaki oz. reference na pojave, s katerimi se je določena jezikovna skupnost srečevala, se na te pojave sklicevala v paremiološkem gradivu in jim s tem v vsakdanjem sporazumevanju pripisovala posebno vlogo, so morali ti pojavi preiti iz konkretnih jezikovnih uresničitvev v t. i. družbena dejstva (Jakobson, Bogatyrev 1971 [1929]: 370). To pomeni, da je morala ista skupnost te pojave sprva podvreči nekakšni »preventivni skupnostni cenzuri« (Jakobson, Bogatyrev 1971 [1929]: 373), s katero je po različnih merilih presojala o ustreznosti posamičnega teksta glede na obstoječo normo. Takemu »skupnostnemu vrednotenju« inovacije ter presoji njene ustreznosti glede na tekste v repertoarjih govorcev iste skupnosti in obstoječe norme je sledil sprejem ali zavrnitev teksta v rabi. Skupnost je tekstu z nadaljnjo rabo odmerila določeno mesto v folklornih repertoarjih in jo umestila v skupni folklorni sistem. Paremije v tem oziru pojmujejo kot sredstva »komunikacije o komunikaciji« oz. kot eno od številnih kulturno konvencionaliziranih metakomunikacijskih sredstev (Bauman 1984: 15–16). Vpliv družbenokulturnih sprememb na preoblikovanje meril skupnosti je pričakovan, zaradi česar je mogoče s semantično analizo sestavin v zgradbi paremioloških enot sklepati o posamičnih dejavnikih, ki so zaznamovali produkcijo in recepcijo žanrskih tekstov – vsaj v obsegu, v

tvorijo 9 %; 8 % gradiva je iz zasebnih zbirk, ki so pri ljubiteljskih zbiralcih nastajale povečini brez sistematičnega ključa za klasifikacijo gradiva; nadaljnjih 7 % predstavljajo pregovori, napaberkovani iz rokopisnih zapuščin; prav tako so 7 % prispevale zbirke zbiralnih akcij, ki so povečini potekale v okviru uredništev časopisov in revij. Gre za dve tretjini gradiva oz. 66 % celotnega korpusa – zadnja tretjina gradiva izhaja iz manjših priložnostnih virov, zlasti ustnih. Redki primeri so iz jezikoslovnih priročnikov, zlasti slovnice, najdemo pa lahko tudi gradivo, ki je bilo v korpus vključeno iz leposlovnih del, a to predstavlja le 3,9 % vsega gradiva.

kakršnem so se ti dejavniki vpisali v obravnavane tekste. Bolj kot žanrske spremembe v diahroni perspektivi nas bo v nadaljevanju zanimalo, ali je mogoče oblikovati in v obliki klasifikacije predstaviti koherenten nabor kontekstualnih dejavnikov, ki nam bi omogočili razmeroma dobro gradivsko podprto sklepanje o naravi folklornega žanra besednih zavez in žanrskih tekstov iz obdobj, o katerih nimamo drugih virov. Na tem mestu je ključno vprašanje, kaj pričakujemo od rezultatov takega pristopa in kako se jih odločimo prikazati po opravljeni klasifikaciji in analizi gradiva.

2 Metodologija

Zaradi nedostopnosti konkretnih kontekstov izjavljanja/izvajanja obravnavanih paremij, je pomembno, da se pri poenostavljenem in okrnjenem pojmovanju pomena paremije, kot ga omogoča gradivo, ne ravnamo glede na potencialne specifične funkcije, ki bi jih ta paremija lahko imela. Razločku med vsakokratnim pomenom *izvajanja* paremije in ustaljenim pomenom *paremije* kot take (Kirshenblatt-Gimblett 1973: 826; Grzybek 1994: 35) dodajamo, da pri semantični analizi upoštevamo koncept »semantičnega potenciala« posamične paremije po Arvu Krikmannu (1984), ki je ločeval med pomenom paremije na ravni možnega, tj. potencial, ki ga omogoča semantična zgradba paremije, in na ravni konkretnega, ko se ta potencial natančno opredeli in dobi nedvoumen pomen z govorno uresničitvijo paremije. To lahko ponazorimo s primerom pregovora, ki ga v nadaljevanju najdemo pod zaporedno številko 62: *Kdor je mož, malo obeta, pa veliko stori*. Če tekst postavimo v domnevno situacijo, v kateri je govorec B govorc A⁵ v istem dnevu obljubil, da bo opravil in presegel dodeljeno delo v treh dneh, lahko govorec A ob neizpolnitvi zaveze uporabi pregovor na različne načine, vsakokrat z nekoliko spremenjeno funkcijo glede na konkretne okoliščine, kar razberemo iz nekaj možnih interpretacij:

- ♦ A sporoča B-ju, da zanj ni mož, saj je obljubljal preveč in tega ni opravil ...
- ♦ A sporoča B-ju, da naj prihodnjič opravi dodeljeno delo brez obljube ...
- ♦ A sporoča B-ju, da naj prihodnjič obljubi, da bo opravil le dodeljeno delo in ga tudi opravi ...
- ♦ A sporoča B-ju, da lahko prihodnjič obljubi, da bo opravil in presegel dodeljeno delo, a le če ga zatem res opravi ...

Klasifikacijo paremij v jedru članka smo pripravili na podlagi zunajjezikovnih kontekstualnih dejavnikov, povezanih glede na razmerje z eno od treh temeljnih časovnih kategorij – te kategorije izhajajo iz delitve procesa jezikovnega oz. folklornega

⁵ V nadaljevanju sta oba akterja poimenovana skladno z metaforično predstavo o *dajanju* in *jemanju besede*, torej kot *besedodajalec* in *besedojemalec*, kar nam omogoča, da v ozkem slovstvenofolklorističnem žanrskem okviru sprti razjasnimo reference na specifično vlogo, ki jo ima posameznik v procesu sklepanja besedne zaveze.

sporazumevanja⁶ na tri stopnje, ki si zaporedno sledijo glede na časovne okoliščine sklepanja zaveze. Tako ločimo a) prvo stopnjo *pred* sklenitvijo zaveze, b) drugo stopnjo *med* sklepanjem zaveze in c) tretjo stopnjo *po* sklenitvi zaveze, ki navadno označuje posredne in neposredne posledice njene (ne)izpolnitve. Ni naključje, da je strukturalistična delitev, ki se zgleduje po temeljnih modelih sporazumevanja (prim. npr. Jakobson 1989: 153–159), hkrati skladna z modelom, povzetim po zgodnjih študijah ritualov in ritualnosti (prim. Turner 1978: 243 po van Gennepe 1909; gl. tudi Grimes 2010: 50). Po Victorju Turnerju bi namreč lahko (folklorni) trenutek izrekanja obljube ali prisege, s katero se govorec A zavezuje govorce B za določeno dejanje v sedanjosti ali prihodnosti, opredelili kot liminalno ali prehodno stopnjo v procesu sklepanja besedne zaveze. Čeprav bomo v nadaljnji klasifikaciji hitro prepoznali, da posamičnih paremij ne moremo vselej nedvoumno uvrstiti k zgolj eni stopnji, je ključno opazovanje semantične napetosti, ki se tvori na relaciji med dvema stopnjama, največkrat med prvo in tretjo ali med drugo in tretjo stopnjo folklornega sporazumevanja.

Naslednji razložek, ki smo ga morali uvesti za namen klasifikacije paremij, sledi pragmatičnojezikoslovni razpravi Vladimirja I. Karasika (2016: 62–63) o komisivih,⁷ v kateri je avtor s pomočjo paremiološkega gradiva pregledno strnil vrednotenje obljub in obljubam sorodnih tekstov v različnih jezikih. Pri tem je izpostavil orientacijo paremije, ki jasno opredeljuje delitev vlog v sporazumevanju med govorce oz. akterji, soudeleženi v procesu sklepanja besednih zavez, in sicer na govorce A, ki »da/daje besedo«, in govorce B, ki jo »vzame/jemlje« in s tem govorce A metaforično »drži za besedo« (gl. op. 5). Gre za drugo relacijo, na kateri se v procesu sklepanja besedne zaveze ustvarja semantična napetost. Orientacija posamične paremije na govorce je pomembna zlasti v kontekstu zgornjega Miederjevega citata o tem, da si pregovori pogosto nasprotujejo, saj se je izkazalo, da je večina dvoumnih ali navidez nasprotujočih si paremij obravnavane tematske skupine mogoče razjasniti ravno z opredelitvijo govorce, na čigar ravnanje se nanaša paremija.

Na podlagi teoretičnega okvira in etnografskega gradiva predhodnih raziskav smo oblikovali besedni seznam s 87 označevalci temeljnih označencev s področja sklepanja zavez. Znotraj seznama smo nadalje oblikovali temeljni nabor 47 iskalnih pogojev,⁸ da bi v zbirki Pregovori lahko poiskali paremije, v katerih se določena beseda pojavlja

⁶ Po Ericu W. Rothenbuhlerju (2016: 76–77), ki obravnava sporazumevanje kot ritual, bi tak proces lahko imenovali tudi ritualnosporazumevalni.

⁷ Komisivi so vrsta govornih dejanj v pragmatičnem jezikoslovju in predstavljajo nekoliko širši pojem, ki ni povsem prekriven s folklorističnim pojmovanjem žanra besednih zavez, a se z njim v veliki meri ujema.

⁸ Na seznam smo ob prvi obravnavi gradiva v zbirki Pregovori, ki je potekala od 22. avgusta 2023 do 1. septembra 2023, uvrstili naslednje ključne besede oz. iskalne pogoje, skladno s pričakovanimi spremembami na glasovni in oblikovni ravni različnih pregibalnih vzorcev: *priseg*/prisež**, *obljub**, *vez*/vež**, *vpreg*/vprež**, *obvez*/obvež**, *privez*/privež**, *spov**, *verig*/veriz**, *zareč*/zarek**, *obet*/obeč**, *strup**, *fig**, *držati*, *trn**, *vab**, *priduš**, *duš**, *pekel/pekl**, *preklet**, **klin*/*klim**, **klet**, *zapisati*, *zob**, *ust**, *dlan**, *rok*/roč**, *sladk**, *vrv**, *besed**, *štrik**, *dolg*/dolž**, *mož*, *kriv**, **vrač**, **vrn**.

kot paremiološka sestavina v zgradbi tistih paremij, ki tematizirajo žanr besednih zavez. Uporaba 47 iskalnih pogojev je s prav toliko iskalnih poizvedb ponudila 5785 konkordanc oz. paremij, v katerih smo po dvojnem pregledu kot tematsko ustrezne opredelili 941 paremij, jih nato ob natančnem ponovnem pregledu 329 označili za neustrezne⁹ in s tem izoblikovali širši nabor relevantnih paremij, ki obsega 612 enot. Na tej točki smo prepoznali potrebo po uvedbi dihotomije, ki obsega a) jedrne paremije, v katerih sestavine vzpostavljajo povezavo z elementi katerekoli od treh faz sklepanja besednih zavez in tvorijo koherentno tematsko skupino tekstov, ter b) obrobne paremije, v katerih sestavine vzpostavljajo povezavo s sorodnimi, bližnjimi koncepti, ki nam lahko ponudijo veliko podatkov o besednih zvezah, a ne poimenujejo konkretnih okoliščin te jezikovno-kulturne prakse. Omenjenih 612 enot torej predstavlja seštevka jedrnih (224) in obrobnih (388) paremij glede na povezavo z obravnavanim folklornim žanrom zavez.

V nadaljevanju navedene paremije so oštevilčene po zaporedju, v katerem si sledijo, pri čemer smo jim v oklepaju dopisali tri metapodatke iz paremiološke zbirke *Pregovori*: a) prvi podatek vsebuje leto ali časovni razpon zabeležitev paremije v pisnih virih korpusa, b) v drugem podatku je povzeto število korpusnih virov, v katerih najdemo paremijo, c) tretji podatek pa podaja enkratno identifikacijsko oznako paremije, s katero lahko posamični tekst naknadno preverimo v korpusu *Pregovori*. Vrstni red navajanja paremij znotraj kategorij uravnava kronološki kriterij, tako da si posamične paremije sledijo od tistih v najstarejših virih do tistih v najmlajših. Če je bilo mogoče, smo pri tem upoštevali tudi število virov in njihov časovni razpon. Edine spremembe v navedbah gradiva so manjše pravopisne prilagoditve, ki obsegajo spremembo ali vstavke ločil, kjer je bilo to potrebno. Navedbe paremij v starejšem črkopisu (npr. v bohoričici) so ohranjene v obliki, kot smo jih našli v zbirki *Pregovori*. Ker je vsem primerom paremij dodana identifikacijska številka, jih je v zbirki mogoče hitro poiskati kljub manjšim pravopisnim prilagoditvam.

3 Klasifikacija paremiološkega gradiva

3.1 Prva stopnja in paremije o ravnanju pred sklenitvijo zaveze

Poudarek paremij, v katerih se vsebina osredinja na stopnjo pred sklenitvijo zaveze, se največkrat nanaša na previdnost ter naslovnika kot potencialnega besedodajalca spodbuja k premisleku in tehtnosti:

1. *Desetkrat obrni besedo, preden jo izrečeš.* (1947–1985, 3; 01861)
2. *Besedo daj na tehtnico, preden jo izgovoriš.* (2000, 1; 16359)

⁹ V več kot dveh tretjinah izločenih primerov gre za vremenske pregovore, v katerih določeno dogajanje v sedanosti obeta, tj. zanesljivo napoveduje dogajanje v prihodnosti, npr. *Sv. Marjeta grom in strelo obeta* ali *Če marec prah okrog pometa, nam dobro letino obeta*.

Oba primera se lahko nanašata tudi na splošne posledice izrečenih besed – v kontekstih, kot smo jih posredno razbrali iz vsebine obrobni paremij, se lahko ta tip paremij nanaša na neslutene posledice govorjenja nasploh, najpogosteje na a) izdajanje informacij zaupne narave, b) nehoteno razžalitev sogovornika ter c) večjo učinkovitost pri sporazumevanju s sogovornikom, tj. izogib besedičenju. Pri obravnavi tovrstnih primerov množstvo interpretacij ni izključujoče, temveč celo nasprotno – glede na poli-funktionalnost in heterosituativnost paremij je lahko ravno njihova vsebina pripomogla k razširjenosti in funkcionalnosti paremije v različnih kontekstih (gl. npr. Krikmann 1984; Mieder 1993: 4–5;), o čemer pa, kot že rečeno, zgolj na podlagi tekstov ne moremo sklepati. Tehten premislek lahko govorcu pomaga dognati zanj ugodne okoliščine za izvedbo zaveze, zaradi česar zaveza ne sme biti prehitra in nepremišljena – svarilo pred preuranjenostjo lahko razumemo tudi kot spodbudo k preudarku o tem, ali je sklenitev določene zaveze sploh potrebna:

3. *Kdor obljube ne prevdari, se spominja poznejši izgube.* (1852–1859, 1; 20708)

V primerih dvoma o smiselnosti sklenitve zaveze in zmožnosti njene izpolnitve, je na prvi stopnji zavezo boljše pravočasno zavrniti oz. se ji izogniti in se s tem izogniti tudi bremenu in posledicam, ki jih lahko uveljavlja besedojemalec:

4. *Gorši je prej odreči, ko besedo snesti potlej.* (1852–1859, 1; 20725)

5. *Kdor nič ne obljubi, mu ni treba nič izpolniti.* (1987, 1; 29716)

Del paremij se nanaša tudi na razmerje med trenutkom sklenitve zaveze in trenutkom njene uresničitve oz. izpolnitve – če se denimo posameznik zaveže, da bo izpolnil nekaj, kar je že opravil, ne more utrpeti zanj neugodnih posledic:

6. *Obljubo treba je pozno delati zgodaj spolniti.* (1987, 1; 29191)

7. *Pred storiti, potem šele obljubiti.* (1987, 1; 29192)

8. *Počasi obetaj, hitro stori.* (1987, 1; 29715)

Ko se posameznik odloča, ali bo sklenil zavezo, mora preudariti, komu se zavezuje oz. s čim, torej h kakšnemu dejanju/ravnanju se zares zavezuje:

9. *Kar ne ljubiš, s tem se ne veži!* (1900, 1; 13610)

3.2 Druga stopnja in paremije o ravnanju med sklenitvijo zaveze

Med pomembne zunajjezikovne dejavnike, ki vplivajo na izvedbo in izpolnitev zaveze, sodi alkohol, kar predstavlja pomembno stičišče raziskav o besednih zvezah in raziskav o splošnem pojmovanju alkohola v paremiologiji (gl. Babič 2022; Babič, Voolaid 2023). Opiti besedodajalec ne more doseči idealiziranega statusa »moža besede«, saj bo na dano zavezo (ali njene dele) najverjetneje pozabil, verjetno pa jo bo v trenutku izrekanja sklenil celo s pretiravanjem njenega učinka oz. posledic:

10. *Vino ni mož beseda, je preveč pozabljivo.* (1987, 1; 30576)

11. *Vino veliko obeta, pa malo da.* (1987, 1; 30027)

Nadalje se je za pomembnega izkazal verbalni in neverbalni kulturni znak *fige*, ki se je razvijal tudi v tesnem odnosu do žanra besednih zvez. Figa na širšem slovanskem

območju velja za simbol koitusa ter pomembno in razširjeno apotropejsko sredstvo, ki je med drugim ščitilo pred neugodnimi posledicami lažnih zavez oz. krivih priseg (Levkievskaja 2004: 26–27). Vlogo tistih, ki so se takega sredstva prekomerno posluževali in s tem izničevali veljavo lastnih besed, tematizira več paremij:

12. *Ko je zajc u bregu in tat pri prisegi, vsim figo molita.* (1852–1859, 1; 21304)
13. *Kdor veliko obljudi, pa malo stori, je figa-mož.* (1885, 1; 03872)
14. *Spredaj z medom maže, zadaj fige kaže.* (1934–1984, 3; 10469)
15. *Kjer figa raste, tam ni moža.* (2000, 1; 14662)

V gradivu imata vidno vlogo tudi *med* in *sladkor* kot metafori »lepih besed« besedodajalca, ki pa v resnici ne prinašajo nič dobrega. S tem se vzpostavlja večja skupina paremij, ki besedojemalca svari pred olepšanim in lažnim obljubljanjem, v splošnem pa tudi pred hlinjenjem in neiskrenim pridobivanjem naklonjenosti:

16. *Beseda zapeljiva je sladka ko sterd, kdor jo pa posluša, mu je pekel odpert.* (1852–1859, 1; 19912)
17. *Kdor med obeta, trnje spleta.* (1934–1984, 3; 07284)
18. *Na med se muhe love, na sladke besede ljudje.* (1952–2000, 5; 11818)
19. *Besedi sladki in medeni nikdar ne verjemi.* (2000, 1; 16337)

Obljubljanje in pridobivanje naklonjenosti s komplimenti je primerjano tudi z ribolovom, kjer je beseda ekvivalent metaforičnemu trnku, s katerim si besedodajalec lahko na svojo stran pridobi besedojemalca. Pri tem ravno trnek, ki simbolizira slab konec za ribo, označuje neiskrenost in škodljivost besedodajalca:

20. *Pravo ribo uloviš na trnek, človeško ribo pa na lepe besede.* (1983–1985, 1; 11589)
21. *Ribe lovimo na trnek, ljudi na besede.* (2000, 1; 35772)

Poseben je spodnji frazem, ki sicer odstopa od ostalega paremiološkega gradiva, odraža pa skrajno zadržanost do očitno zavajajočega obljubljanja in ima lahko celo humorni učinek, ki gradi na oksimoronu biološke nezdržljivosti plodu z drevesom, na katerem raste:

22. *Obeta hruške na vrbi.* (1934, 1; 16906)

Del paremij izraža skepsa besedojemalca v odnosu do besedodajalca in njegovih obljub – ali po številu (sklepanje več zavez, ki ustvarja kvantitativno razsežnost) ali po navidezno ugodni nesorazmernosti vsebine zavez za besedojemalca (kvalitativna rasežnost). Gre za neposredno negativno vrednotenje besedodajalca, njegova ocena pa temelji na razmerju med dano zavezo (besedo) in njeno izpolnitvijo (dejanjem):

23. *Kdor veliko obeta, pa malo da.* (1852–1859, 1; 19521)
24. *Kdor veliko obeta, malo da.* (1852–1987, 2; 21226)
25. *Od velikih obetov in dolgih svetov je malo upati.* (1852–1974, 3; 07565)
26. *Kdor dela veliko obljub, rad na vse pozabi.* (1987, 1; 29180)
27. *Kdor večkrat obljublja, malokrat da.* (2000, 1; 35889)

Zgornja skupina paremij (23–27) o posamezniku, ki veliko oz. preveč obljublja, se v delu vsebine delno prekriva in ujema s spodnjo skupino paremij (28–30). V spodnji skupini se znova pojavi možnost več interpretacij – paremije se namreč lahko nanašajo na sklepanje besednih zavez, a njihovo vsebino lahko brez dodatnega konteksta

enako verjetno interpretiramo tudi kot splošno negativno vrednotenje prekomernega govorjenja oz. klepetanja, ki zmanjšuje zmogljivost soudeleženih v delu (prim. z 31):

28. *Stvora malo, kjer je mnogo besedi.* (1887–1974; 06096)

29. *Mnogo besedi, malo prida.* (1934, 1; 33582)

30. *Dosti besed – malo prida.* (2003, 1; 35525)

Spodnji primer potrjuje predstavo, da je opravljenega manj dela, kjer posamezniki klepetajo in se pri delu raje kratkočasijo, zato ga ne moremo povezati s pregovori o besednih zvezah:

31. *Kdor besedo vleče, mu tudi delo iz rok neče.* (1852–1859, 1; 19989)

Zgornjim primerom (28–30), ki govorijo o določeni vrsti pretiravanja, lahko poiščemo analoške primere, v katerih učinkuje mehanizem antonimije, ki se vzpostavlja med prvim in drugim stavkom povedi – tudi te smo v različnih kontekstih našli v povezavi s pretiranim obljubljanjem, ki ne privede do obetane rezultata, a jih zgolj na ravni teksta ne moremo nedvoumno povezati le s sklepanjem besednih zvez:

32. *Mnogo dima, malo ognja.* (1914–1968, 2; 03073)

33. *Velik grom, majhen dež.* (1927–1938, 3; 28822)

Veliko bolj neposredno lahko komentar povezave med pretirano oz. prekomerno uporabo tekstov besednih zvez, ki povzroča nasproten učinek, da se torej zaupanje izgublja, namesto da bi se z ustrezno rabo teksta zaveze krepilo:

34. *Kdor prerad obeta, zaupanje drugih od hiše pometa.* (1852–1859, 1; 19221)

35. *Kdor se vedno priduša, temu vest shujša.* (1887–1940, 3; 05244)

36. *Kdor se rad priduša, rad laže.* (1987, 1; 29714)

Razmerje med besedami in dejanji je v gradivu odkrito predstavljeno na več mestih; najizraziteje z nizom paremij, v katerih se odraža kontrast med lahkostjo besede in težo dejanja:

37. *Ne besede – dejanja govore.* (1974, 1; 08607)

38. *Od besed do dejanja je pot dolga.* (1983, 1; 24114)

39. *Dejanja povedo več kot besede.* (2001, 1; 27320)

40. *Boljše je enkrat pomagati z roko kakor dvakrat z usti.* (2002, 1; 29487)

41. *Dejanja so glasnejša od besed.* (2003, 1; 35524)

42. *Lahko obljubiti – težko storiti.* (2006, 1; 25750)

Naslednje, prav tako dihotomično razmerje, ki tvori samostojno kategorijo z izrazi-tim razmerjem med drugo in tretjo stopnjo besednih zvez, je razmerje med pisnim in govornim sklepanjem zvez, ki se po dolgoročnem vplivu in resnosti posledic nagiba na stran pisnega, zaradi česar je slednje bolj cenjeno. Spodnje paremije torej vsebujejo podatek o tem, *kako* je lahko v določeni skupnosti sklenjena zaveza:

43. *Tinteži so vinteži.* (1934–1974, 2; 07132)

44. *Pisana beseda ostane.* (1974, 1; 07524)

45. *Pismo velja, beseda nič.* (1987, 1; 29056)

46. *V tinti se je že marsikteri utopil.* (1987, 1; 30280)

47. *Papir je že požrl več gruntov kakor voda in ogenj.* (1987, 1; 29131)

48. *Papir drži.* (2000, 1; 16453)

Kot izrazit kontrast se vzpostavlja tudi ena od obrobnih kategorij, ki poudarja nepomembnost, minljivost in nemoč govornice besede:

49. *Beseda je sapa.* (1868–1912, 1; 34405)

50. *Beseda je veter – tebi iz ust, meni mimo ušes.* (1974, 1; 08931)

51. *Beseda se kmali da – nič ne velja.* (1987, 1; 29193)

3.3 Tretja stopnja in paremije o ravnanju po sklenitvi zaveze

Opazovanje posameznikov v njihovem preteklem izpolnjevanju zavez se v paremiološkem gradivu odraža v označevanju posameznika na podlagi njegovega ugleda. Ugled med drugim izhaja iz performansa, tj. izvedbe tekstov zavez v konkretnih okoliščinah, ki lahko neposredno vpliva na posameznikovo sklepanje zavez v prihodnosti. Sankcija skupnosti za posameznika v vlogi besedodajalca pomeni izgubo zaupanja, če se pri njegovem ravnanju pojavi razkorak med izrečenim in storjenim:

52. *Praznih miz in praznih besed se človek hitro naveliča.* (1963: 02429)

V kontekstu tradicionalne kulture je treba v tem segmentu gradiva poudariti, da je besedno zavezo, dogovor, kupčijo ipd. izpolnjeval pretežno gospodar, ki je bil mož-ki. Več paremij označuje soodvisnost med njegovim družbenim statusom in veljavo njegovih besed:

53. *Kakrjhin je mósh, taka je befféda.* (1684, 1; 03223)

54. *Tič se po petji, človek po besedi spozná.* (1864, 1; 18097)

55. *Kakršen mož, taka beseda.* (1929–1974, 1; 08800)

Mož beseda, posameznik, ki je zvest svojim zavezam in jih v odnosu do besedoje-malca izpolni, je dobrodošel v skupnosti, ki ji pripada:

56. *Mož beseda brez premisleka – preseda.* (1974, 1; 10732)

57. *Pošteniga moža vsi štímajo.* (1999, 1; 36962)

Positivna podoba in družbeni status takega posameznika sta izpostavljena v številnih primerah:

58. *Mož – beseda več velja, kakor kup zlata.* (1876, 1; 03715)

59. *Mož beseda več velja kakor kup srebra.* (1934–1974, 2; 07331)

60. *Mož beseda velja.* (1934, 1; 33587)

61. *Mož beseda mora biti le tisti, ki se pogodi, ne pa tak, ki se pogaja.* (1987, 1; 30260)

62. *Kdor je mož, malo obeta, pa veliko stori.* (1987, 1; 29182)

Če gradivo, povezano z izjavljanjem, torej z drugo stopnjo sklepanja besednih zavez, svari pred olepševanjem in pretiravanjem (23–27), takšno sklepanje zavez povzroča tudi negativno karakterizacijo tistih pripadnikov skupnosti, ki zaradi svoje nezanesljivosti izgubljajo ugled. Spodnje paremije opozarjajo na nevarnost sklepanja zavez s takšnimi besedodajalci:

63. *Boj se tega, ki rad prisega.* (1887–2001, 5; 04911)

64. *Boj se sladkega človeka.* (1993, 1; 03105)

65. *Boj se tistega, ki rad prisega.* (1996–1999, 2; 13470)

66. *S fig-fag človekom ni se meniti, najboljšo se mu je daleč ogniti.* (2001, 1; 23681)

Sloves dvomljivega človeka, na čigar zaveze se ne gre zanašati, označuje tudi primerjalni frazem:

67. *Drži besedo, ko canja vodo.* (1852–1859, 1; 20117)

Posebno mesto v zbirki imajo paremije, ki med posameznike s posebnim odnosom do sklepanja besednih zavez štejejo bolnike. Zanje je značilna šibka moč zaveze, saj imajo zaradi krhkosti zdravja izmed vseh izpostavljenih kategorij besedodajalcev najmanj vpliva na svoje ravnanje:

68. *Bolnik obljubi vse, samo da ozdravi.* (2001, 1; 36534)

69. *Bolnikovih besed ne tehtajmo na tehtnici za zlato.* (2001, 1; 36548)

Spodnji pregovor osvetljuje človeški pohlep in daje uvid, kako lahko posamezniki gledajo na zaveze, ki se izpolnijo v povezavi s smrtjo posameznika:

70. *Kdor po smerti obeta, mu smert žele.* (1852–1859, 1; 20701)

Zaupanje obljubam dvomljivih, obrobni posameznikov v izbranem gradivu izkazujejo neprevidni in naivni ljudje.¹⁰ Naivnost je primerjana z norostjo, naivnež pa z norcem. S tem skupnost ne kaznuje le tistega, ki daje dvomljive obljube, temveč tudi tistega, ki presoji skupnosti pri označevanju posameznika ne zaupa:

71. *Obljubvanje, obljubvanje, norca verovanje.* (1983, 1; 11736)

V spodnjih primerih je poudarek na rezultatu zaveze, ki je označen kot dolg, ki ga mora besedodajalec poplačati ob izpolnitvi zaveze. Vsem paremijam te kategorije je skupno poudarjanje resnosti sklenjene zaveze:

72. *Obljuba dolg dela.* (1825–1999, 21; 01125)

73. *Obet dolg dela.* (1887–1934, 2; 05864)

74. *Obetanje, najtežje dolgovanje.* (1902, 1; 18374)

75. *Beseda navadno dolg dela.* (1974, 1; 08850)

Različno pojmovanje izpolnitve med govorci se kaže na neposredni relaciji med drugo stopnjo (kaj obljubiš?) in tretjo stopnjo (s čim obljubljeni izpolniš?). Spodnja paremija odraža poskus posameznika, da bi poiskal bližnjico in se karseda hitro in enostavno rešil bremena zaveze:

76. *Kdor obljublja moko, misli, da je dovolj, če da otrobe.* (2004, 1; 21527)

Sorodna je paremija, ki kaže na pohlep posameznika, ki kljub dani zavezi spreminja svoje ravnanje zaradi vpliva denarja:

77. *Kadar zlato govori, vsaka beseda slabi.* (1832–1983, 7; 00914)

Gotovost se v paremijah pojavlja kot mehanizem pomiritve posameznika, ki je zadovoljen s tistim, kar je zagotovljeno in kar ima, namesto da bi se oziral po obljubljenem. Gre za kategorijo, ki je izrazito ambivalentna, saj lahko služi tudi kot svarilo,

¹⁰ Posebno podskupino predstavljajo otroci, ki bodo zlasti v vlogi besedodajalcev obravnavani v nadaljevanju. Za primerjalni vidik o tem, kako se v splošnem kaže podoba otrok v paremiološkem gradivu gl. Huzjan 2024, 75–92.

naj besedojemalec ne sklepa zavez, ki bi dajale lažno upanje, zaradi česar bi spodnje primere lahko uvrstili tudi h gradivu prve stopnje:

78. *Bolj je kos terdega kruha v roki kakor obljubljen hleb.* (1852–1859, 1; 18957)

79. *Tisti denar je gotov, ki ga človek v pesti drži.* (1934–1974, 2; 08826)

80. *Groš v roki je dostikrat več vreden ko sto obljubljenih goldinarjev.* (1987, 1; 29174)

81. *Suha skorja bolj nasiti ko obljubljena pogača.* (1987, 1; 29185)

Metafore za oprijemljivost vezi, ki se s sklenjeno zavezo vzpostavi med posameznikoma, se z določeno mero naklonjenosti nagibajo k besedojemalcu, ki ima v rokah vajeti ali roge, besedodajalec pa je enačen z živino, z nekom, ki je v izrazito podrejenem položaju, s katerim upravlja drugi:

82. *Možje se vežejo z jeziki, voli s štriki.* (1934–1974, 2; 09868)

83. *Konja za uzdo, moža za besedo držijo.* (1934–2001, 3; 08328)

84. *Konja za brzdo, moža za besedo držijo.* (1983–1985, 1; 12444)

85. *Vol se veže za roge, človek za jezik.* (1986, 1; 04309)

86. *Vola primeš za roge, človeka za besedo.* (2003, 1; 35527)

Ravnanje besedodajalca, ki si izprosi, kar želi, ko pride čas za povračilo, pa svoje ravnanje povsem spremeni, tvori posebno kategorijo paremij, ki so povezane s posojjo denarja in z zavezami pretežno finančnega značaja:

87. *Kdor na posodo prosi, zlata usta nosi; kdor vračuje, hrbet obračuje.* (1887–1934, 2; 05327)

88. *Dokler prosi, zlata usta nosi; ko vrača, hrbet obrača.* (1974, 1; 06675)

89. *Kdor prosi, zlata usta nosi, kdor vrača, hrbet obrača.* (1975, 1; 11500)

Splošno zavedanje resnosti posledic se kaže v paremijah, ki opozarjajo na izpolnjevanje dane zaveze; v pravnih kontekstih je lahko kazen tudi smrtna:

90. *Kriva prisega človeku po življenju sega.* (1868–1912, 1; 34546)

91. *Zmeron drži basedo, ki jo daš!* (1998, 1; 16041)

Ustaljena besedna zveza *zarečeni kruh* ima kot znak, ki povzema izkušnjo pre-nagljenega sklepanja zavez, posebno mesto med paremiološkim gradivom. Resda je zvezo pogosto mogoče uporabiti za označevanje pre-nagljenih izjav nasploh, a ker se najpogosteje rabi v nikalnih zvezah, torej da se govorec zaveže k nedejanju česa, pa stori ravno to,¹¹ smo skupino uvrstili v obravnavo, predstavljamo pa jo le s tremi spodnjimi primeri, saj je skupno relevantnih pregovorov in njihovih variant s sestavino »zarečen« kar 32 in predstavljajo zabeležene primere rabe v časovnem razponu od leta 1789 do leta 2009.

92. *Sarezhen kruh fe rad rieshe.* (1789, 1; 00012)

93. *Zarečenega kruha veliko se sne.* (1852–1934, 3; 06282)

94. *Zarečenega in zapečenega kruha je največ snedenega.* (1934, 1; 33901)

¹¹ Enako na sodobnih primerih iz besedilnih korpusov navaja SPP, in sicer pregovor *Zarečenega kruha se največ poje* »opisuje situacijo, ko kdo deluje v nasprotju z lastno izjavo, obljubo, zagotovilom; izraža, da je težko predvideti, da kdo česa ne bi storil, počel« (SPP).

3.4 Vpisanost starostne, spolne in socialne diferenciacije v paremiološkem gradivu

V zadnjem delu klasifikacije izpostavljamo paremiološko gradivo, ki bi ga lahko uvrstili v eno od zgoraj navedenih stopenj izvajanja besednih zvez, a je z zadostnim številom tekstov v zbirki izoblikovalo nove kategorije, na katerih bomo ponazorili dinamiko med folklornim žanrom in tremi dihotomijami, in sicer a) *starostno* med odraslimi in otroki, b) *spolno* med moškimi in ženskami ter c) *socialno* med nižjimi in višjimi družbenimi sloji.

Otroci so v paremiološkem gradivu o besednih zvezah vedno poimenovani spolno nezaznamovano in pretežno v množini. Prva je izpostavljena percepcija otroške naivnosti in pozabljivosti, ki je hkrati edini primer, ko lažno, neiskreno obljubljanje odraslih ni sankcionirano (prim. s siceršnjim odnosom do otrok v Huzjan 2024):

95. *Po zimi se otrokom češnje obetajo.* (1852–1859, 1; 19699)

96. *Po zimi otrokom črešnje obetamo.* (1887, 1; 05782)

Primer najverjetneje odraža skupni vzrok, ki je povzročil tudi, da drugih primerov, v katerih bi bila izpričana vloga otrok v procesih sklepanja zvez, v zbirki nismo našli. Vloga otrok je tako v celoti potisnjena na stran, obljubljanje, priseganje in nasploh zavezovanje začne posameznik preizkušati kot mladostnik, in sicer po paremiološkem gradivu izključno v ljubezensko-intimnih kontekstih, v katerih se izreka o svoji domnevni zvestobi. Gre za nastavke nekakšne žanrske kompetence, ki jih lahko spremljamo vse do preloma, ki pomeni prehod mladostnika v moža, ko postane zaupanja vreden in se uveljavi v svoji skupnosti. S tem lahko odloča in sklepa tudi pomembnejše zaveze, ki denimo zadevajo družino in dom. Med paremije, ki opredeljujejo mladostniške in ljubezenske zaveze, uvrščamo:

97. *Fant dekletu vse dobro obeta, dedec z babo vsaki dan pometa.* (1987, 1; 29317)

98. *Fantje zvečer obetajo, zjutraj pozabijo.* (1987, 1; 29189)

99. *Fante obetajo dekletom zlate gradi, pa še črne kajže ni.* (1987, 1; 29172)

100. *Fantovska obljuba deklinška poguba.* (1987, 1; 29181)

101. *Fantovska obljuba, zimska odjuga, spomladni sneg.* (1987, 1; 29183)

102. *Muhe se love na med, dekleta pa na sladke besede.* (2000, 1; 35788)

103. *Prisege ljubezni imajo kratke noge.* (2001, 1; 36750)

Kot smo nakazali zgoraj, paremiološko gradivo izkazuje velike razlike v položaju moških in žensk tudi v odnosu do »družbene veljave« besede. Tradicijska prepričanja so negativno vrednotila ženske zaveze, medtem ko so pri moškem spolu negativno vrednotene le zaveze spolno nezrelih posameznikov v obdobju adolescence in so navadno poimenovane kot *fantovske obljube* (gl. 95–101). Da je ženitev predstavljena kot pogodba s hudičem, smo našli le en primer:

104. *Kdor se oženi, se hudiču zapiše.* (1987, 1; 29487)

Prvi pokazatelj razlike v družbeni veljavi besede glede na spol izjavljalca v družbi je kvantiteta, kar se sklada s podobo sklepanja zvez, kot je izražena v paremijah; če so besedna dejanja moškega maloštevilna, so pri ženskah znatno številnejša, poleg

tega pa so razlike v določenih primerih preprosto poenostavljene na raven bioloških znakov (105):

105. *Kjer ni moškega glasa, niti ni mož besede.* (1934–1974, 2; 08805)
106. *Mož beseda, žena besednjak.* (1963, 1; 03448)
107. *Mož beseda – žena besednjak.* (1983, 1; 24243)

Genezo poštenega moža gre po paremijah iskati v zavezi, ki traja do smrti in jo mož potrjuje z vsakim dnem zvestobe ženi – zakonski stan torej nekako posredno daje veljavo in moč drugim besednim zavezam moškega kot besedodajalca:

108. *Povezan mož, krepak junak, razvezan mož – slabič za nič.* (1947, 1; 01817)
109. *Da se mož rodi z dnem, ko se poroči.* (2001, 1; 36836)

Sestavine, povezane s poroko imajo med obravnavanimi paremijami posebno mesto, saj poročni obred v svoji ritualni zgradbi vsebuje sklenitev besedne zaveze. V spodnji metafori je razkorak med pričakovanim (zaobljubljenim) stanjem pred poroko in dejanskim stanjem po poroki ponazorjen z metaforo vida:

110. *Pri poroki roke vežejo, po njej se pa oči odvežejo.* (1934–2000, 3; 08630)
111. *Pri poroki roke zvežejo, po poroki se oči odvežejo.* (1987, 1; 27540)

Metafora vezi se kot temeljna metafora potrjuje tudi v primeru zakonske zveze:

112. *Sákon je jekléno vesílo.* (1833–1999, 2; 00539)

Sklenjena zaveza pa se kaže kot enaka za oba udeleženca, ne glede na to, kdo sta:

113. *Staremu možu se priseže pred altarjem ravno taka ljubezen kakor mlademu.* (1987, 1; 30244)

Meščanstvo in plemstvo je v gradivu zbirke na več mestih prikazano kot izrazito negativno zlasti v odnosu do besednih zavez z nižjimi družbenimi sloji, npr. s kmečkim prebivalstvom. Zanje so značilne vse negativno vrednotene lastnosti, ki smo jih našli v paremijah o zavezah na drugi stopnji: gospoda ni usmiljena in trdno drži za besedo, ne daje »moških« besed, torej se danih obljub ne drži, v svojih obljubah pogosto pretirava in na koncu ne preide na dejanja, kar pomeni, da ostane le pri besedah. Kot kontrast se v pregovoru (116) kaže kmečki stan, ki pa drži svojo besedo in je v pregovornem gradivu predstavljen kot pravična protiutež krivičnosti višjih slojev.

114. *Vifoki gopjudi, veliki obeti.* (1789, 1; 00043)
115. *Gosposko je obetat, kmętovsko dershat.* (18. stol., 1; 04163)
116. *Tudi beseda gosposka ni vselej moška.* (1852–1859, 1; 19939)
117. *Obljubiti in storiti ni gosposki.* (1852–1859, 1; 20194)
118. *Gospoda streže na besedo ko prasica na govno.* (1934–1974, 2; 07986)
119. *Gospoda streže na besedo ko pes na kost.* (1974, 2; 07987)

4 Nekaj sklepnih misli

Odgovorili smo na uvodoma postavljeno vprašanje in prišli do ugotovitve, da nam v razpravi predstavljeni pristop lahko ponudi nove poglede na razmeroma znano gradivo.

Paremiološko gradivo se izkazuje za nenadomestljivo, saj omogoča prepoznavo in vsaj delno interpretacijo številnih zunajjezikovnih kontekstualnih dejavnikov, ki so v preteklosti spremljali rabo besednih zvez. Do teh dejavnikov ne moremo dostopati drugače kot s semantično analizo sestavin v zgradbi paremij. Ostaja vprašanje, ali lahko na klasificirano tematsko skupino paremij, ki govori o določenem folklornem žanru, gledamo kot na zbir žanrskih napotkov, smernic in priporočil govorniku, torej kot na neke vrste »žanrski priročnik«, s katerim lahko govorec enkrat kot besedodajalec in drugič kot besedodajalec z »ustrezno« rabo žanra krmari med folklornosporazumevalnimi praksami skupnosti. Na vprašanje lahko z zadostno razčlenjenimi in gradivsko podprtimi kategorijami paremiološke klasifikacije odgovorimo pritrdilno, vendar dodajamo, da je treba rezultate tovrstnih analiz upoštevati le v okviru obširnejših raziskav, ki se ne opirajo izključno na en tip gradiva in lahko zaradi tega ponudijo celovitejšo predstavo o rabi žanra, omejeno na specifične lokalne skupnosti. Pomembno ostaja neodgovorjeno vprašanje, zakaj so se v zgradbah paremij ustalili ravno izpostavljeni kontekstualni dejavniki in v kakšnem odnosu so do ostalih kontekstualnih dejavnikov konkretnih govornih uresničitev posamičnih paremij.

Zahvala

Članek je nastal v okviru raziskovalnega programa *Etnološke, antropološke in folkloristične raziskave vsakdanjika* (P6-0088) ter projekta *Tradicionalne paremiološke enote v dialogu s sodobno rabo* (J6-2579).

Viri in literatura

- Babič, Saša, 2015: *Beseda ni konj. Estetska struktura slovenskih folklornih obrazcev*. Ljubljana: Založba ZRC, ZRC SAZU (Ethnologica – Dissertationes, 6). DOI: <https://doi.org/10.3986/9789612547660>.
- Babič, Saša, 2022: Raziskovanje paremioloških enot in nastajanje znanstvene zbirke pregovorov v okviru Inštituta za slovensko narodopisje ZRC SAZU. *Slavistična revija* 70/3, 289–298. DOI: <https://doi.org/10.57589/srl.v70i3.4019>.
- Babič, Saša; Voolaid, Piret, 2023: A Sociocultural View of Estonian and Slovenian Proverbs on Alcohol and Drinking. *Folklore* 134/3, 370–394. DOI: <https://doi.org/10.1080/0015587X.2023.2175972>.
- Bauman, Richard, 1984: *Verbal art as performance*. Long Grove: Waveland Press.
- Bauman, Richard; Briggs, Charles L., 1990: Poetics and Performance as Critical Perspectives on Language and Social Life. *Annual Review of Anthropology* 19, 59–88.
- Ben-Amos, Dan, 1993: "Context" in Context. *Western Folklore* 52/2–4, 209–226.
- Dundes, Alan, 1980: *Interpreting Folklore*. Bloomington: Indiana University Press.
- Grimes, Ronald L., 2010: *Beginnings in Ritual Studies*. Waterloo: Ritual Studies International.
- Grzybek, Peter, 1994: Foundations of Semiotic Proverb Study. V: Mieder, Wolfgang (ur.), *Wise Words: Essays on the Proverb*. New York, London: Garland Publishing, 31–71.

- Grzybek, Peter, 2014: Semiotic and Semantic Aspects of the Proverb. V: Hrisztova-Gotthardt, Hrisztalina; Varga, Melita Aleksa (ur.), *Introduction to Paremiology: A Comprehensive Guide to Proverb Studies*. Warsaw, Berlin: de Gruyter, 68–111.
- Huzjan, Vanja, 2024: Ethnological Analysis of Lexeme 'otrok' (Child) in the Collection of Slovenian Paremiological Units on Clarin.si. V: Babič, Saša; Jakop, Nataša; Mrvič, Rok (ur.), *Paremiologija med tradicijo in sodobnostjo/Paremiology between tradition and innovation*, Ljubljana: Založba ZRC, ZRC SAZU, 151–179. DOI: https://doi.org/10.3986/9789610508861_10.
- Ivančič Kutin, Barbara, 2011: *Živa pripoved v zapisu. Kontekst, tekstura in prekodiranje pripovedi Tine Kravanja iz Bavšice*. Ljubljana: Založba ZRC, ZRC SAZU (Ethnologica – Dissertationes, 3). DOI: <https://doi.org/10.3986/9789612543518>.
- Jakobson, Roman, 1989: *Lingvistični in drugi spisi*. Ljubljana: ISH (Studia humanitatis).
- Jakobson, Roman O.; Bogatyrev, Petr G., 1971 [1929]: Fol'klor kak osobaja forma tvorčestva [Die Folklore als eine besondere Form des Schaffens]. V: Bogatyrev, Petr G. (ur.), *Voprosy teorii narodnogo tvorčestva*, 369–383. Moskva: Iskusstvo.
- Karasik, Vladimir I., 2016: Komissiv kak postupok: lingvokul'turnye karakteristiki. *Žanry reči* 14/2, 56–66. DOI: <https://doi.org/10.18500/2311-0740-2016-2-14-56-66>.
- Kirshenblatt-Gimblett, Barbara, 1973: Toward a Theory of Proverb Meaning. *Proverbium* 22, 821–827.
- Krikmann, Arvo, 1984 [1974]: On Denotative Indefiniteness of Proverbs. *Proverbium* 1, 47–91.
- Kržišnik, Erika, 2005: Frazologija v luči kulture. V: Stabej, Marko (ur.), *Večkulturnost v slovenskem jeziku, literaturi in kulturi*. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete Univerze v Ljubljani, 67–81.
- Levkievskaja, Elena E., 2004: Kukiš. V: Tol'stoj, Nikita I. (ur.), *Slavjanske drevnosti III*. Moskva: Meždunarodnye otnošenija, 26–27.
- Meterc, Matej, 2017: *Paremiološki optimum. Najbolj poznani in pogosti pregovori ter sorodne paremije v slovenščini*. Ljubljana: Založba ZRC, ZRC SAZU (Lingua Slovenica, 10). DOI: <https://doi.org/10.3986/9789610504153>.
- Meterc, Matej, 2020–: *Slovar pregovorov in sorodnih paremioloških izrazov*. Spletna objava na portalu Fran.si. Na spletu: <https://fran.si/216> (dostop 12. 2. 2024).
- Rothenbuhler, Eric W., 2016 [1998]: *Ritualno komuniciranje: od vsakdanje konverzacije do medijsko posredovane ceremonije*. Ljubljana: Fakulteta za družbene vede in ZRC SAZU (Theoria).
- Stanonik, Marija, 2001: *Teoretični oris slovstvene folklore*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9616358316>.
- Stanonik, Marija, 2015: Slovenski pregovori kot kulturna dediščina. *Traditiones* 44/3, 171–214. DOI: <https://doi.org/10.3986/Traditio2015440309>.
- Telija, Veronika N., 1996: *Russkaja frazeologija. Semantičeskij, pragmatičeskij i lingvokul'turološki aspekti*. Moskva: Škola Jazyki russoj kul'tury.
- Turner, Victor, 1991 [1969]: *The Ritual Process: Structure and Anti-Structure*. Ithaca (NY): Cornell University Press.
- van Gennep, Arnold, 1960 [1909]: *The Rites of Passage [Les rites de passage]*. Chicago: University of Chicago Press.
- Wolfgang Mieder, 2004: *Proverbs: A Handbook*. Westport (CT), London: Greenwood Folklore Handbooks.



III

Etnolingvistična analiza
paremioloških enot
Ethnolinguistic analysis of
paremiological units



The naked truth for a good ending: The end in Slovenian paremiological units

Gola resnica za dober konec: konec v slovenskih paremioloških enotah

Konec predstavlja v naši kulturi eno od meja, je ena od usodnih meja. Konec je prelomnica in eden najbolj doslednih in močnih pojmov. Jezik je eden najpomembnejših nosilcev tega pojma, predvsem zato, ker je neoprijemljiv in povezan z izkušnjami in načinom življenja ljudi. Jezik, zlasti v kratkih folklorih oblikah, ohranja koncept in stereotipe, povezane s koncem. Z uporabo etnolingvističnega pristopa in semiotike lahko prodremo v globlje strukture in pomene besedil ter v njih celo zasledimo družbene stereotipe. V prispevku so predstavljene paremije iz zbirke Pregovori v repozitoriju Clarin.si, ki vsebujejo besedo 'konec', njihove pomenske plasti in semiotični vidiki obravnavanih pregovorov kot celote.

• **Ključne besede:** konec, pregovor, stereotip, paremiologija, folkloristika, mitologija

In our culture, the end represents one of the boundaries, it is one of the fatal boundaries. The end is a breaking point and one of the most consistent and powerful concepts. Language is one of the most important media for this concept, mainly because of its intangible nature and connectedness with experience and ways of life. Language is good at preserving the concept and stereotypes, related to the end, especially in short folklore forms. Ethnolinguistic approach and semiotics help us get a detailed look of the inmost structures and meanings of concepts and social stereotypes. The article presents paremies from the Proverbs section of the Clarin.si repository, which include the word 'konec' (end), their semantic layers and semiotic aspects of discussed proverbs as a whole.

• **Keywords:** end, proverb, stereotype, paremiology, folkloristics, mythology

Today, time is regarded as a linear chain of events, a line with a starting point somewhere in the distant past, from where it continues linearly towards its end. This linear temporal concept appears to act as a contemporary mask, covering a more profound and more lasting understanding of the cyclic nature of time, which can be detected in different folklore genres. There is a connection between our understanding of time and our understanding of the end. The end is one of the key modalities of relations to just about everything: the world, others, ourselves as a part of this world (Zupančič 2019: 7). The end as a concept is all around us, on every level of our lives: in the yearly cycle of activities, in spatial and temporal perceptions, in discontinuation of activities and starting new ones ... It relates to special and temporal perceptions, and predominantly marks the outer edges of different phenomena or boundaries.

The end is a frequent part of folklore literature, and represents an important compound of myths of eschatology in mythology. Myths of eschatology deal with the end of the world, the cosmos and, as a matter of fact, the destruction of the human race,

with the exception of a few individuals who later became the progenitors of the occurrence of a new civilisation (Eliade 1963: 54). The mythical end is a flooded earth or destruction by fire, an occurrence which pushed the Earth back into chaos, which in turn again followed by cosmogony. Myths often make it difficult to discern, whether they are portraying past destruction, ending of the world, or future ones (Eliade 1963: 56), which further endorses the cyclic concept. In our society, people's proximity to the cosmic context of nature is expressed in rituals related to yearly and life cycles (despite the fact that the meaning of many has been lost), and the experience has been embedded in a variety of genres and types of folklore literature, especially in myths and fairy tales, but it can also be found in tales, proverbs, riddles, incantations ... (Šmitek 2001: 33). This could also be applied to calendars, as they divide time into days, weeks, months, while naming days after saints divides our time into shorter and longer cycles, which eventually end and are succeeded by another cycle upon (day, week, month, year ...). Marking the beginning and the end mainly deals with the world order and consequently with time and space as fundamental anthropological constants. Myths of eschatology indicate that the end (of all existence) is inevitable (Šmitek 2001: 37), and this mind-set is ingrained in the very essence and existence of human beings, while the image of the end in a way mirrors the image of the beginning of the world: both the beginning and the end are characterised by chaos, which once again restores life, rather than creating it anew (Dux 1992: 228).

The end is a frequent subject of folklore literature: the end is explicitly exposed even in fairy tale formulas *In živela sta srečno do konca svojih dni / živela sta srečno do konca svojih dni* (And they lived happily till the end of their days / they lived happily till the end of their days) or *In počil je lonec, pravljice je konec* (And the pot cracked and the fairy tale ended), which are used to close the fairy tale and implicate a possible continuation of the story, which the fairy tale doesn't provide (e.g. prince's and princess's married life). Some animal tales approach the end of the world by including mistaken animals, fleeing because *the world is about to fall apart* (ATU 20C: *Animals flee in fear of the end of the world*). While the first case has to do with time, the second relates to time, as well as spatial and action concepts.

The end in riddles is conceptually connected with mythology: the end is signified as something representative of (our) world (*Bog nima, svet ima, hudič bi rad in ne dobi? – Konec* (God doesn't have it, the world has it, the Devil wants it but doesn't get it? – The end)) and as the shortest thing (*Kaj je najkrajše? – Konec* (What is the shortest? – The end)). The end as part of the riddle's question is used in idiomatic form "vzeti konec" (to meet one's end), as in 'cease to exist'. (*Mrzel sem, vendar grejem, kar bi sicer konec vzelo. – Sneg* (I'm cold, yet I keep warm what would otherwise meet its end)).¹

¹ Online: <https://www.folklore.ee/Slovenianriddles/tyybid> (accessed 11.12.2023).

These examples from different genres of folklore literature point to the extent of perceiving the end: it marks both restoration and cessation of existence, the end depends on temporal and special aspects, processes and activities. Paremiological material partly differs from the aforementioned statements, as this folklore material remains part of everyday speech. For this reason, approaching the end relates primarily to the perception of time, space, activities and processes. This discussion will focus on paremiological material, with the use of semiotic approach; individual proverbs will first be defined with the help of semiotics and then analysed as to what is signified by the sign ‘konec’ (the end), found in the paremiological material.

1 Semiotic approach to paremiological material

Proverbs are short folklore forms, characterised and defined by rhyme, metaphor and a concise message (Matičeto 1956; Stanonik 2001; Babič 2015); they maintain the authority of our ancestors’ wisdom in society (Mieder 2004), several generations which relied on observation and experience in their attempt to arrange the world and systematise it in pre-prepared statements. For this reason, proverbs are considered a convenient rhetorical means which convey social rules and values in a more elegant way.

Proverbs are usually defined as relatively stable forms, which preserve words and meanings. When speaking about proverbs and semiotics, dyadic relations therefore need to be considered (Morris 1938: 6): pragmatic, syntactic and semantic dimensions. Pragmatics deals with relation between sign and users of signs, syntactics is directed at formal relations among signs, and semantics focuses on signs’ relations to objects (Grzybek 2014: 69).

Morris and later Grzybek (2014) both stressed high interconnectedness of these three dimensions, indicating that they can only be separated and studied with separate emphasis for heuristic purposes, but nevertheless not really isolated, regardless of the use of signs or research of use. Semiotic approach predicts a sign consisting of a signifier, a signified and an interpreter. If the word “konec” (the end) is the signifier in our case, what is the signified?

The end is - obviously - an abstract concept, to the degree that, when standing alone, without specification, “the end of something” is described and defined only in (language) dictionaries. In any case, the expression “the end” serves as a sign denoting a pending (final) boundary in a temporal and spatial unit and points to the perception of these two anthropological constants.

On the semiotic level, Permjakov also emphasizes the concept of situation or context, which give rise to and provide a platform for folklore texts: according to Permjakov’s interpretation, the sign alone solves a situation, yet doesn’t depend on it; his chosen example are proverbs and proverbial phrases as “signs of situations or a certain type

of relation between objects” (Permjakov 1979: 20). The figurative meaning of the so-called “cliché text”, such as a proverb, for instance, namely enables folklore texts to be used also as rules for recognising novel and previously unfamiliar situations (Permjakov 1979: 272). From a semiotic perspective, this makes a proverb, as verbal constructions, a sign: the sentence carries a (metaphorical) meaning as a whole; as such, a proverb can also be defined as a so-called super sign, i.e. a complex sign or a sign complex (Grzybek 2014: 85). The meaning is created by context in which it is used, yet some words nevertheless retain their primary meaning and define temporal and spatial relations within a sign complex or super sign. These words allow us to discern the worldview and attitudes to basic social levels and relations. One of such signs or words is ‘konec’ (the end), as it retains its meaning and at the same time directs the meaning of a context from temporal, spacial and processing perspectives.

1.1 Development and meaning of the word ‘konec’ (end)

End is signified by two words in the Slovenian language: ‘konec’ (end) and ‘kraj’ (place/end) - the first is nowadays more frequently utilised in everyday use, the latter is signified as obsolete in this sense, while it is used in standard language for defining location or geographical settlement unit.² The end signifies a concept usually linked to a specific phenomenon, and, as defined in the Dictionary of the Slovenian Standard Language, defines 1) a part, area, spatially most distant from the starting point, beginning, centre (*biti na koncu sveta / iti na konec sveta*) (to be at the end of the world / to go to the end of the world), 2) a larger or a smaller part of an area, space (*zame so to španski kraji*) (these are Spanish places for me), 3) a piece, a part of something (*začeti na pravem koncu, potegniti za krajši konec*) (to start at the right end, to pull the short end), 4) something most distant from the starting point, regardless of time, action, existence (*niti na konec / kraj pameti mi ne pade*) (it doesn’t even fall at the end / place of my mind).³

The etymological development of the Slovenian word ‘konec’ (end) is supposedly rooted in Proto-Slavic *konьsь ‘end, beginning’ (< kōn ‘end, beginning’ and Proto-Slavic *čęti ‘to begin’), which semantically encompasses the concept “where one thing (a field, a stick, a rope) begins, it also ends” - if one looks from the opposite side (ES 2015, 27.12.2023). On the other hand, the Slovenian word ‘kraj’ (place / end) developed from Proto-Slavic *krājь, which is related to *krojiti ‘to cut, to tailor’ and originally meant ‘that which is cut off, where it was cut off’. The meaning first concretised as ‘end, edge’, from there moved to ‘border’ and ‘border area’ and later developed into ‘bordered area, land’ and into (inhabited) place’ (ESS 2015, 27.12.2023). ‘Kraj’ (place / end) in formal language therefore primarily marks a settlement. (E.g. *V domačem kraju so odprli novo pošto*. (A new post office has opened in the home place.)), while

² The word ‘konec’ (end) in this discussion is used to represent both ‘konec’ (end) and ‘kraj’ (place / end).

³ Online: <https://fran.si/iskanje?View=1&Query=konec> (27.12.2023)

the word ‘konec’ (end) can mark a location in colloquial language. (E.g. *V katerem koncu živiš?* (Which end do you live in?)). The connectedness of both terms to both beginning and end is clearly demonstrated by the idiom *brez konca in kraja* (with no end or place), which marks that which never ends or the absence of borders.

A similar conceptual development of an expression for the end can be traced in several other Germanic languages, where the expression ‘end/das Ende’ supposedly came from the Proto-Germanic *andiaz, which primarily signified “the opposite side”, and in turn developed from the Indo-European *antjo (end, border) with the lexical root *ant- “in front, forehead” and derivative meanings “in front, before”.⁴ The expression for end in Roman languages comes from the Latin base of *fine* in the sense of “a distinguishing point, something changes after the end; what ends, ceases to exist; a border in time and space”.⁵

It is evidently clear that the meaning of the sign ‘konec’ (end) relates to most distant parts of finishing and beginning, i.e. borders, spatial, temporal and processing, that end used to mark both beginning and end, where it could signify another beginning, which points to a cyclical perception of the course.

Contemporary predominant meaning of the words ‘konec/kraj’ (end / place, end) depends on the comprehension of conclusion: [biti] konec česa/koga (to be the end of something/someone) relates to reaching the outermost borders of space, time, process, which we are about to leave, conclude, noting that the outermost borders are not always defined and can mean any point, which is indicated by the idiom na vseh koncih in krajih (in all ends and places), meaning “everywhere”.

1.2 Methodology

To analyse the conceptual aspect of the sign ‘konec’ (end), I took a closer look at paremiological units in the scientific Collection of Slovenian paremiological units Pregovori 1.0, available at the repository of Clarin.si.⁶ Using the NoSketch concordancer I searched for paremiological units containing expressions ‘konec’ or ‘kraj’, and excluding proverbs using the words ‘konec’ or ‘kraj’ as settlement signifiers. This left 239 paremiological units suitable for analysis. The material was grouped by meaning, and represented in this article by a selection of variants (the complete number is significantly higher than the present selection): I tried to read the context and determine which nuance of the word ‘end’ was realized. Conceptual categorisation was then used in an attempt to identify semantic compounds which the word ‘konec’ could express, while I was looking for connections with myths of eschatology. Semiotics was used to detect both nuance in the meaning of the word, as well as the meaning of the supersign that is the formula: sign = signifier + signified (i.e. the word konec/kraj).

⁴ Online: <https://www.etymonline.com/search?q=end> (27.12.2023).

⁵ Online: <http://www.etimo.it/?term=fine&find=Cerca> (27.12.2023).

⁶ Online: <http://hdl.handle.net/11356/1455> (27.12.2023).

2 Analysis of paremiological units containing the word 'end'

Material comprised of 239 paremiological units is divided into six groups: five conceptual groups belonging to proverbs, and an additional group containing weather paremies - they represent a separate and autonomous group which falls in the intersection of experience, prediction and superstition. While proverbs express a society's world view, instructions for actions and morality of several generations (Mieder 2004), weather proverbs predominantly express expectations of future weather and harvest, and are highly cyclical and full of eternal return of time (Eliade 1963). Both types of paremies are generally classified as proverbs due to their structural characteristics, i.e. rhyme, rhythm and sentence structure, and partially also due to metaphorical expression (in case of weather proverbs mostly in the form of personification), while another common characteristic should not be overlooked, namely that all paremies include "several generations' experience and observations" and are passed down from one generation to another in the form of ancestral wisdom (despite the fact that many of them are questionable when it comes to the aspect of traditionality) (Mieder 1996, 2004). This is what endows them with authority and social importance. Expressing a paremiological unit might indeed be using a pre-prepared statement or reaction, but at the same time the statement makes the expressed thought more significant.

All units where "konec/kraj" signified a geographical unit, a settlement, were removed from the selected material, despite their conceptual underpinning of "where I live is the end/edge of "our" world"; this concept is further confirmed in e.g. incantations, as well as in mythical schemes, where forest, ocean or anything beyond the village border represent another, wild, not-our world. By using the sign kraj/konec as geographical units, the primary conceptual base described here has been lost to the extent that it makes them irrelevant for this discussion. The formation of conceptual groups containing selected paremiological units was governed by the meaning which the sign in question took on in an individual sign complex, i.e. paremiological unit.

I. Beginning

Beginning is more often signified by the word 'kraj' than the word 'konec', the latter is used in this sense in only one case. The conceptualisation of beginning makes it clear that it is a point at the border, simply a place where a process begins.

1. *Od kraja moli boga in od kraja kruh jej.* (Prey from the beginning and eat bread from the beginning.)
2. *Smrt kosi od kraja. / konca.* (Death mows from the beginning.)
3. *Smrt je slepa, od kraja pobira, nič ne izbira.* (Death is blind, it takes from the beginning, it does not choose.)
4. *Kdor kosi od kraja, psa pusti, da laja.* (The one who mows from the beginning, leaves the dog barking.)

II. Edge

Proverbs where the sign 'konec' / 'kraj' means a border area, running along the edge. The edge is an area, not a single point. While the edge can mean distance, invisibility (6.), it can also mean greater exposure to sun, in this case positive one (8.), the absence of the phenomenon at borders can mean total absence (5).

5. *Strah je v sredi votel, ob krajih ga pa nič ni.* (The fear is hollow inside, at the ends there is none of it.)
6. *Tukaj na kraju, tam na sredi v raju.* (Here at the end, in paradise in the middle.)
7. *Konec jezika se rado opotika.* (The end of the tongue likes to knot.)
8. *Lepe črešnje na koncu vej raso.* (Good cherries grow at the end of the branch.)
9. *Ako pri kraju ne vidiš dna, ne bređi čez vodo!* (If you don't see the bottom at the edge, don't wade across the water!)

III. Side

A special category emerged, wherein 'konec' or 'kraj' mark a side, with the focus mainly on the two-sidedness of an object (stick, knife ...). There is a clear perception of binary in this group: everything has two sides, these two sides can be cut into, while a side can also denote spare next to something (12.):

10. *Palica ima dva konca / kraja.* (The stick has two ends.)
11. *Je nož, ki na dva kraja reže.* (The knife cuts two ways.)
12. *Kraj suhega drevesa tudi sirovo zgori.* (By the dry tree the fresh one burns also.)

IV. Conclusion

The most common conceptual interpretation of the word 'konec' in today's everyday life is conclusion, so it comes as no surprise that the majority of paremiological units were conceptually related to this meaning (conclusion of a process, of time). In these cases, the end marks a final point, a moment when an outermost limit is reached.

13. *Konec dober, vse dobro.* (All's well that ends well.)
14. *Vsak doseže svoj konec, a cilja ne.* (Everyone reaches their end, but not the goal.)
15. *Vsaka reč ima svoj konec.* (Everything has its end.)
16. *Kjer začno ženske, zlepa ni konca.* (Where women begin, there is no end.)
17. *Vsaka pesem ima svoj konec.* (Every song has its end.)
18. *Začetek ljubezni je sladak, konec pa grenak.* (The beginning of love is sweet; the end is sour.)
19. *Ako eno drevo pade, ni gozda kraj.* (If one tree falls, it is not the end of the forest.)

The paremiological unit *Vsaka reč ima svoj konec, klobasa ima dva* (Everything has its end, the sausage has two) points to a close conceptual connection within meanings expressed by the word 'konec' / 'kraj'. Conceptually, it always denotes a marginal, border area which marks the borders of existence, and at the same time calls attention to either multiple ends or the relativity of a point of view, while stressing the binary concept of two sides.

V. Death

Death represents the end of life (or the end of the known). Since death is a fairly taboo phenomenon, it is often replaced with other, euphemistic expressions. One of them is “end”, which, when used for human beings, marks death.

20. *Zlobni človek ne vzame kon'c.* (Evil person does not take the end.)
21. *Kdor ne bo lačnemu kruha podal, lahko bo od žeje še konec jemal.* (The one that does not give the bread to the poor, might take the end from thirst.)
22. *Pijanec vzame slab konec.* (A drunk takes a bad end.)
23. *Dolgi boleznj konec motika.* (After long disease the end is slipping.)

In this sense “end” refers mainly to a bad ending and such proverbs are used as admonitions (21, 22) or warnings of (seemingly) eternal persistence of evil (20, 23). In no case does the ‘end’ refer to a positive outcome.

VI. Weather paremies

The analysed material contained approximately 21% of weather paremies. In most cases, they referred to an end of a specific period: year, season, month. This exposes a connection of end with the temporal perspective.

Anticipation of a season’s end is most obvious in the case of winter: the end of winter means the arrival of a less difficult period, the promise of harvest, warmth and more light. Weather paremies either warn that the winter will not be over soon:

24. *Če na svečnico (2. 2.) od strehe kane prej ko od sveče, ne bo še kmalu konec zime.* (If on Candlemas it drips from the roof sooner than it drips from the candle, the winter will not end anytime soon.)
25. *Če sv. Petra stol (22. 2.) mrazi, še dolgo cajta zime konca ni.* (If on St Peter’s day the chair is cold, for a long time there is no end to the winter.)
26. *Če je mračno in mrzlo, gre medved nazaj v brlog, ker zime še ne bo konec.* (If it is dark and cold, the bear goes back to the den, because winter is not over yet.)

or ensure that the winter is over and a new abundance is coming:

27. *Sv. Matija led razbije (24. 2.) – konec zime.* (St Matthias breaks the ice – the end of winter.)
28. *Ljubi maj, krasni maj, konec zime je tedaj.* (Beloved May, wonderful May, the end of winter is then.)

Individual Saints predict time which should last throughout a season:

29. *Vreme, ki ga Lovrenc (10. 8.) naredi, do konca jeseni drži.* (Weather on St Lawrence lasts till the end of autumn.)

or a month:

30. *Če se Servac (13. 5.) kislo drži, do konca meseca ni sončnih dni.* (If Servac has a sour face, there will be no sunny days till the end of the month.)
31. *Če se sv. Medard (8. 6.) kislo drži, do konca mesca ni pravih sončnih dni.* (If St Medardus is sour, then no real sunny days will come till the end of the month.)

Future weather is also determined by the specific time period (on a certain day or month):

32. *Maja mora biti mrzlo, če ni na začetku, je ob koncu tako.* (May should be cold, if not at the beginning, it will be so at the end.)
33. *Če je na drugi dan julija lepo, tako do konca meseca bo.* (If the second day of July is nice, it will stay like that till the end of the month.)

The end of the month can also predict harvest, which is key for prosperity. Predictive are the ends of February, October and December:

34. *Če konec svečana sever brije, dobre letine nam zasije.* (If at the end of February the north wind blows, good harvest shines for us.)
- Konec vinotoka dež – rodovitno leto.* (Rain at the end of October - a fruitful year.)
35. *Če konec leta burja hruje, slabo letino napoveduje.* (If at the end of the year northern wind roars, a bad harvest is forecasted.)

St Andrew and St Sylvester are the ones predicting weather till the end of the year:

36. *Na Andreja (30. 11.), ko se v kozarcih vino smeja, vreme tako ali onako ostane do konca leta enako.* (On St Andrew, when wine smiles in glasses, the weather will remain the same till the end of the year, one way or another.)
37. *Kakor vreme na Silvestra (31. 12.) kane, tako potem do konca leta ostane.* (As the weather is on St Silverster, so it will continue till the end of the year.)

The end also plays a role on both summer and winter solstice. In both cases, the solstice relates to a period of several days rather than to a single day. The weather paremy referring to the summer solstice focuses on the spacial aspect, and the winter on the temporal aspect of the duration of darkness:

38. *O sv. Vid (15. 6.) se na vsakem konc dan vid'.* (On St Vitus' day you can see the day at every end.)
39. *Od 21. decembra do konca decembra se utegne dan, kolikor petelin skoči čez prag.* (From 21 December to the end of December, the day stretches as far as a rooster jumps over a threshold.)

Weather paremies clearly depict the anticipation of an end, which can bring relief, as they bring something new, better (27, 28).

When talking about weather paremies, 'konec' relates to a conclusion of a certain time period, a cycle, which is followed by a new period, a new cycle. People were predicting weather conditions and harvest in the coming period based on phenomena in the expiring season. To a certain extent, weather paremias function as instructions (akin to true proverbs) concerning future actions, according to weather forecast. Namely, weather paremies attempt to systematise natural phenomena, create rules governing the 'cosmisated' world, as our world is governed by rules, while the other, not-our world, the one that follows the end, chaos and lawlessness prevail. "Za koncem" (beyond the end) even weather paremies lose their power.

3 Conclusion

The end is an all-pervasive concept. It is a part of everyday and ceremonial life, heavily present in both mythology and daily conversation. In modern times, 'konec' predominantly marks a conclusion of an action, an irreversible state. This is apparent in the units classified in the contemporary paremiological optimum: *To ni konec sveta / Vsega (lepega) je enkrat konec* (This is not the end of the world / Everything (nice) ends sometime) (Meterc 2017), with the emphasis on the aspect of passing and apparent irreversibility. A lesser known fact is that 'end' in Slovenian folklore and mythology also relates to a cyclical pattern: a new birth as well as cyclical understanding of time. This conceptual foundation is apparent in the very etymology of the word 'konec' and 'kraj', and even in individual paremiological units, where the two words explicitly mark a beginning. The mythological aspect is reflected in Slovenian riddles, as 'konec' partly relates to chaos, that which is not present in our "cosmisated" world. However, the mythological chaos is the earliest cause of something new - cosmos can only arise out of chaos and only chaos can give rise to something new. In modern times, cyclical perception of time along with a new continuation after an end, and active realisation of the proverb *Vsak konec je nov začetek* (Every end is a new beginning) has retreated into the background and left the stage to mere linear perception and irrevocable conclusion (this proverb is very often used as consolation).

The origin of the word 'konec' in various Indo-European language groups indicates that the concept of an end relates to borders, to a place of "fracture" and the start of another, new. This is corroborated by proverbs with discernible semantic groups beginning, edge, side, conclusion, death, that which can be overhanging. The sign "konec" thus includes signifiers 'konec' and 'kraj', and the signified relates to perception of extreme parts of conclusion and beginning, i.e. to borders, may they be spatial, temporal or process-related; the end is thus both the beginning and the conclusion, where it can mark a new beginning. The end is a border between the known, from spacial, temporal and processing aspects alike. So it doesn't come as a surprise that "konec" also serves as a euphemism for death, and in this case in no way relates to the concept of good, on the contrary, used in proverbs it always serves as a warning of a difficult death or an evasion of a bad ending (or a person or a physical condition). When marking the border in the sense of edge or side, the end is perceived as a wider strip or area (border, peripheral area), while in the case of the end of the process and often also of time, such as death, it denotes a point.

When talking about weather proverbs, temporal marking of conclusions of certain periods stands out the most. Judging by exposure, the conclusion of winter is the most anticipated. Namely, there are no proverbs articulating the end of spring or the end of summer, as they are periods of abundance and growth; proverbs discuss the end of autumn, which at the same time marks the beginning of winter, and the end of winter,

which at the same time means the beginning of spring; based on other narrative material one can conclude that this is based on the mythological perception of “the end of chaos and dangerous times, the beginning of cosmos and prosperity”. In any case, weather paremies are possibly signified by the perception of time as cyclical: beginning with the fact that weather courses are repeated every year or that individual days foretell a specific weather course in the near future, are strong indicators of the social perception of the cyclical nature of time. It is also obvious that ends mean conclusions of certain periods and simultaneously beginnings of new ones. Expectation of a conclusion is usually related to the duration of a less favourable period, whose end we are looking forward to: the end of winter or a long illness, as this will be succeeded by a new, better beginning. At the same time the weather proverbs use beginnings and endings to mark our “cosmised” world, which should function “according to rules”, in contrast to the not-our world, where chaos reigns and where there is no beginning and no end.

‘Konec’ also exposes our perception of binary: the existence of two sides as two extremes: a beginning and an end (e.g. 9, 18) or left and right (10), in the same way that they mark beginning and concluding, but not everything in between: just like in fairy tales featuring a prince marrying his chosen one, we read *In živela sta srečno do konca dni* (And they lived happily till the end of days) without learning anything about the couple’s married life. This binary or focus on two extremes is also evident in paremiological units which made it into the collection of the most familiar Slovenian paremies and M. Meterc classified in the so-called paremiological optimum: *Kakršen začetek takšen konec / Konec dober, vse dobro / Slab začetek, dober konec* (As it begins, so it ends / Good end, all good / Bad beginning, good ending) (Meterc 2017: 237–250).

Proverbs as super signs undoubtedly point to diachronic self-perception of society and the way it functions in time and space (chronotope), with specific expressions or signs within this complex super sign direct the meaning and point to relations among individual phenomena. ‘End’ as a concept is difficult to pinpoint, as it is usually related to actual phenomena and is not examined independently in discussions. But in each and every case it expresses its etymological base: conclusion and beginning, beginning and conclusion, i.e. the turning point which gives rise to the new.

Acknowledgements

The research for this paper was supported by the Slovenian Research and Innovation Agency (ARIS): *Traditional Paremiological Units in Dialogue with Contemporary Use* (J6-2579), *Language, culture and values: economic images of everyday life in folklore forms* (2023-2026, J6-50197), and *Ethnological, Anthropological and Folklore Studies Research on Everyday Life* (research core funding, P6-0088).

References

- Babič, Saša, 2015: *Beseda ni konj. Estetska struktura slovenskih folklornih obrazcev*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612547660>.
- Babič, Saša, 2018: *Slovenian Riddles*. Online: <https://www.folklore.ee/Slovenianriddles/index> (accessed 11.12.2023)
- Babič, Saša; et al., 2022: Collection of Slovenian paremiological units Pregovori 1.0, Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1455> (accessed 27.12.2023).
- Dux, Günter, 1992: *Die Zeit in der geschichte: Ihre Etwicklungsojik vom Mythos zur Weltzeit*. Frankfurt am Mein.
- Eliade, Mircea, 1963: *Myth and Reality*. New York: Harper & Row Publishers.
- Eliade, Mircea, 1992: *Kozmos in zgodovina. Mit o večnem vračanju*. Ljubljana: Nova revija.
- Lauhakangas, Outi, 2014: Categorization of Proverbs. In: Hrisztova-Gotthardt, Hrisztalina; Varga, Melita Aleksa (eds.), *Introduction to Paremiology: A Comprehensive Guide to Proverb Studies*. Warsaw, Berlin: De Gruyter Open Poland. DOI: <https://doi.org/10.2478/9783110410167>.
- Grzybek, Peter, 2014: Semiotic and Semantic Aspects of the Proverb. In: Hrisztova-Gotthardt, Hrisztalina; Varga, Melita Aleksa (eds.), *Introduction to Paremiology: A Comprehensive Guide to Proverb Studies*. Warsaw, Berlin: De Gruyter Open Poland. DOI: <https://doi.org/10.2478/9783110410167>.
- Kingsbury, Stewart A.; Kingsbury, Mildred E.; Mieder, Wolfgang, 1996: *Weather Wisdom. Proverbs, Superstitions and Signs*. New York, Washington D.C., Baltimore, Bern, Frankfurt am Main, Berlin, Vienna, Paris: Peter Lang.
- Matičeto, Milko, 1956: Pregovori in uganke. In: Legiša, Lino; Gspan, Alfonz; Slodnjak, Anton (eds.), *Zgodovina slovenskega slovstva I*. Ljubljana: Slovenska matica, 115–118.
- Meterc, Matej, 2017: *Paremiološki optimum. Najbolj poznani in pogosti pregovori ter sorodne paremije v slovenščini*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610504153>.
- Mieder, Wolfgang, 2004: *Proverbs. A Handbook*. Westport, Connecticut, London: Greenwood Press.
- Morris, Charles W., 1938: *Foundations of the Theory of Signs*. Chicago: The University of Chicago Press.
- Permyakov, Grigoriy, 1979 [1970]: *From Proverb to Folk-Tale. Notes on the General Theory of Cliche*. Moscow: Nauka.
- Snoj, Marko, 2015: *Slovenski etimološki slovar*. Online: <https://fran.si/193> (accessed 11.12.2023).
- Stanonik, Marija, 2001: *Teoretični oris slovstvene folklore*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9616358316>.
- Šmitek, Zmago, 2001: Time in Myth and Fairytale. In: Zmago, Šmitek; Brumen, Borut (eds.): *Zemljevidi časa/Maps of Time*. Ljubljana: Oddelek za etnologijo in kulturno antropologijo, Filozofska fakulteta Univerze v Ljubljani, 33–59.
- Zupančič Žerdin, Alenka, 2019: *Konec*. Ljubljana: Društvo za teoretsko psihoanalizo.
- eSSKJ, 2016–: *Slovar slovenskega knjižnega jezika*. Online: <https://fran.si/201> (accessed 11.12.2023).

Clothing in proverbs: Between symbol and meaning

Oblačila v pregovorih: med simbolom in pomenom

Short folklore forms are a genre of everyday speech: they capture material or content from everyday life, not only from the world of objects and existence, but also from values and understandings of the world. Clothes are an indispensable part of our everyday life. They are such an important part of our existence that they also have great economic significance, which is why the appearance of different items of clothing can also be traced in folklore forms. This paper presents an overview of clothing items and the crafts connected with making them in Slovenian proverbs and their connection with symbolic meanings. The specific cultural horizon of life in the past will be covered through analysis and interpretation, and will also reveal social and cultural stereotypes, images and concepts related to the economic image of society.

♦ **Keywords:** proverbs, clothing, crafts, everyday life, metaphors, symbols

Folklorni obrazci so zvrst vsakdanjega govora: zajemajo snov iz vsakdanjega življenja, ne le iz sveta predmetov in bivanja, temveč tudi iz vrednot in razumevanja sveta. Oblačila so nepogrešljiv del našega vsakdanjega življenja. So tako pomemben del našega bivanja, da imajo tudi velik gospodarski pomen, zato lahko pojavljanje različnih oblačil zasledimo tudi v folklornih žanrih. V prispevku je predstavljen pregled oblačil in z njihovo izdelavo povezanih obrti v slovenskih pregovorih ter njihova povezanost s simbolnimi pomeni. Z analizo in interpretacijo bo zajeto specifično kulturno obzorje življenja v preteklosti, razkriti pa bodo tudi družbeni in kulturni stereotipi, podobe in pojmi, povezani z gospodarsko podobo družbe.

♦ **Ključne besede:** pregovori, oblačila, obrt, vsakdanje življenje, metafore, simboli

1 Introduction

Language can serve as a lens through which we can explore the culture and ideology of a nation. Disciplines that study language, such as ethnolinguistics and paremiology, offer valuable insights into this complex subject. It is indisputable that proverbs and sayings serve as precise reflections of a nation's way of life, traditions and other distinctive characteristics. Embedded within a language's memory, they encapsulate the everyday realities that a nation has encountered over centuries. Not just proverbs but also other folklore genres reveal rich ethnographic details, ranging from various tools and attire used in daily life to the unique features of the geographic environment, encompassing landscapes, climates, flora and fauna. Folklore genres, being rooted in everyday speech and taking substance from everyday life, reflect not only the tangible elements of the material world but also the beliefs, values and understanding of the world at the time when they were created, transmitted and used. Sayings and proverbs stand out from other short folk genres as potent echoes of societal values and beliefs (Bobkina 2004; Babič 2015).

Clothes and clothing culture are a very important element of everyday life. Clothes not only reflect an individual's taste, personality, social status, profession, age and gender, but also other cultural and economic factors (Gordon 2013; Miller 2016: 29). They are such an important part of our lives that they also have great economic importance. This is why various items of clothing and other aspects connected with making clothes can be traced in various folklore genres.

Clothing and the textile industry were even more important in the past. Historically, clothing served as a powerful marker of social status, wealth and identity. In many societies, including the one present on Slovenia's ethnic territory, sumptuary laws regulated what individuals could wear based on their social class or rank (Vilfan 1949; Baš 1970.). Clothing styles, fabrics and accessories can be seen as communication tools, as they indicated one's occupation, affiliation with a particular group or community, or religious and cultural norms, and not only social class but also marital status (Barnard 2005).

Some clothing items have had strong symbolic meanings, mainly because of their role in marking both social and marital status, as well as local or even national affiliation. For example, red clothing and shoes were a long-standing privilege of the aristocracy (Gordon 2013: 152–153), while various head coverings like *šaplji* and *parte* in the Slovenian rural environment were worn by unmarried girls. The symbolism of the *avba*, a special type of bonnet worn by women after getting married, has even been captured in the phrase 'to get under the *avba*,' meaning to get married (Žagar 2004: 19). Clothing has played an important identifying role, not just for individuals but also for entire nations. Certain items of clothing have become significant symbols of national belonging through the creation of national, regional or ethnic costumes (Eriksen 2004; Knific 2010) like the Slovenian national costume, the Norwegian *bunad*, the Bavarian-Austrian *dirndl*, and many other traditional costumes.

Many elements of traditional clothing carried symbolic meanings in their colourings, patterns, embroidery and accessories, as they often represented specific aspects of local traditions and beliefs. As some researchers have noted, certain colours may symbolize fertility, prosperity or spirituality, while specific motifs reflected cultural narratives or folklore. Under certain circumstances, clothes could also have magical properties (Doljak 2013: 73–75; Sedakova 2016).

Before the industrial revolution, textile production was primarily artisanal and localized, as ethnographic data shows. Communities often had their own traditions in textile manufacturing, with skilled artisans handcrafting fabrics, garments and accessories using traditional techniques. The textile industry provided livelihoods for numerous artisans, weavers, dyers, tailors and seamstresses, contributing to local economies and cultural heritage (Geramb 1934; D'Orlandi and Perusini 1988; Makarovič 2007 etc.).

The textile trade played a central role in the economy over a long period of time and in many cultures. Textiles were among the most valuable commodities traded

along ancient trade routes such as the Silk Road, and the demand for luxury textiles, such as silk, but also cotton, wool and linen drove commerce and cultural exchange between distant regions, facilitating economic growth and cultural diffusion (Bogansky 2013; Beckert 2014). The rise of urban centres and industrialization during the 18th and 19th centuries transformed the textile industry profoundly. Innovations such as the spinning jenny, power loom and mechanized textile factories revolutionized production methods, leading to the mass production of textiles and causing a decline in localized artisanal production and thus also the variety of regional types of clothing that had been worn until then. This industrialization of textile manufacturing reshaped urban landscapes, fuelled population migrations and transformed social and economic structures. The growth of the textile industry during the industrial revolution also brought significant social and labour challenges. Factory workers, including many women and children, often endured harsh working conditions, long hours and low wages. This led to labour movements, trade unions and social reforms aimed at improving working conditions, advocating for workers' rights and addressing issues of labour exploitation and inequality (Spielvogel 2005: 583–608). Overall, the historical influence of clothing culture and the textile industry on everyday life and the economy is characterized by the interplay of social, economic, technological and cultural factors, shaping societies and the course of history.

In this article, the terms phrasemes and proverbs refer to multi-word, relatively established and also semantically transferred linguistic units. We should not view phrasemes and proverbs as forms that have remained fixed over time. As will be shown in the article, in many cases the phraseological structure allows for the existence of variants, as well as certain morphosyntactic changes and modifications. Just as the form of phrasemes could change over time, so could their meaning. Some units found in the collection of proverbs stored at the Institute of Slovenian Ethnology are no longer relevant today due to changes in lifestyle, customs and the circumstances related to textile production, which has led to a loss of the connection with the representation that motivated the phrase (Jakop 2022: 130–133). On the other hand, some proverbs have remained important because they allow dual interpretation, both literal and metaphorical.

Proverbs are important carriers of cultural connotation and content, which can be more or less expressed or obscured and may change over time. In this sense, paremiology and phraseology units can also be seen as cultural heritage, keepers of social memory and indicators of the social conceptual world, as they often contain cultural content and can be indicators of changing social circumstances and cultural norms and values (Babič 2020: 125–127; Jakop 2022: 130). They therefore store information about the functioning of society and of the changes that transform it over time.

Through the study of clothing culture per se and its echoes in language, especially in folklore and phrasemes, we can not only gain insights into the economic, social, political and cultural conditions in which people lived in a certain period, but also the

symbolic meanings and beliefs connected with it, as demonstrated through research into other aspects of everyday life and their reflections in proverbs (Babič 2018; Babič and Voolaid 2023). In this article, we will look at what individual textile elements that appear in proverbs indicate, both as elements of everyday life in the past and as an echo of a past craft and form of manufacturing.

2 Methodology

The research was conducted on the collection of proverbs of the Institute of Slovenian Ethnology at ZRC SAZU, which was digitized in the project entitled Traditional Paremiological Units in Dialogue with Contemporary Use. The collection contains 35,706 units and is fully digitized and accessible on the Clarin.si repository.¹

The research started with a random search for words connected with clothing pieces that appear in well-known Slovenian proverbs, e.g. clothes, shoe and shirt. Here are some examples:

Navada je železna srajca (Habit is an iron shirt)

Brez muje se še čevelj ne obuže (Without an effort even a shoe can't be put on)

Obleka naredi človeka (Clothes make the man)

These proverbs are part of the three hundred best known Slovenian units (Meterc 2017).

The second stage was conducted more systematically. The presence of certain words related to clothing culture was checked. For example, words connected with the manufacture of clothes and technology (sewing, spinning, spinning wheel, distaff, etc.), the professions (tailor, seamstress, weaver), raw materials (linen, wool) and fabrics (linen, velvet, silk). Searches were also conducted for names of clothing items and parts of clothing, from general items, such as coats, skirts and shirts, to more specific items such as gloves, socks or even parts of clothing, like pockets.

Most of the proverbs we found were connected with general items of clothing or concepts related to dressing. Meanwhile, proverbs that included specific words, such as the names of pieces of clothing, techniques or materials that were strongly present in the past, e.g. cotton, blouse, scarf, lederhosen, could not be traced.

3 Analysis

Six words were found that appeared in more than fifty units: obleka (Eng. clothes; 89 units), čevelj (Eng. shoe; 68 units), srajca (Eng. shirt; 60 units), žep (Eng. pocket; 53 units), hlače (Eng. trousers; 52 units) and klobuk (Eng. hat; 50 units).

¹ Proverbs (a collection of Slovenian proverbs): <https://www.clarin.si/ske/#dashboard?corpname=pregovori>.

The large number of proverbs that contain these pieces of clothing is mainly due to a large number of variants. These proverbs are among the most well-known proverbs in the Slovenian language and there are therefore many variants that share the same linguistic structure. Here are a number of variants of the Slovenian version of the proverb *Clothes make the man*:

Obleka naredi človeka. / Obleka dela človeka. / Obleka res naredi človeka. / Obleka stori človeka.

The last three mentioned items of clothing that have the most appearances, žep (Eng. pocket), hlače (Eng. trousers) and klobuk (Eng. hat), have a greater number of semantic variants. We can see this in the example of a few units of proverbs that contain the word ‘pocket’.

Mrtvaška obleka nima žepov. (A mortuary gown has no pockets.)

Vsak žep ima svoje dno. (Every pocket has a bottom.)

Vsak gleda za svoj žep. (Everyone looks after their own pocket.)

Puhla glava, prazen žep. (Empty head, empty pocket.)

4 Textile crafts in proverbs

Even today, clothing culture and the textile industry intersect with various aspects of our everyday life and the economy, shaping consumer behaviour, social dynamics and even the economy and environmental issues. The textile industry is still one of the largest and most diverse industries in the world, producing and selling clothing, fabrics, yarns, fibres and other products. In the past, it played a significant role as a driving force behind the Industrial Revolution that started in the middle of the 18th century. Moreover, until the end of the 19th century, local non-industrial production on the territory of present-day Slovenia was also important and was closely connected to self-sufficiency. Certain processes were part of the domestic economy on farms, such as preparing wool and linen fibres for further use. Craftsmen and women, such as weavers, seamstresses, tailors and cobblers were regular visitors who went from house to house. This was especially typical of certain regions until the early 20th century (Makarovič 2007, 2009, 2011).

The significance of textiles and textile crafts, their deep connection with everyday life, and their integration into economic currents, have been gradually woven into Slovenian proverbs. This is evidenced by around 800 proverbs found in the Proverbs collection of the Institute of Slovenian Ethnology at ZRC SAZU. Research into Slovenian proverbs has shown that all aspects of clothing and textile production can be found in proverbs, including fibres, crafting tools, artisans, names of garments and also parts of garments.

Although the majority of textile-related proverbs contain names for clothing items, there are also some mentions of craftsmen. A small number of proverbs mention seamstresses, tailors, shoemakers/cobblers and weavers. Textile crafts were a significant

portion of all the crafts in the past. As many as 14–18% of all craftsmen made or sewed clothes, or were specialized in making special pieces of clothing such as underwear, corsets, ties, etc., between the two world wars (Cerar 2003: 158).

Pri šivilji ni svile kupovati. / Od mojškne židanje draga. (Don't buy silk from a seamstress.)

Krojač naredi gospoda. / Krojači so naredili že precej velikih gospodov. (A tailor makes a man.)

Naj čevljar pri svojem kopitu ostane. / Čevljar ostani pri svojem kopitu! (Let the cobbler stay at his last).

Brez mandranja ne zleze tkalcu platno na berdo. (No linen is made without hard work).

Proverbs not only describe the specific characteristics of professions, but also their attributed qualities and their status in society, such as the notion that buying fabric from a seamstress is expensive, that they are exploiters like innkeepers or that a well-made garment bought from a tailor makes a gentleman. They also encapsulate the seasonal aspects of professions such as shoemaking and weaving, or certain seasonal tasks related to textile production. For instance:

Ko kukavica zapoje, je tkavec ob delo svoje. (When the cuckoo sings, the weaver stops working.)

or *Kolovrat pozna, da nismo klali. (The spinning wheel knows we did not slaughter.)*

These probably indicate that the spinning wheel was more in use when there was not a lot of additional work, such as slaughtering animals and preparing meat, during the winter.

Ethnographic data describes the seasonality of work, such as spinning, flax processing and weaving. For instance, spinning was primarily the work of women and girls, often done during winter evenings, especially in January and February, all over Europe. This task had an important social aspect, as women frequently helped each other. It was not uncommon for the activity to conclude with dancing, feasting and general socializing (Novak 1964: 248; Makarovič 2007: 22; Gordon 2013: 123). Collective spinning typically ended around the time of 'pust' (Carnival), which is reflected also in the proverb

Kadar se začne post, vrzi preslico v gozd. (When fasting begins, throw the distaff into the forest.)

These are borderline paremiological examples, as they offer concrete instructions about what a person must do, rather than ethical-moral guidance, which is typical of proverbs.

Specific tools are also associated with textile production, such as the loom, distaff and spinning wheel. These are all tools needed for preparing yarn and later fabric:

Vsaka preja pride v statve! (Every yarn goes into the loom!)

Kolovrat, skrinja in povesmo – nevesti najlepša pesem. (Spinning wheel, chest and bundle - the bride's most beautiful song.)

Pripravi vreteno in preslico in Bog ti bo poslal lan. (Prepare the spindle and the distaff, and God will send you flax.)

The proverbs also contain terms related to the basic tools and gadgets connected with clothes-making, such as the needle and thread:

Sito očita šivanki, da ima luknjico. (The sieve accuses the needle of having a hole.) and

Prej pojde kamela skozi šivankino uho, kakor nevoščljivec v nebeško kraljestvo. (It is easier for the camel to go through the eye of a needle, than for an envious person to reach the heavenly kingdom.)

Thread is commonly mentioned in European folklore and primarily symbolises the connection between the worlds of the living and the dead, as well as the arrival of children's souls into this world, and death or their departure back to the other world. The motif of spinning or the association with yarn, thread, balls of yarn, wool or spinning tools is connected to various mythical beings, which are usually female. This is understandable as spinning has always been in the female domain (Mencej 2010).

In the collection, most of the proverbs that mention threads are a metaphor for something very thin and fine:

Niti ženske zvijače so finejše, kakor predivo svilene gosenice. (The threads of a woman's cunning are finer than the thread of a silken caterpillar.)

A similar metaphor is also used in proverbs that refer to someone's fortune or property:

Premoženje gre curkom dol, po niti pa gor. (Assets go down in spurts, but up by a thread.)

Imetje se po niti nabira, po vrvi zapravi. (Possessions are gathered by a thread and wasted by a rope.)

Dolg pride po vrvi, odhaja po niti. (Debt comes by rope and leaves by thread.)

In European folklore the thread is often connected with the perception of human life or destiny. In ancient mythology we can find stories about the Moirai, the spinners who wove the destinies of human lives. The Greeks understood fate and its phases, such as misfortune, marriage, old age, illness, or death, as threads spun by the gods or the Fates. Similar motifs can also be found in other European mythologies. For example, there are old Slovene stories about the Sojenice or Fates, which suggest similar beliefs (Kropej 2008: 237–240; Mencej 2010: 160–162).

Traces of the connection between spinning and fate are also evident in language. For example, the Slovenian saying 'presti niti človekove usode' means spinning determines someone's fate (SSKJ 2014). The connection between thread and illness is even more explicit in numerous variations of proverbs, which suggest that disease comes quickly but healing is often more difficult and takes longer.

Bolezen gre v človeka z vozom, iz človeka po niti. (The disease goes into a person by cart, and leaves the person by a thread.)

Bolezen po vrvi, zdravje po niti. (The disease by rope, health by a thread.)

Zdravje curkom doli, po niti gori. (Health goes down in a stream and up by a thread.)

The proverbs also mention the fabrics and raw materials that are used for producing textile fabrics. Although yarn and its symbolism are revealed to a greater extent in other folklore forms, the logical connection of high-quality yarn with high-quality fabric is most strongly expressed in Slovenian proverbs. The collection contains some variations of the proverb 'As the yarn is, so will be the linen' in the form 'Kakršna preja, tako bo platno' or 'Kakor preja, tako platno'.

The raw material is important for producing first the fabrics and then the clothes. For many centuries, flax and wool were the primary materials from which clothes and other textiles were made. Archaeological findings testify that the deliberate cultivation of flax and the breeding of sheep for wool were an important source of income for some places in the Alpine region, and integral parts of textile production since prehistoric times (Gleba and Mannering 2012; Grömer 2016). Wool and flax were also an important part of the Slovenian peasant household economy for many centuries. The peasant population in Slovenia produced woollen and linen fabrics mainly until the arrival of industrialization in the middle of the 19th century. Until then, people had mostly made their clothes and other textile products at home, but the development of the textile industry and technological innovation in the 19th century led to the mass production of textiles in factories. This led to a breakthrough in the way textile products were manufactured and people began buying products that were produced in factories (Makarovič 2007: 33).

Linen fabrics continued to be produced for longer by the Slovenian peasant population than woollen fabrics. Due to the general shortage during the world wars, people again began growing and spinning flax to make textiles in some places in the middle of the 20th century (Makarovič 2007: 21, 2009: 23).

Lan je lan, z njim je posla leto in dan (*Flax is flax, there is work with it all the year round.*) says the proverb and thereby articulates the reality of making linen. Early flax is sown in March and late flax is sown in May, when the early flax has already grown. In June the early flax is harvested and plucked, thrashed to remove the seed pods, retted and prepared for scutching. The late flax is harvested in August, and in the summer linen textiles woven during the winter are bleached and softened. In November it is time to dry the flax, scutch it and comb it, and then spin it over the winter until February. Then the cycle is repeated (Makarovič 2007: 22).

Just how important flax and linen were for the Slovenian peasant household economy is expressed by the proverb *Najlepši zakladi so za gospodinjjo, če ima s platnom napolnjeno skrinjo*. (*The most beautiful treasures are for the housewife if she has a chest filled with linen.*) Desire for good linen is also reflected in the following weather proverb: *Svečnica lep, jasen dan, leto nam dobro bo za lan*. (*Candlemas a beautiful, clear day, we will have a good year for flax.*)

The breeding of sheep for wool is also reflected in a proverb, which articulates the basic difference between breeds bred specifically for meat and wool, and dairy breeds: *Dobra ovca da malo mleka, pa veliko volne*. (*A good sheep gives little milk, but much wool.*)

Silk is also mentioned, in addition to wool and linen. *Raševine ne šivamo s svilo* (*We don't sew rags with silk*), says one of the proverbs. It is clear from the proverbs that silk is a valuable material that is almost equivalent to gold and the two appear together in one of the proverbs: *Rajši knez v razcapani ciganski obleki kakor sluga v zlato in svili!* (*Better a prince in tattered gypsy clothes than a servant in gold and silk!*).

But there is also the proverb: *Ni vse svila, kar je svilnato. (Not everything that is silky is silk)*, which is similar to the proverb *Ni vse zlato, kar se sveti (Not everything that glitters is gold)*. The following proverb also shows that it was (and still is) a valuable and expensive material: *Svila in žamet pogasita ogenj na ognišu (Silk and velvet put out the fire in the fireplace)*. This probably means that beautiful clothes spoil a woman and distract her from caring for her family and home, or that money brings other priorities and the family is no longer in first place. However, we can also understand this proverb as criticising people who get money quickly and for whom presenting luxury on the outside is more important. This could be so important for them that they run out of funds for their basic needs. In both cases the proverb indicates the great value of silk and velvet fabrics.² In fact, in some regions silk was so highly valued and in such demand that it was used as a form of currency and means of payment like gold or money (Hansen and Rong 2013). Throughout history, silk was considered a precious textile that was reserved for the aristocracy, and its use symbolised authority and power. Although silk production in Europe began as early as the 7th century AD (the ancient Greeks and Romans were known as importers of silk from China), throughout history silk was always seen as a luxury that only the wealthiest could afford (Callava 2018). Silk and velvet are therefore primarily metaphors for wealth in these contexts.

Velvet produced from silk fibres is also a symbol of wealth in proverbs. However, proverbs comparing wit and wisdom to velvet argue that the former are worth more: *Pamet je boljša kakor žamet. (Wisdom is better than velvet.)* (Meterc 2020). We can see that the proverb can also express the meaning that it is better to be smart than rich. This could be seen as a criticism of the elites who can afford silk, but what good are silk and money if you are stupid? *Riches serve a wise man but command a fool*³ says an English proverb, which suggests that wealth is only beneficial if one has the wisdom to use it properly. There is also a biblical proverb, which states that it is better to be poor and wise than to be a rich fool, from which these proverbs may have derived: *Better a poor man who walks with integrity than a fool whose lips are perverse.*⁴ This suggests that the Slovenian proverb is actually biblical and adapted to the Slovenian context of life and language.

5 Clothes in proverbs

Slovenian proverbs do not only mention various aspects of textile production but also clothes. Various items of clothing that are mentioned also have a metaphorical meaning, which helps to construct a wider field of connotations. The most frequently mentioned word in proverbs is *obleka* - clothes, suit or garment.

² Velvet was historically made of silk fibres, but is nowadays more commonly made of cotton or synthetic fibres.

³ <https://quoterich.com/list-proverbs-foolishness-wisdom>.

⁴ <https://biblehub.com/nlt/proverbs/19-1.htm>.

Clothes play a vital role in many aspects of our daily lives, such as interpersonal relationships, social situations and in the professional environment. They are important in establishing an individual's mood, commanding respect and often create first impressions (Hutson and Rodriguez 2016). The importance of nice clothes that help make a good impression is also reflected in the proverb: *Obleka (ne) naredi človeka.* (*Clothes (do not) make the man.*) The Latin form of the proverb *vestis virum faciti*, meaning *clothes make the man*, was recorded by Erasmus of Rotterdam, the Dutch humanist, Catholic theologian and philosopher in his work entitled *Adages* in 1520 (Atkins 2017). The proverb expresses that clothes are essential for a person's appearance, dignity and reputation, and are also an identification element (Meterc 2020).

There are also proverbs, which express the idea that beautiful clothes cannot hide an unpleasant character, a person's nature, or turn a bad person into a good one:

Kdor nima v srcu lepote, ga tudi lepa obleka ne polepša. (*Whoever lacks beauty in the heart, even beautiful clothes do not make them beautiful.*)

En madež pokvari tudi najlepšo obleko. (*One stain ruins even the most beautiful outfit.*)

Just how highly valued and important neat, clean, socially acceptable clothing is in western culture can be seen in the following proverbs:

Obrabljena obleka ima malo ugleda. (*A worn suit enjoys little reputation.*)

Lepa obleke ne more vsak nositi, čeden pa mora vsak biti. (*Not everyone can have nice clothes, but everyone must be decently dressed.*)

People are expected to dress appropriately, and above all to be well-presented if they want to be accepted. Dirty clothes and dirt are generally undesired by society.

Clothes and dress can metaphorically indicate something new and interesting, but not for long because they are like interesting news, the novelty of which quickly wears off as new events unfold or as time passes:

Še iz vsake praznične obleke je nastala vsakdanja. (*Every festive suit becomes an everyday suit.*)

Lepa obleka kratko traja. (*A beautiful dress does not last long.*)

In proverbs and other short folklore forms, we can also detect gender divisions and stereotypes (Babič 2018). Trousers are often a metaphor for men, but also a symbol of power and control in the family and in other relationships.

Kjer žena hlače nosi, si mož kruha prosi. (*Where the wife wears the trousers, the husband begs for bread.*)

Gorje možu, ki mu je žena slekla hlače. (*Woe to the husband whose wife took off his trousers.*)

Skirts and aprons are items of clothing that are distinctively connected with women. They are not just metaphors for women but also symbols of motherhood:

Majhni otroci te cukajo za krilo, veliki pa za srce. (*Small children tug at your skirt, while big ones tug at your heart.*)

Mati mora imeti velik predpasnik, da pokrije vse napake svojega otroka. (*A mother must have a big apron to cover all her child's mistakes.*)

6 Weather proverbs and clothes

Weather proverbs form a special group of proverbs that is in no way connected with dressing or the production of textiles and clothing. They indicate weather phenomena but quite a number of clothes items are mentioned in them including skirts, caps, hats, and coats. These items of clothing are not meant to be worn. Instead, they symbolise weather phenomena such as fog, wind, clouds over peaks, etc. This personification is based on the principle of conceptual similarity: a hat covers the head, a cloud covers a mountain, a fog hugs us like a cloak, etc.

Če se sveti nad Gorico, je Italijanka dvignila krilo. (If it (the sun) shines over Gorica, the Italian woman has raised her skirt.)

Če ima Šmarna gora kapo, bo dež. (If Šmarna gora has a cap, it will rain.)

Če je Martinov plašč v meglo zavil, bo vsak od nas še zime sit. (If Martin's coat is shrouded in fog, all of us will get tired of winter.)

Meanwhile, the mention of gloves, fur, coats and shirts on a particular day indicates or predicts warm or cold weather. These items of clothing are connected with direct predictions, with the shirt being a light garment and gloves being winter clothing. However, gloves are not as warm as fur, which is the warmest and most prestigious garment.

V februarju rajši vidim na polju volka - kakor hoditi v srajci moža. (In February, I'd rather see a wolf in the field than a man walking in a shirt.)

Svetega Lovrenca v rokavicah, svetega Jerneja v kožuh. (Saint Lawrence in gloves, Saint Bartholomew in a fur coat.)

These proverbs show the importance of clothing in everyday life, emphasising its self-evident presence and adaptation to changing weather conditions. In weather paremias, the protective function of clothes is emphasised. For example, the mention of warmer clothes indicates colder weather, as it is evident that such clothing protects against the cold and retains body heat. At the same time, specific clothes can also symbolise particular weather conditions precisely because of their protective properties.

7 Conclusion

Words and textiles are inextricably linked. In English this connection is even clearer, as the word text actually derives from the Latin *texere*, which means 'to weave'. Words and symbols have been written on or in textiles for centuries. Silk was a medium for written documents in China in ancient times; in medieval Europe people got to know Bible stories from huge tapestries hung in churches and other public and private spaces; from at least the 17th century women have stitched verses about morality, virtue and family piety into their needlework (Gordon 2013: 216–217). In the period spanning some 150 years, from the French Revolution to 1930, an original form of folk art emerged

and flourished in the countries of Central Europe. Various types of embroidered napkins were produced to adorn walls and tables. The pictures on these works of art were complemented by proverbs that expressed the moral values of the family and society of that time (Villani 2016). Meanwhile, colourful, commercially printed wrap garments known as kanga (also referred to as lesò, kikoy or pagne), which combine imagery and proverbs, are popular among married women in East Africa. These clothes are often used by women as a form of coded expression, conveying messages through both the words and images displayed on the fabric (Gordon 2013: 212–215).

The very existence and number of proverbs mentioning items of clothing, crafts or other elements related to the manufacture of clothes show us the importance of these material goods in everyday life in the past. They reveal the specificity of textile crafts, including clothing and fabric production, and the seasonal timing of certain steps (e.g. spinning was mostly done in winter). They also indicate techniques such as weaving and sewing and mention tools that are used in various phases of textile production, such as the distaff used for yarn production, the loom for fabric and the needle for making clothing. Proverbs also reveal the great value and appreciation of quality clothing products, highlighting the importance of precise craftsmanship by weavers, tailors and seamstresses, the use of good materials, such as wool and linen, and the high value of quality fabrics like silk and velvet.

Certain clothes also have a clear identification role, which was probably established through history by the regulations that limited and prescribed clothing according to social class for a long time. As a result, gold and crimson red are still seen as symbols of aristocracy, while velvet and silk symbolise wealth. Over time, other symbols were also developed, which reflect different social identities. For example, blue shirts or collars became associated with city office workers, and the *avba*, a Slovenian folk head covering, became a symbol of married women. Different materials are associated with specific occasions: silk for festivities, coarse linen for everyday use, finer, more expensive materials, often embellished and embroidered, for celebrations, and cheaper, heavier, rougher and sturdier materials for daily use. This distinction further separates the everyday from the holidays.

The clothes mentioned in the proverbs, especially in the weather paremias, allude to their properties, such as protecting the body from the cold, while also personifying them as weather phenomena. Some items of clothing serve as symbols for individuals and their roles in society. For example, the skirt and apron metaphorically represent motherhood, and the suit signifies not only a person, but often an upper class man.

The analysed proverbs actually express all the economic, social and cultural spectrums that are connected with textile production, clothing use and also the symbolic meanings of garments.

If we return to the proverb that appears most often in the collection of proverbs and its dual nature: *Clothes make or don't make a person*, the answer is clear after analysing all the proverbs connected with textile and clothing production. There is no

longer any doubt regarding whether clothes define an individual or not. Through the lens of proverbs, clothes are undeniably a symbol for humanity, and even civilisation and its norms. The mention of clothing or attire often goes beyond the literal sense and carries deeper, symbolic meanings.

Acknowledgements

The research for this paper was supported by the Slovenian Research Agency: *Traditional Paremiological Units in Dialogue with Contemporary Use* (J6-2579) and *Ethnological, Anthropological and Folklore Studies Research on Everyday Life* (research core funding, P6-0088).

References

- Atkins, Alexander, 2017: What is the Origin of “Clothes Make the Man”? Online: <https://alex-65670.medium.com/what-is-the-origin-of-clothes-make-the-man-7f75e070bf45> (14.12.2023)
- Babič, Saša, 2015: *Beseda ni konj: estetska struktura slovenskih folklornih obrazcev*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789612547660>.
- Babič, Saša, 2018: Stereotipi in predsodki o ženskah in moških v šaljivih vprašanjih. In: Jambrešić Kirin, Renata (ed.), *Humor u svakodnevnoj komunikaciji*. Zagreb: Institut za etnologiju i folkloristiku, 81–95.
- Babič, Saša. 2020. Waste and dirt in short folklore forms. *Traditiones* 49/1, 125–139. DOI: <https://doi.org/10.3986/Traditio2020490107>.
- Babič, Saša; Voolaid, Piret, 2023: A sociocultural view of Estonian and Slovenian proverbs on alcohol and drinking. *Folklore* 134/3: 370–394. DOI: <https://doi.org/10.1080/0015587X.2023.2175972>.
- Baš, Angelos, 1970: *Noša na slovenskem v poznem srednjem veku in 16. stoletju*. Ljubljana: Mladinska knjiga.
- Barnard, Malcolm, 2005: *Moda kot sporazumevanje*. Ljubljana: Sophia (Sodobna družba).
- Beckert, Sven, 2014: *Empire of Cotton: A Global History*. New York: Vintage Books.
- Bobkina, Jelena, 2024: Clothing, sayings and proverbs taken through the prism of the Russian old-believers. Conference paper from 4th Interdisciplinary Colloquium on Proverbs, International Association of Paremiology (AIP-IAP) in Tavira, Portugal. DOI: <http://dx.doi.org/10.13140/2.1.1927.9685>.
- Bogansky, Amy, 2013: *Interwoven Globe: The Worldwide Textile Trade, 1500–1800*. New York: Metropolitan Museum of Art.
- Callava, Trini, 2018: *Silk Through the Ages: The Textile that Conquered Luxury*. London, New York: LID Publishing.
- Cerar, Estera. 2003: *Tekstilna obrt na Slovenskem med obema vojnama: magistrsko delo*. Ljubljana.
- Doljak, Justina, 2013: Ženska oblačila. V: Šuligoj Brevec, Miroslav (ed.), *Oblačilna kultura na Trnovsko-Banjški planoti*. Lokovec: Kulturno turistično društvo, 40–81.
- D’Orlandi, Lea; Perusini, Gaetano, 1988: *Antichi costumi friulani*. Gorizia, Udine: Editrice Goriziana, Società filologica friulana.
- Eriksen, Thomas Hylland, 2004: Keeping the Recipe. *Focaal* 44, 20–34. DOI: <https://doi.org/10.3167/092012904782311308>.

- Geramb, Konrad, 1934: *Steirisches Trachtenbuch*. Graz: Leuschner & Lubensky.
- Gleba, Margarita; Mannering, Ulla (eds.), 2012: *Textiles and textile production in Europe. From Prehistory to AD 400*. Oxford: Oxbow Books.
- Grömer, Karina, 2016: *The Art of Prehistoric Textile Making: The Development of Craft Traditions and Clothing in Central Europe*. Wien: Naturhistorisches Museum. DOI: https://doi.org/10.26530/oopen_604250.
- Hansen, Valerie; Rong, Xinjiang, 2013: How the Residents of Turfan Used Textiles as Money, 273–796 CE. *Journal of the Royal Asiatic Society* 23/2, 281–305.
- Hutson, Matthew; Rodriguez, Tori, 2016: Dress for Success: How Clothes Influence Our Performance. *Scientific American Mind* 27/1, 13–13. <http://dx.doi.org/10.1038/scientificamericanmind0116-13a>.
- Jakop, Nataša, 2022: Govor in prostor skozi frazeološka očala. In: Podbevšek, Katarina; Žavbi, Nina (eds.), *Govor in prostor*. Ljubljana: Založba Univerze v Ljubljani, 127–137. DOI: <https://doi.org/10.4312/9789617128826>.
- Knific, Bojan, 2010: *Folklornikom s(m)o vzeli noše: kostumiranje folklornih skupin - med historično pričevalnostjo in istovetnostjo*. Ljubljana: Založba ZRC, ZRC SAZU.
- Kropej, Monika, 2008: *Od ajda do zlatoroga: slovenska bajeslovna bitja*. Celovec, Ljubljana, Dunaj: Mohorjeva založba.
- Makarovič, Marija, 2007: *Obleka predela človeka: Oblačilna kultura v kmečkem okolju na severu in jugovzhodu Štajerske od 18. do 20. stoletja*. Maribor: Litera.
- Makarovič, Marija, 2009: *Po sledeh bele noše v Beli krajini*. Ljubljana: JSKD.
- Makarovič, Marija, 2011: *Oblačilna kultura v Zgornji Savinjski dolini*. Ljubljana: JSKD.
- Mencej, Mirjam, 2010: Simbolika niti in predenja v evropski folklori. *Studia mythologica Slavica* 13, 151–70.
- Meterc, Matej, 2020: Slovar pregovorov in sorodnih paremioloških izrazov 2020–. Internet: www.fran.si (14. 12. 2023)
- Meterc, Matej, 2017: *Paremiološki optimum: najbolj poznani in pogosti pregovori ter sorodne paremije v slovenščini*. Ljubljana: Založba ZRC, ZRC SAZU.
- Miller, Daniel, 2016: *Materialna kultura*. Ljubljana: Studia humanitatis.
- Novak, Anka, 1964: Domače tkalstvo v Gornjesavski dolini. *Slovenski etnograf* 16/17, 245–256.
- Sedakova, Irina, 2016: Clothes as symbolic and magic objects in slavic childbirth customs. *Folklore: Electronic Journal of Folklore* 66, 165–174. DOI: <https://doi.org/10.7592/FEJF2016.66.sedakova>.
- Spielvogel, Jackson J., 2005: *Western Civilization*. Boston: Cengage Learning.
- SSKJ 2014: *Slovar slovenskega knjižnega jezika: spletna izdaja*. Ljubljana: ZRC SAZU. Online: <https://fran.si/133> (14.12.2023).
- Vilfan, Sergij, 1949: Predpisi o obleki in blagu v policijskih redih 16. stoletja. *Slovenski etnograf* 2: 26–48.
- Villani, Stelio, 2016: Geschickte Hände und ein frohes Herz – Gestickte Sprüche für ein „gemütliches Heim“. In: Konrad Vanja et al. (ed.), *Arbeitskreis Bild Druck Papier, Tagungsband Graz 2015*. Münster, New York: Waxmann, 122–130.
- Žagar, Janja, 1994: *Oblačilna kultura delavcev v Ljubljani med prvo in drugo svetovno vojno*. Ljubljana: Mladika.
- Žagar, Janja, 2004: *Pokrivala: zbirka Slovenskega etnografskega muzeja*. Ljubljana: Slovenski etnografski muzej.

Ethnological analysis of lexeme 'otrok' (child) in the Collection of Slovenian paremiological units on Clarin.si

Etnološka analiza leksema 'otrok' v Zbirki slovenskih paremioloških enot na Clarin.si

This paper presents an ethnological analysis of the word 'otrok' (child) in Slovenian proverbs in order to research the socialisation of children in Slovenia over the past centuries. The analysis is based on the collection of proverbs available at the research repository Clarin.si, which brings together records from every Slovenian region, starting from the 16th century. For centuries, the majority population in Slovenia consisted of peasantry with a thoroughly different approach to childhood. Independence and ability of family members to work were highly valued. Young children were dependent and unable to work, which put them in a socially inferior position. They only gained value through physical strength and experience once they were able to earn their keep. Because of the historically high mortality rate among young children, they represented a risky but worthwhile investment for parents, as they could contribute to their care later in their life.

• **Keywords:** children, proverbs, Slovenia, Clarin.si

V prispevku je predstavljena etnološka analiza leksema 'otrok' v slovenskih pregovorih z namenom raziskovanja socializacije otrok na Slovenskem v preteklih stoletjih. Analiza temelji na zbirki pregovorov, dostopni v raziskovalnem repozitoriju Clarin.si, ki združuje zapise iz vseh slovenskih pokrajin, začenši s 16. stoletjem. Večino prebivalstva na Slovenskem je stoletja sestavljalo kmečko prebivalstvo s povsem drugačnim odnosom do otroštva. Visoko sta bili cenjeni samostojnost in delovna sposobnost družinskih članov. Majhni otroci so bili odvisni in nesposobni za delo, kar jih je postavljalo v družbeno podrejen položaj. S fizično močjo in izkušnjami so pridobili vrednost šele, ko so bili sposobni zaslužiti za svoje preživetje. Zaradi zgodovinsko visoke umrljivosti majhnih otrok so bili za starše tvegana, vendar koristna naložba, saj so lahko pozneje v življenju prispevali k njihovi oskrbi.

• **Ključne besede:** otroci, pregovori, Slovenija, Clarin.si

1 Introduction

Studying folklore literature, folklore forms, including proverbs,¹ equals studying language. Both linguists and folklorists agree that language acts as an inventory of cultural phenomena. Folklorist Saša Babič notes that 'language is culture's material, while at the same time it also functions as its metalanguage' (Babič 2021: 81). Linguist Erika Kržišnik describes a similar dialectical relationship between language and culture: 'This relationship is by no means one-sided in the sense of being limited to culture leaving its trace in language signs, on the contrary, it works both ways:

¹ Use of 'proverb' in this paper signifies paremiological material (proverbs, sayings and a variety of other expressions of 'internal truths' found in the repository at Clarin.si.

language signs also help support collective notions through their reproductive power of intergenerational transitioning' (Kržišnik 2005: 74). Gradual changes in culture and language is the result of the conservative aspect of the human mind. This proverbial slow pace of change (*Pregovori so modrost stoletij / Proverbs are wisdom of centuries*)² leads to proverbs 'carrying information' (Babič 2021: 81) of a past culture or 'cultural connotation' (Kržišnik 2005: 67) of past practices. Several centuries old proverbs remain in use today,³ e.g. *Otroci in golobi (ptice) (p, o)serjejo, ((u)mažejo, ogrdiijo, onesnažijo) hišo / Children and pigeons (birds) shit (over), tarnish, mar, soil a house*. In her monograph *Beseda ni konj*, Babič lists a series of definitions and functions of proverbs, spanning seven pages (!), as their heterogeneity makes it difficult to separate them from other 'sententious statements'. Her conclusion defines a proverb as a 'rounded text' appearing in a broader text 'in daily communication' (Babič 2015: 11). It's characterized by 'stability as a result of rhythm, repetition, rhyme, briefness' (ibid.). The relative continuity ensures the conventionality of a culture's 'linguistic realisation' (Kržišnik 2005). This makes a proverb a stereotype 'carrying figurative or generalised meaning' (Grzybek 2012: 137), and has transferred from generation to generation. The transfer 'is a testament to their extended chronological (historical) presence in a linguistic community' (Kržišnik 2005: 69). Most cultures need proverbs as statements recommending, advising, warning, commanding, prohibiting, etc. (Babič 2018: 82), to maintain every consecutive social and cultural order, as it should be (ideal lifeworld), and express commonplace life experience (realistic lifeworld). It's up to proverbs to transfer cultural practices to the next generation,⁴ thus creating a culture through repetition and gentle alterations, which is apparent from their many existing versions. A change or a difference in transfer is the result of changed subjective or objective circumstances. Babič split proverbs in two groups, namely true proverbs expressing stereotypical principles, e.g. *Brez muje se še čevelj ne obuže / Even a shoe doesn't get put on without an effort*, maxims and instructions, e.g. *Brez dela ni jela / There is no food without work* and *Bolje zlepa kot zgrda / The nice way is better than the hard way*. They are all ready-made, refined through censorship. (Babič 2015: 58–59, 75).

Ethnological analysis in this article focuses on socialisation and regards proverbs as fragments of multigenerational tradition, in the same way as e.g. lullabies (cf. Huzjan 2002). Literary particles from the past have the power to reveal previous habitual

² All proverbs in this paper are written in cursive.

³ Nowadays, proverbs are used for humorous rather than teaching purposes. 2014 research of their use in regions of Styrian, Pannonian and Carinthian dialect groups showed that proverbs mostly represent a passive part of Slovenian language speakers' vocabulary, that they are more frequently used in spoken and less frequently in written language, and that there is a connection between the use of proverbs and the speakers' level of education and age (Lipavc Oštir, Vasiljevič in Koletnik 2014).

⁴ The child is at the centre of intergenerational cultural transfer, and in this regard proverbs are »metaphors teaching children« abstract thinking, cultural norms and integration of social (and religious) rituals with the help of language use (Penfield in Duru 1988: 119). Language is a primary tool of socialisation.

living patterns, as the language of proverbs 'reflects common language' (Babič 2015: 49) and the accompanying common reality, despite being metaphorical in nature. My interest lay in their textual level, in reading 'verbatim', specifically in the analysis of the category 'otrok' (child) ('zibel' (cradle), 'hči' (daughter), 'sin' (son)) in the sense of a human being at the start of life and not in the metaphorical sense of e.g. product (e.g. *Iz šale se rode norci / A joke gives birth to fools; Nevoščljivec je umrl, a je zapustil mnogo otrok / An envious person died and left behind many children*) or consequence (e.g. *Napuh je otrok nevednosti in oče neolikanosti / Vanity is the child of ignorance and the father of impoliteness*). The leading question is: why did 'otrok' appear in a specific proverb and what was its function at the communicative level or, as Kržišnik puts it, what was the 'motivational background [...] of its internal form' or 'picture of the world' or 'perceptive base' (Kržišnik 2005: 67, 69, 77) of the proverb. The choice of lexeme is an important part of drafting a proverb; any proverb could take on a different form or employ any other lexeme. If that had happened, a different culture would have been 'inscribed' in the proverb.

Nevertheless, deliberation on the ethnological analysis of the lexeme 'otrok' runs into a research-wise unsolvable conundrum, the issue of how valid or contradictory proverbs actually are. As folklorist Arvo Krikmann pointed out, the meaning of a proverb can only manifest itself in 'concrete actualisations' (Krikmann 1985), hence in actual use, 'which is why proverbs in collections are nothing but listings, useless for analysis' (Mieder 1993). And he was right: analysing proverbs without the speaker is akin to interpreting dreams without the dreamer's associations. If a proverb only makes sense when expressed in (proverbially diverse) use, it's characterised by, as established by Wolfgang Mieder, by hetero-situativity, poly-functionality and poly-semanticity (Mieder 2004: 132). In addition, proverbs are valid, as pointed out by expert on Slavic languages Peter Grzybek, when they relate to a real-life situation (Grzybek 2008: 30). This leads me to conclude that the validity of proverbs is defined by their use, therefore their contradictory aspect could stem from a change in the speaker's stance. As folklorist Roger D. Abrahams wrote (1975), proverbs were created to express different ideas and positions. They serve as a pragmatic tool for argumentation, while their various roles, stances and meanings simultaneously express contradictions in social and cultural daily life. On one hand, every child is welcome: *Otrok in kozarcev ni nikoli preveč v hiši; otroci pomrjo, kozarci pa se razbijejo / There can never be too many children or glasses in a house; children die, glasses get smashed*, while on the other hand every child is a nuisance: *Siromaku kuj tako otrok ne vmerje, če bi ga čez streho vergel, pa dva nazaj prideta / A poor man's child doesn't die easily, if he were to throw it over the roof, two of them would come back*. Considering this, searching for consistency in proverbs outside of their context seems absurd; the same life situation can produce a host of different experiences, which are articulated in proverbs as 'correct' or 'appropriate' messages. As Babič writes: 'There are as many proverbs as

there are life situations' (Babič 2015: 55). Or, in the words of a proverb: *Pregovori so otroci vsakdanjih izkušenj / Proverbs are children of daily experience*. If regulating a community (Vilfan 1961) through transmitting a system of values (Babič 2015) used to be the essential role of proverbs, the regulation would have been pragmatic.

Linguist Juraj Dolnik strived to understand the origin of antinomies in proverbs. In his view this was how human thinking developed, 'from contrary, polar opposites in a multipart class, to becoming aware of the complementary' (Dolnik 2007: 99, in Meterc 2013: 361). Naturally, contradictions disrupt logical thinking, disturb conscience, while the subconscious, which is subject to primary processes, allows for coexistence of contradictions (Freud 1987: 147–148); testing reality is therefore not effective. Traditional folk literature can be analysed as dreams without the dreamer (Freud 1977: 351; 2000: 328); they are both inhabited by perceptive complexes and desires, which comprise an ideology as a collection of values and beliefs. In this way, proverbs can be seen as censored, rounded forms, with content sourced from the subconscious ideological store room, which is expressed through proverbs. Linguists Alja Lipavac Oštir, Anja Vasiljevič and Mihaela Koletnik skilfully described the way ideology technically operates in proverbs: 'Sociolinguistics sees proverbs as a type of anomaly (cf. Abrahams 1975: 64), as they were originally used primarily as a stylised communication tool, used to transfer the discourse from personal to impersonal level, from the present to the past, from an informal code to a formal and rhetorical one, etc.' (Lipavac Oštir, Vasiljevič, Koletnik 2014: 49). Kržišnik speculates that 'expressions in language serve as a type of "corpus delicti"' (Kržišnik 2005: 68) of thought processes in a specific language / cultural group, thus exposing more about members of this group than they themselves would like to reveal. From this point of view, an ethnological analysis of the lexeme 'otrok' in Slovenian proverbs, despite their lack of consistency and questionable reality, as their use was not part of my research, proves to be a challenging scientific endeavour. Therefore, I will take the risk of entertaining fragments of past authorities (on what is correct and what is the appropriate course of action), that had been 'embedded' in proverbs.

Most proverbs in the discussed collection at the Clarin.si research repository were entered between 16th and 21st century through manuscripts, grammars (e.g. Gutsman 1777), dictionaries, practice and calendars, as well as other periodicals, fiction (e.g. Slomšek 1855), literary collections of proverbs (e.g. Kocbek 1887; Kocbek, Šašelj 1934; Prek 1972; Bojc 1974⁵), expert surveys (e.g. Makarovič 1975⁶), etc. They were

⁵ Etbin Bojc, a Slovenian language professor, included Bishop Anton Martin Slomšek's proverbs in his collection of several thousand proverbs. Slomšek was a 19th century preacher and patriot. Some of his proverbs were common in his time, others were the result of his teaching tendencies (Babič 2015: 78; Meterc 2021: 48).

⁶ Ethnologist Marija Makarovič was interested in more than just recording proverbs and included information on their use, as well, which makes her collection the most relevant to date (Meterc 2021: 59), despite the fact that proverbs and sayings have long lost their former standing, as far as their traditional scheme and function (Terseglav 2002: 56).

recorded by linguists, patriots, 'with the writer's pragmatic and moralistic attitude often playing an important part' (Meterc 2013: 374), as well as folklorists, ethnologists, etc. The collection brings together proverbs used by the local majority population, i.e. Slovenian proverbs, as well as proverbs loaned from other cultural or linguistic regions (German, Latin).

In centuries past, recording proverbs served a pedagogical and distributive purpose. As language used in proverbs is often authoritarian and moralising,⁷ they were a convenient linguistic tool for sermons, i.e. they served as instructions and accessories for spreading the ruling ideology. Collecting and recording proverbs in the 19th century also had an affirmative function in the scope of socio-political tendencies of the time; much like any other form of literary tradition, it was considered 'nation's commodity', practically evidence of the existence of the Slovenian language and consequently the Slovenian nation. In the foreground of this was the category of a nation, which was to be healthy, strong and cultural. E.g. the number of children was a societal norm, primarily expressed as divine affection (*Kolikor otrok, toliko blagoslova / The number of children equals the number of blessings*), and secondly as a political project (*Veliko otrok imej, da ostali bomo večno gospodarji teh planin in sten / Have plenty of children, so we can forever remain lords of these mountains and slopes*). 'Throughout the 19th and 20th centuries, collectors chiefly copied Slovenian paremiological expressions from one collection to another, without any information about the source, and enriched copied material with material they discovered during fieldwork, but seldom clearly divided the two types of material' (Gryzbek 2008: 43). The collection of Slovenian proverbs at Clarin.si is a conglomerate of different time periods, social contexts and cultural surroundings of both recorders and speakers. This led paremiologist Matej Meters to question whether we could read their meaning 'with a fairly high degree of accuracy [...], due to the ostensible nature of their transparency in meaning' (Meterc 2021: 46). The question is indeed an important one; to offer a slightly humorous response, I would argue that experimenting with thoughts is central to all research, as long as this is done within the framework of responsibility to both public and science.

To conclude: the analytical question of this article focuses on the concept of childhood as an elementary thinking tool for personnel, responsible for socialisation. I'm interested in the answer to the question of what people who recorded proverbs or those who used them and passed them on through oral transmission (family, teachers, priests) thought of the signified, i.e. 'otrok'. The question is, what metaphorical language of proverbs conveys about socialisation, what kind of image of a child is (was) this socialisation based on. Socialisation demands the subject to submit to the existing normative order. Ways in which socialisation is realised depends on adults'

⁷ This is illustrated by Heda Jason's (1987) classification of proverbs: summary of life's experiences; advice based on experience; advice to the listener on their conduct; advice on desired conduct; warning about undesired conduct; criticism of a person's behaviour (Babič 2015: 56).

attitude towards children. This attitude is defined by cultural norms, which depend on social-economic conditions prevalent in specific periods and related family survival strategies, which govern intergenerational relations within a family (s.v. Turk Niskač 2021; Huzjan 2023). Analysis of approximately 600 proverbs resulted in the following semantic groups, which define the category 'otrok', care and socialisation of children within patriarchal discourse.

2 *Otrok je neumen in nedolžen, odrašeni [odrasli] je pameten ali poln greha, kateri je bolj srečen? (A child is stupid and innocent, an adult is smart and filled with sin, which one is happier?)*

The answer to the question in the title isn't complicated. The one characteristic of children the adults are often envious of, is carelessness: (*Otrok in pijanec imata srečo / A child and a drunkard are lucky; Pijanci in otroci imajo posebnega angelja varuha / Drunkards and children have a special guardian angel*), therefore *Del ko je človek lahko otrok, del mu je lepo / The longer a man can be a child, the longer he has a good time*. And who is a child? Children as a subject are endowed with characteristic attributes: *Tudi kraljevi otroci so otroci / Even the king's children are children*, yet at the same time there are differences among them: *Otroci so kakor prsti na roki: vsi so tvoji, pa dva nista enaka. / Children are like fingers of a hand: they are all yours, but not two of them are the same*. The first noticeable characteristic is their physical weakness: *Težko junaku, ki ga branijo otroci / Woe to a hero who is defended by children* and *Kdor vrže pijanca ali otroka, naj se preveč ne baha / Whoever overpowers a drunkard or a child should not brag too much*, followed by them being uncultured: *Kdor se ukvarja z otroki, si umaže srajco / Those who deal with children get their shirt dirty* and *Otroci in golobi umažejo hišo / Children and pigeons soil the house*. Which characteristics of children upset the world of adults?

Firstly, it's a child's crying: *Otrok raste z jokom / A child grows with crying*, which is also close to the elderly: *Otroci in stari ljudje kmalo (brž) jok s(iz)tresejo / Children and old people quickly shake out crying*. It sometimes happens that *Dete plače, mati ne čuje / A child cries, mother doesn't hear*. This is better from the point of view of upbringing, as *Smerkov otrok, prevzeten cvet / A snotty [crying] child, a conceited flower* and modesty was considered a virtue, practically *lepa čednost / best policy*. Other forms of (defensive) rationalisation of a child's crying are revealed by a comparison between children and parents: *Bolje, da joka otrok, nego oče / It is better for a child than a father to cry; Boli je da se derejo otroci kak pa starši / It is better for children than for parents to scream [cry]*. Adults supposedly have more serious reasons for crying and are harder to console, as *Otrok ima jok in smeh v jednem (enim) meh(u) / A child has crying and laughter in one bellows*.

Next on the list of irritating qualities when it comes to children, are the noise and racket that comes with them: *Otročaji so kričaji / Children are screamers; Otok je slavček⁸ v hiši / A child is a nightingale in the house; Vrane⁹ čivkajo na strehi, otroci pa v hiši / Crows chirp on the roof, children chirp in the house*. This is understandable as *Ni zdrav otrok, ki se ne igra in ki ne odpira ust / A child that doesn't play or open their mouth is not healthy*, therefore *Mati molčecemu detetu nič ne veruje / A mother has no trust in a silent baby*. Because *Mlada kri mirno ne stoji / Young blood does not stand still*, it is *Slabo znamenje, če so otroci mirni / A bad omen, when children are calm* and *Če so otroci tiho, so nekaj ušpičili / When children are silent they have been up to mischief*.

Children are inquisitive (*Otroci so najboljši opazovalci / Children are the best observers; Majhni vrči imajo velika ušesa / Small jugs have big ears; Otroci imajo majhne oči, pa velika ušesa / Children have small eyes and big ears; Otroci imajo velika ušesa in dolg jezik / Children have big ears and long tongues*) and truthful (*Otroci ne povedo nič drugega kot to, kar so slišali / Children never say anything other than what they've heard; Resnica prihaja iz ust otrok / Truth comes out of the mouths of children; Če hočeš zvedeti resnico, vprašaj otroka / If you want to know the truth, ask a child; Vino in otroci govorijo resnico / Wine and children speak the truth; Otroci (norci; neumniki; pijanci; vino; stare klepetulje) govore (povedo) resnico / Children (fools; nitwits; drunkards; wine; old babblers) speak (tell) the truth; Otroci in starci ne znajo lagati / Children and old people can't lie; Otroci in norci ne lažejo / Children and fools don't lie; Resnice poiđi iskat k svetnikom, k otrokom, k pijancem in k drugim nespametnim ljudem, pri modrijanih je ne najdeš / Seek the truth with saints, children, drunkards and other foolish people, you are not going to find it with wise men, because *Otok se moti malokdaj, večkrat se motiš ti, starec / A child is seldom wrong, you, old man, are wrong more often. Človek se mora pri otrocih neumen narediti / A man must play dumb with children*, because a child's inquisitiveness and truthfulness provide adults with information about the community and thus a feeling of control: *V mlinu in od otrok se vse izve / One can find out everything in a mill and from children; Kjer so bedaki, otroci in vino, tam nič ne ostane dolgo skrito / Where there are nitwits, children and wine, nothing stays hidden for long*. There is one problem with this, namely that *Kar otroci slišijo doma, kmalu poleti v svet / What children learn at home soon flies out into the world*, one should therefore be careful about what they say.*

Preacher Anton Martin Slomšek realised that a stupid person is one beyond good and evil (Slomšek 1855: 67). Children are not capable of sound judgement and are therefore stupid: *Otokom vse moreš dat, somo pameti ne / You can give children everything*

⁸ A nightingale is a stereotypically small bird with a piercing voice.

⁹ Crows are stereotypically clever and merging, as well as loud and screeching.

but cleverness; *Lahko daš otroku kruha (če je lačen), pameti pa (mu) ne (moreš dati) / You can give bread to a child (if they are hungry), but cleverness you can't (give) (them)*. They are at their most unreasonable when it comes to managing money and time, like 'fools': *Le otroci in norci mislijo, da sto goldinarjev in sto dni nikoli ne mine / Only children and fools think that a hundred guldens and a hundred days never come to an end; Le otroci in norci mislijo, da dvajset let in dvajset petic [srebrniki] večno terpi [traja] / Only children and fools think that twenty years and twenty silver coins last forever ; Deca in duraki [norci] mislijo, da dvajset orlažev in pa dvajset let bez konca so / Children and fools think that twenty coins and twenty years have no end; Če grede otroci in norci na trg, kramarji služijo / If children and fools go to the market, pedlars get paid; Lahko trgovcu, kadar imajo otroci denarja / Lucky is a merchant when children have money; Gda majo norci no deca peneze, te Židje tržijo / When fools and children have money, Jews are trading; Kadar imajo otroci dosti denarja, imajo kramarji dober beret / When children have lots of money, pedlars have good business; Kadar (če) imajo otroci (in norci) denar (novac), imajo(takrat so) kramarji (štacunarji; krošnjarji; trgovci) semenj (veseli) / When (if) children (and fools) have money, (then) pedlars (shopkeepers; hawkers; merchants) have a (merry) fair; Ko bi imela deca denar, bi bili kramarji bogati / If children had money, pedlars would be rich*. As noted by Kržišnik (2005: 76), time and money are related to work,¹⁰ which is why children's irrational actions when it comes to time and money stem from the lack of their ability to work. Children need to be guided and taught due to their inexperience (*Kure in otroci ne znajo pota / Chicken and children don't know the way*)¹¹ A foolish person is one who deviates from cultural norm, who is not (yet) socialised.¹² If *Odrašene [odrasle] varuje pamet, otroke Bog / Grown-ups are guarded by cleverness, children by God, if Bog daje pametnim srečo, boben (igrač) otrokom / God gives happiness to the smart and the drum (toys) to children*, then *Neumnemu detetu niti mati ne more pomoči / Even the mother can't help a stupid child*.

Considering the aforementioned examples, the phrase 'stupid child' seems tautological. Yet the tone of a proverb reminds us that *Deset bistrih (umnih) otrok - enajst sreč (pri hiši) / Ten bright (smart) children - eleven counts of luck (in the house)*, and that *Pametni otroci so največja bogatija staršev / Smart children are the parents' greatest wealth*. This confirms the existence of 'stupid' and 'smart' children. What did being bright, smart mean? Surely such a child was more experienced and prudent, more cultured. They were raised 'right': to be industrious, obedient, docile, cooperative.

¹⁰ The industrial age brought a change in the way work was expressed. If it used to 'typically relate to time it required', it was now expressed in money – wage (Kržišnik 2005: 76).

¹¹ For more on this, see section *Kar ni za zrelega moža, ni za otroka / That which is unsuitable for a mature man, is unsuitable for a child*.

¹² In this regard, a fool or a stupid child do not personify a special needs child, as per Alicja Fidowicz's (2021) weak arguments. In folklore, foolishness is a wider concept, covering special needs children, as well.

Subordinate and at the same time restless, impatient, ill-tempered: *Brihten otrok je kakor brencelj (zolja)*¹³ / *A bright child is like a horse-fly, paradoxically – indomitable.* Saša Babič defined a smart person in 19th century proverbial tradition as one who is educated, like 'gentlefolk' (Babič 2018: 84), adding that cleverness was considered 'the most valuable trait a person can possess' (Babič 2018: 85). In the case of a child that would mean that they could get educated and become 'gentlefolk', thus acquiring a higher social status and a more comfortable life. Having a clever child gave parents hope for economic and social welfare, as the clergy provided education to smart children from poor families, thus securing their survival. Therefore: *Moder sin razveseljuje očeta, neumen sin je žalost svoje matere* / *A wise son is his father's joy, a stupid one his mother's sorrow.*

The categories 'inquisitiveness', 'truthfulness' and 'stupidity' all fall under the umbrella category of 'innocence'. *Kdor misli kot otrok, ne misli slabo* / *One who thinks like a child, doesn't think badly*; *V otroku moraš videti otroško srce in dušo* / *You must see a child's heart and soul in a child*; *Otroci so otroci - bavijo se z otročarijami - Children are children - they occupy themselves with childish things*; *Ko sem bil otrok, sem govoril kakor otrok, mislil kakor otrok, sklepal kakor otrok*¹⁴ / *When I was a child, I spoke as a child, thought as a child, reasoned as a child.* In their innocence, a child is stupid, truthful and inquisitive. The innocence is related to being god's: *Otroci so božja dediščina* / *Children are god's heritage*; *Otrok je božji dar* / *A child is a gift from god*; *Vonj otroka prihaja iz raja* / *The smell of a child comes from heaven*; *Otrok ne razume božjega imena, a ga Bog vseeno ljubi.* / *A child doesn't understand god's name, but god loves them just the same.* Children are god's property, they are only on Earth in foster care: *Bog da staršem otroke, ali mu jih bodo morali vrniti.*¹⁵ / *God gives children to parents, but they will have to give them back.* The fact that *Otroška molitev oblake predira* / *A child's prayer pierces the clouds*, means that *Bogu kos vsaki otrok, junak pa nobeden* / *Every child can equal god, but no hero can.* By their divine nature, children are able to protect a house: *Vrag ne vstopi v hišo, ki je polna otrok* / *The devil doesn't enter a house full of children* and *V hišo, ki je polna otrok, hudič nima vstopa* / *The devil has no access to a house full of children*, and they deserve the utmost respect: *Otroku gre največje spoštovanje* / *A child should get the utmost respect.*

The idea that children are without sin, innocent and authentic, i.e. god's messengers, is part of a romantic discourse about children. A romantic vision of a child, personifying the 'natural' and 'emotional' principle as opposed to the 'material world' of industrial production, emerged as an answer to a rising industrialisation of urban

¹³ A botfly is an insect 'katere ličinke se zajedajo / whose larvae burrow themselves' (Fran: <https://fran.si/iskanje?View=1&Query=zolja>) into cattle, horses, sheep, etc.

¹⁴ The text is taken from Paul's first letter to the Corinthians (<https://www.biblija.net/biblija.cgi?m=1+Kor+13%2C8-13&compact=1&id13=1&pos=0&set=2&l=sl>).

¹⁵ This proverb provides solace in the case of a child's death.

areas (Burcar 2007; Kroflič 2011). In his texts instructing religious upbringing, bishop Slomšek paradoxically wasn't shy about enforcing the necessity of corporal punishment (Košir 2023), which common sense subsequently realised in disciplinary procedures aimed at taming the wild nature of children,¹⁶ because *Otroci vohko starše sprajijo na struap ali pa u gruab / Children can drive their parents up the wall or into a grave; Raje z možem v pekel, kakor z otroki v nebesa / Sooner to hell with a husband than to heaven with children*, and was closer to the upbringing ideal of the Enlightenment (cf. Cunningham 1996).

The romantic imaginery disappears in the moment one focuses on differences among children. They differ in material living conditions and gender. *Sitega očeta otroke (deco) je lahko spoznati / It's easy to identify children of a father who is full*, although this doesn't mean that children from noble families are also raised right: *Najboljši starši imajo često nevezgojene otroke / The most distinguished parents often have ill-mannered children* or *Pametnan mož ima po navadi naumne otroke / A clever man usually has stupid children*. The attitude of an adult toward a rich or a poor child differed, as fawning over a rich child could prove beneficial to a poor person: *Bogatim novemu otroku se vse mili, reveževega pa vsaki peha od sebe / Everyone is sweet to a rich man's child, while everyone pushes a poor man's child away*. Grammatically, the word 'dete' (baby/child) is neutral in gender, meaning that such a being is non-differentiated in terms of gender, has no societal role and is supposedly outside the cultural norm (cf. Babič 2021: 82). The moment a cultural environment ascribes gender to a child is the start of gender segregation, as well as comprehending and dealing with a child according to their gender. Or the other way around: 'son' and 'daughter' appear in order to allow the speaker to utter the gender difference with accompanying gender role, as well as the intergenerational relationship: *Sin je tvoj, hči pa ne / A son is yours, a daughter isn't*. A son will carry the family name and inherit property (patrilineality), a daughter leaves the house (patrilocality). If *Sin lahko postane knez, hči mati / A son can become a prince, a daughter a mother*, that means that a son's future is less certain, but he also has more possibilities to actualise his life. This points to inferiority of the female gender: *Ko je hčerka rojena, je še mravlja razžaljena / When a daughter is born, even the ant gets offended*. From the moment she is born, a daughter secures the family's emotional economy (offence), and this will remain her mission: *Kdor nima hčeric, nima (tudi) solzice / The one who hasn't a daughter, also hasn't a tear*, i.e. doesn't emote. There is another way to read this proverb: 'The one who has a son has no regrets.' In patriarchy, the name, property and power over family members are automatically passed down from father to son.

¹⁶ For more on educational guidelines, see section *Kar ni za zrelega moža, ni za otroka / That which is unsuitable for a mature man, is unsuitable for a child*.

Children are emotional, inexperienced, reckless and unreliable (cf. Babič 2018), as well as weak, and therefore unsuitable for work (cf. section *Otroci so večni jarem / Children are an eternal yoke*). As they don't understand categories of the living environment, which are essential to adults – time and money, i.e. work –, they remain carefree and are possibly envied by adults, which often manifests in expressions of inferiority. Women are ascribed similar qualities (*Kadar se otroci valjajo in ženske kregajo, bo dež / When children roll on the ground and women squabble, rain is coming; Če imaš cibe, otroke in babo, imaš vse pogoje za se kregati / If you have chicken, children and a broad, you're in condition to argue ; Žena in otroci – to so računi / Women and children – they are expenses*), as are fools (*Tri pote sem norel: ko sem bil otrok, ko sem se oženil in ko sem zasajal murbe! / I was crazy three times: when I was a child, when I got married and when I was planting mulberry trees*), drunkards (*Otroci in pijanci resnico govore / Children and drunkards speak the truth*), old people (*Otroci in stari ljudje so za vkup / Children and old people belong together; Star dvakrat otrok / Once old, twice a child; Kmet mora biti dvakrat pastir: kot otrok in na stara leta / A peasant has to be a shepherd twice: as a child and in old age; Vsak mora biti le enkrat otrok, če ne mlad, pa star / Everyone must be a child at one time, if not when young, than in old age; Kratek čas delajo leta: otroka kratkočasijo igrače, fanta dekleta, moža gospodarstvo, starca peč in otroci / Diversion comes with age: a child is diverted by toys, a boy by girls, a man by economy, an old man by a furnace and children*), even sailors (*Nebesa varujejo otroke, mornarje in pijance / Heaven protects children, sailors and drunkards*), pigs (*Detetu in prasetu ni mogoče nič dopovedati / There's nothing one can get across to a child or a pig; Kdor se meša med otroke, ga svinje pojedjo / The one who mingles with children gets eaten by pigs*, as *Otroku je za kratek čas svinjka, starcu pa še naga dekle ne / A child is entertained by a pig, an old man not even by a naked girl*), poultry (*Zaradi otrok in kur se ne prepiraj / Don't argue over children and poultry*) and dogs (*Z otrokom in psom se ni za bost / It's not worth quarrelling with a child and a dog*).

The abovementioned subjects (women, old people, fools, drunkards and animals) are on the margins of a patriarchal family. Children used to be the last link of hierarchical social stratification: *Kjer so otroci, ni ne sorodstva ne prijateljev / Where there are children, there is neither family nor friends*, so it doesn't come as a surprise that: *Če trka beda na vrata, se to po deci najbolj pozna / If misery comes knocking, the children feel it the most* and *Dedeji sredico mudlajo, otroci skorjo grudijo / Old men eat the soft centre, children eat the crust*.

3 Otroci so večni jarem (children are an eternal yoke)

Taking care of a child, i.e. feeding and clothing them as well as taking care of hygiene fell to women: *Matere še ne naredi rojstvo otroka / A birth of a child does not make a mother*; which means that *Otroka ni težko (je lahko) roditi, temveč (a težko) vzgojiti / It's not hard (it's easy) to give birth to a child, it's hard to raise them*. Women used to work in the house, barn and fields. They either brought their children along or left them home, where they were cared after by a grandmother, a female farmhand or an older sister (Žagar 1997; Knežević Hočevar 2013: 108). They were a handful (*Dokler so otroci, dela ne zmanjka / As long as there are children, one never runs out of work*; *Majhni otroci in majhni dolgovi dajo veliko opraviti / Small children and small debts provide ample labour*) and a source of considerable worry (*Ko se otrok rodi, nastanejo skrbi / When a child is born, worries come*; *Otrok je skrb, naj bo poreden ali dober, bolan ali zdrav / A child means worry, whether naughty or good, sick or healthy*; *Kjer so otroci, tam so skrbi / Where there are children, there are worries*; *Ki nima otrok, ne pozna nadlog / The one who has no children, knows no bother*; *Vsak otrok velja (mater) en zob / Each child costs (the mother) one tooth*; *Kolikor otrok, toliko očenašev / The number of children is the number of the Lord's Prayers*), because they had to be kept alive – *Otroci in cvetje imajo radi nego / Children and flowers like care* – and socialise: *En otrok, en vrag / One child, one devil*.

Children were seen as an economic investment: *Otroci še nikogar niso naredili bogataša / Never have children made anyone rich*. They had to be fed: *Otrok ne joka od veselja, ampak od glada / A child doesn't cry with joy, they cry with hunger*; *Vsak otrok prinese s seboj žlico na svet / Every child brings a spoon along into the world*; *Bog je ustvaril otroka in z njim na svet poslal tudi žlico / God created a child and sent a spoon along into the world, thus 'eating into' (sucking, pecking, eating) the family economy*: *Otroci sesajo mater, ko so majhni, očeta pa, ko odrastejo / Children suck on the mother when they are little, and on the father when they grow up*; *Otroci in piščanci morajo nenehno kljuvati / Children and chicken must peck constantly*; *Malo dete če jesti z devetimi kosci / A small child wants to eat with nine harvesters*; therefore *Če imaš otroke, niso vsi grizljaji tvoji / If you have children, not every bite is yours*. Children presented a cost: *Ena napačnost potrebuje več stroškov, kakor dva otroka / One wrongness requires more costs than two children*, and the more there were, the greater the burden on family economy: *Prvi otrok je zlato, drugi srebro, tretji pa blato / The first child is gold, the second silver, the third mud*,¹⁷ *Veliko otrok, veliko dolgov / Many children, many debts*, because *Odkod je kaša, otrok ne vpraša / A child doesn't ask where porridge has come from*; *Otroci ne vprašajo, po čem je kruh / Children*

¹⁷ This is contradicted by the following proverb: *Prvi otrok je igračka, drugi je čačka, tretji je šele otrok / The first child is a toy, the second is a doodle, only the third one is a child, thus using the category of 'child' to illustrate the proverb V tretje gre rado / Third time's a charm*.

don't ask about the price of bread. Hence the instruction: *Od koruzne setve pravijo: Kdor ima veliko otrok, naj seje gosto; kdor jih pa ima malo, naj seje redko!* / *This is said of sowing maize: Those who have many children, should sow densely; those who have few, should sow sparsely!*

For the majority of rural population, children presented a risky investment due to their high mortality: *Krsta je sestra zibeli* / *A coffin is a cradle's sister*. When parents were referencing dying of children, it was said that: *Kdor ima veliko otrok, naj seje proso redko, kdor ima pa veliko živine, naj ga seje gosto* / *Those who have many children should sow their millet sparsely, those who have a lot of livestock should sow it densely; Če seješ za otroke, sej žito redko, če seješ za žival, na gosto* / *If you sow for children, sow sparsely, if you sow for livestock, sow densely*. Raising livestock reimbursed the costs, while children can be unprofitable. Child mortality was caused by hunger, poor housing conditions, disease, injuries, epidemics, wars, and also a lack of loving closeness in the first months of life, because *Otrok potrebuje več kot le zibelko* / *A child needs more than just a cradle*.

Veliko otrok sive lasi dela / *Many children [due to the aforementioned reasons] cause grey hair*, until the time they start working: *Otrokove roke in svinjsko korito morajo biti zmeraj polni* / *A child's hands and a pig trough should always be full*, because *Vsak otrok prinese s seboj roke za delo, žlico za jelo* / *Every child brings along hands for working and a spoon for eating*, so they can feed themselves, i.e. *Vsak otrok svoj kos kruha na svet prinese* / *Every child brings their own piece of bread into the world*. An agricultural holding thus included every member of the household and diligence was highly valued. On account of being weak and unreliable, young children helped around the kitchen, stable or pasture, because *Gorje mu, ki s teleti orje in z otroci kmetuje* / *Woe betide him who ploughs with calves and farms with children* and *Kdor s teleti orje in z otroki dela, ga rada glava boli* / *Those who plough with calves and work with children often have headaches*, while older children were performing increasingly more responsible chores. When it came to children, becoming independent and able to work as fast as possible was appreciated, while older members of the family were expected to retain these abilities for as long as possible (Turk Niskač 2021: 63–67; Knežević Hočevar 2013: 109).

Otroci zanesljive skrbi negotove ugodnosti / *Children [were] certain worries [and] uncertain benefits*. What, apart from the aforementioned child labour, are the benefits in question? Since children presented a cost, they were living in debt until their parents grew old: *Otrok šele takrat plača materi, kadar ji glavo umije* / *A child only repays their mother when they wash her head* and *Otroci staršem ne dajejo, ampak vračajo* / *Children don't give parents, they give them back*.¹⁸ Their debt is relative to the initial

¹⁸ The fourth commandment – Honour your father and mother – ascribes a child's debt to parental sacrifice, which is revealed as an expense in the microeconomy of a rural home.

investment: *Od otrok pričakuj le toliko, kolikor si jim dal / Only expect as much from your children as you have given them.* The children only became aware of their debt when they became parents themselves: *Koliko si dolžan svojim staršem, se zaveš šele takrat, ko imaš sam otroke / You only realise how much you owe your parents, when you have children yourself.* Investing in a child therefore meant investing in one's old age, as children were expected to take care of their incapacitated parents. I am going to revisit this motif in the segment *Kar ni za zrelega moža, ni za otroka / That which is unsuitable for a mature man, is unsuitable for a child*, suffice it to say that there was great concern about debt being annulled, reflected in the instruction: *Ljubi svojega otroka, dokler je majhen, da te ne bo sovražil, ko odraste / Love your child while they are young, so they don't hate you when they grow up.* Fear of being abandoned in old age: *Podpore otrok se nadejati je naravno, a goljufivo / It is natural, yet misleading to hope for one's children's support*, meant worries actually began with young adults: *Prave porodne bolečine pridejo šele, ko otroci zrastejo / True labour pains come when the children are grown, although Poročen otrok je le še pol otroka / A married child is only half a child.* At that time parents no longer wield direct control over their child's actions (*Sin (je) moj, (pa) um (ima) svoj / A son (is) mine, (but has) mind of his own*), which worsened their worries regarding their own future. The difference between caring for a small child and worrying about a young adult is depicted in a comparative between 'small' and 'big': *Mali (majhni) otroci/k (dete), male/a skrb(i), velik(i) otroci/k, velike/a skrb(i) / Small child (children, baby), small worry/ies, big child (children), big worry/ies; Dokler so otroci majhni, so z njimi majhne skrbi, ko so pa veliki, pa velike / As long as children are small, they present small worries, and big ones when they are big; Majhen otrok - majhna skrb, velik otrok sive lase dela / A small child – a small worry, a big child makes grey hair; Majhni otroci stopajo materi na krilo, veliki (pa) na srce / Small children step on their mother's skirt, (but) big ones on her heart; Majhni otroci te cukajo za krilo, veliki pa za srce / Small children tug at your skirt, big ones at your heart; Mihni otroci lulajo na krivo, vejki otroci pa na srce / Small children pee sideways, big children on the heart; Majhni otroci jedo juho, veliki pa srce svojih staršev / Small children eat soup, big ones their parents' heart; Majhni otroci jedo kašo, veliki požirajo srce / Small children eat porridge, big ones devour the heart; Zaradi majhnih otrok boli glava, zaradi velikih srce / Small children bring on a headache, big ones a heartache; Otroček stopa roditeljem na prste, odrasel na srce / A small child steps on their parents' toes, a grown one on the heart; Majhne punčke raztrgajo materi predpasnik, velike pa srce / Small girls tear their mother's apron, big ones her heart; Majhni otroci ne dajo spati, veliki ne dajo dihati / Small children won't let you sleep, big ones won't let you breathe; Majhni otroci ne pustijo spati, veliki pa ne živeti / Small children won't let you sleep, big ones won't let you live; Najprej nosiš otroka na rokah, nato v naročji in slednjič na hrbtu / At first you carry a child in your hands, than in your lap and in the end on your back.* Small children need to be taken

care of: they need help and care: they step on a skirt, on toes, they tuck at a skirt, tear an apron, need to be held in arms, they pee sideways, eat soup, porridge, won't let you sleep. With older children, the parents' agony only intensifies: they step and urinate on the heart, eat, devour, tug, tear the heart, bend the back, suffocate and generally don't let live. This is about gradual loss of control over one's own future.

Hence *Otroci so veselje in žalost (so greh in smeh) / Children are joy and sadness (are sin and laughter)* and *Kjer ni otrok, ni muk ne radosti / Where there are no children, there is no anguish and no joy* or *Dom z otroki - bazar, dom brez otrok - grobnica / A home with children – a bazaar, a home without children – a tomb*. Due to high child mortality and obstinacy in older children, they represented a risky investment: *Polna hiša otrok je kot polna košara jajc / A house full of children is like a basket full of eggs*; they had a high (economic) value: *Kdor otroku roko da, seže staršem do srca / The one who gives a child their hand, touches the parents' heart*, yet their 'frailty'¹⁹ made them an unreliable pair of working hands and support for their parents once they reached old age. Proverbs focusing on caring for children in the context of an agricultural holding instructed the use of economic categories. The discourse on children was an economic one.

4 Kar ni za zrelega moža, ni za otroka (That which is unsuitable for a mature man, is unsuitable for a child)

As opposed to African tradition which stipulates that it takes a village to raise a child, socialisation in Slovenia was a family matter (*Kjer je veliko žensk (babic), tam je otrok kilav / Where there are many women (midwives), the child is crummy*)²⁰ and a matter of the church. Upbringing was left to women and men, but they occupied different roles: the woman took over essential care and corrected small offences, while the men represented fear and respect, a punitive authority feared by the women, as well. A child's upbringing began with their parents' upbringing: *Otroka lahko vzgoje le vzgojeni starši / A child can only be brought up by well-bred parents*; *Največja nesreča za otroke je, če imajo nezrele starše / Immature parents are the greatest misfortune that can befall children*, because *Kadar se kregajo starši, najbolj trpijo otroci / When parents argue, children suffer the most*. Therefore: *Kjer ni mirnih ljudi, tam ni božjih otrok / Where there are no calm people, there are no children of god*. Parental quality was of vital importance for growing children, as parental encouragement or reprimand held an

¹⁹ Eggs were seldom featured in a peasant diet and were reserved for holidays, important guests and the infirm, while most eggs were intended for sale.

²⁰ The proverb says that 'kjer sodeluje preveč ljudi, ni pravega uspeha / where too many people participate, there is no real success' (Keber 2015; https://fran.si/192/janez-keber-frazeoloski-slovar-slovenskega-jezika/4215434/T4MVC_System_Web_Mvc_ActionResult).

enormous (magic) power over them: *Materin (očetov) blagoslov otrokom hiše zida (utrjuje), očetova (materina) kletev jih pa (do temeljev) podira / A mother's (father's) blessing builds (fortifies) children's houses, but a father's (mother's) curse knocks them down (to the foundations) and Starišev blagoslov zida otrokom hiše / A parents' blessing builds houses for children!*

Children's upbringing started very early: *Upogibaj drevo, dokler je mlado / Bend the tree while it's young*, as they believed that children were more susceptible to teaching measures in their early years: *Kar se otrok naučiš, to boš (bodeš) v starosti imel(a); kar v mladosti zamudiš, tega starec (stara) ne boš žel(a) / What you learn as a child, you (will) have in old age; what you miss in your youth, you will not reap in old age; Kdor se kdaj - ali zdaj otrok uči, pomnil bo vse žive dni / Whatever a child learns at one time or another, they remember for the rest of their days. Kar (česar) se Janezek nauči (od svojih staršev), to (bo) Janez (Janže) zna(l) (celo življenje) / What Janezek learns (from his parents), Janez (Janže) (will) know (his whole life); Kar se ne uči Janezek, tudi Janez ne bo znal / That which Janezek doesn't learn, even Janez won't know; Kar se naučiš v mladosti, to obvladaš v starosti / What you learn in your youth, you master in old age. Upbringing, 'good' or 'bad' influenced a child's future. It defined the adult's character and actions: *Kakor se mlado drevesce nagne, tako bo raslo staro drevo / As a young tree slants, the old one will grow;*²¹ *Česar se mlado dete privadi, tega se starček odvadil ne bo / What a young child gets accustomed to, an old man will not break from; Kakor otroka navadiš, tak(i) ostane / As you get your child accustomed, such they will stay; Kdor že kot otrok izgubi dušo, jo bo v starosti težko našel / The one who loses their soul as a child, will have a hard time finding it in old age; Če se otrok dela nauči, bo odrasel imel manj skrbi / If a child learns to work, they will have less worries as an adult. Our ancestors believed that 'good' upbringing (resulting in work habits) will avert poverty in the future: *Dobra vzgoja brani uboštvo pod streho / Good upbringing keeps poverty from under one's roof*, while a 'bad' one doesn't: *Kriva vzgoja otrok je slabih časov mati / Wrong upbringing of children is the mother of difficult times.***

The educational ideal pictured a descendant who would surpass their parent both economically and socially: *Človek želi, da je od vsakega boljši, a od sina, da je gorji / A man want to be better than everyone, but worse than his son*, to prevent the following sequence of events, among other things: *Starši nabirajo, otroci tratijo, vnuki (pa) stradajo? Parents gather, children spend, (and) grandchildren starve; Kar sezidajo skrbni roditelji, razden(ej)o slabi otroci / That which caring parents build, bad children pull apart*, and create a more fitting succession: *Oče polje skrči, sin ga orje, vnuk z*

²¹ Frequent comparisons of a child to a tree are hardly surprising from a psychoanalytic point of view; a tree is a phallic symbol and consequently symbolises a child in the equation 'feces-child-penis' in the scope of a child's sexual fantasies. Sigmund Freud realised 'that these elements are often dealt with in the subconscious mind as equivalent and could easily act as substitutes' (Freud 2006: 99).

njim bogati / The father clears a field, the son ploughs it, the grandson gets rid of it. Upbringing was seen as more distinguished and demanding than marriage (Imeti ženo je človeška zadeva, vzgajati otroke pa je božje delo / Having a wife is a human affair, bringing up children is god's work), as an adult must have the ability to learn from experience and plan: Starši v skrbi za otroke nazaj (v preteklost) gledajo, pa naprej (v bodočnost) vidijo / Worrying about their children, parents look back (into the past) and see forward (into the future). For instance, Mati kregaje hči nevesto svari / A mother scolding her daughter [in foresight] is warning the bride. The educational goal was embodied by a prudent and at the same time obedient and subordinate child: Mirnega otroka ima še mačeha rada / Even a stepmother loves a calm child;²² Čedne otroke Bog da, pridne morajo starši izrediti / Handsome children are given by God, good ones have to be reared by parents, therefore Najhujša staršem bolečina, če mora sram jih biti sina / Parents' worst pain is having to be ashamed of their son.

There were few educational instructions regarding care for the physical needs of a child: *Kaša otročja paša, leča pa otroke meče / Porridge children's pasture, and lentils toss children; Vino je za otročnico ne pa za otroka / Wine is for the new mother, not for the child; Mladega otroka pa kruh moraš zmirom greti / A young child and bread must always be warmed; Kdor otroku zaklepa kruh, ga navaja h kraji / The one who locks bread away from a child is teaching them to steal. A peasant child would be relatively fed, clean and dressed. As summarised by Janja Žagar, a rural child was "surrounded" by family, never alone, yet no-one paid attention to them, at least not in the sense of holding, cuddling and playing' (Žagar 1997: 10), which led to frequent injuries (burns, cuts, snake bites, near-drownings, near-suffocations, etc.) It seems that the multi-generational tradition focuses heavily on parental supervision: *Dober nadzor je pol vzgoje / Good control is half of upbringing; Dober oče ima oči tudi na hrbtu / A good father also has eyes on his back; Na otroke in na denar je treba dobro paziti / Children and money should always be carefully guarded; Kadar je kal poln kač, ne puščaj dece blizu / When the pond is full of snakes, don't let the children near it; Vžigalic in škarij otroke varij! / Guard children from matches and scissors!; Na dve stvari (reči) posebno pazi: na otroka pa na ogenj / Take special care of two things: a child and fire; Otroci in norci (ogenj) so potrebni varuha / Children and fools (fire) need a guardian; Otroci in ogenj niso nikoli prevarovani / Children and fire are never guarded too much. Because: Ogenj se zatrosi po otrocih, po norcih in pa po nesreči / Fire is scattered on children, on fools and by accident. And once a child got burnt, the fear of fire became internalised (through experience): Opečen otrok ogenj upije / A burnt child cries fire.**

²² Therefore: an obedient child can be loved by people, even if they are not blood related, a highly valued condition in society of the time. In folklore, a stepmother is a wicked mother. For more on this, see: Huzjan 2008: 119–120.

Accidents happened more frequently when children were left without adult supervision: *Kadar staršev ni doma, imajo otroci semenj / When parents are away from home, children have a fair.* Warmer temperatures provided opportunities for unsupervised play: *Poletje je čas, ko otroci zapirajo vrata, ki so jih vso zimo puščali odprta na stežaj / Summer is a time when children close doors they left wide open all winter;* *Kakor hitro je kopno, je otrokom v hiši tesno / As soon as the land thaws, children feel cramped inside the house.* The metaphorical line separating inside from outside, was a fence: *Otroci imajo za plotom raj / There's paradise for children outside the fence.* Outside the 'fence' temporarily erased orders, cold looks and sticks; outside, experience was the teacher: *Otroci že dobijo za plotom rajdo / Children are sure to find company outside the fence;* *Otroci imajo za plotom rihtarja / Children have their arbiter outside the fence;* *Najboljši sodnik otrokov je za plotom / The best judge of a child is outside the fence;* *Otročji sodnik je za plotom / A children's judge is outside the fence;* *Deca (otroci) imajo za plotom sodnika / Children have a judge outside the fence.*

Educational instructions admonished poor care of children: *Ljudje z večjo skrbjo gojijo zeljnate glave kot pa svoje otroke / People take greater care when growing cabbage than they do rearing children* and criticised excessive care: *Preveč skrbi uniči otroka / Too much care ruins the child;* *Če se otroku vedno ugodi in ščenetu kar naprej streže, iz nobenega ne bo kaj prida / If a child is always complied with and a mongrel is always catered to, nothing much will become of either;* *Otroku na voljo, sebi v (žalost in) nevoljo / At a child's disposal, to one's own (sorrow and) annoyance.* On one hand they stressed a loving attitude towards children: *Otrok je kakor sekira, če (te) rani, jo vseeno daš na ramo / A child is like an axe, if it wounds (you), you still put it on your shoulder;* *Ljubi svoje otroke kakor Bog angele / Love your children like God loves angels;* *Otrok najbolj potrebuje ljubezen takrat, ko jo najmanj zasluži / A child needs love most when they deserve it least.* Parental love was expected to lead to prosocial behaviour, which in turn should earn the child popularity, pet names: *Otrok, ki je ljubljen, ima veliko imen / A child who is loved has many names.* On the other hand, the proverbial tradition criticised expressions of parental love: *Starši, ki preveč božajo otroke, jih ne vzgajajo h krepostim / Parents who caress their children excessively don't teach them virtues;* *Kdor otroka(e) preveč ljubi, (ta) ga (jih) pogubi / The one who loves a child (children) too much, ruins them;* *Kdor ima svoje otroke rad, (naj) jih (nikar) preveč ne hvali / The one who loves their children, (should never) praise them too much.*

As stated above, the educational goal embodied a subordinate child, one who would fear unquestionable adult authority: *Če starejši govori – otrok naj molči / If an elder is talking - a child should remain silent;* *Ko stari govore, otroci naj molče / When old people talk, children should remain silent,* because *Vsaka zjamrana beseda otroka, je udarec na glavo matere / Every whiny word from a child is a blow to the mother's head.* With the use of such instructions a child soon realised their place in the patriarchal structure of the world, where social power was distributed hierarchically and where

violence functioned as defence from losing one's own social position. This gave rise to centuries of authoritarian education, aimed at securing submission: *Dober otrok vboga na skok / A good child obeys instantly; Največja sila je moč vzgojene volje / The greatest force is the force of a well-brought up will*, and disciplinary educational interventions: commands, scaring, intimidating looks, isolation, kneeling, denying food and beatings (cf. Puhar 1982; Huzjan 2021). Folk tradition says: *Pogled očeta je zakon za otroke / A father's look is the law to children; Mrzel pogled otroka v hipu ohladi / A cold look instantly cools a child down; Očetovi opomini zaležejo več kakor materine batine / Reprimands from a father are more effective than beatings by a mother* (cf. Govekar 1926: 92; Kremenšek 1970: 33; Brumen 1995: 143; Dolničar 2014: 6; Huzjan 2021: 59); *Naj toraj šiba otroke modro strahuje, potem se mladenču palice bati treba ne bo in tudi moža ptuje vrata po petah tolkle ne bojo / So let a birch intimidate children wisely, so a young man won't have to fear a rod and a man won't have foreign doors beat him on the heels; Kder je pri otrokih šibe premalo, bo palice preveč / Where there is not enough birch with children, there will be too much rod; Človek, ki ga bol ne vzgaja, ostane vedno dete / A man not educated by pain will forever remain a child* (cf. Puhar 1982; Dolničar 2014: 9; Turk Niskač 2016: 152; Huzjan 2021: 59–60). A beating was an expression of kindness and love: *Kdor ne kaznuje svojih otrok, jim ne želi dobro / The one who doesn't punish their children doesn't wish them well; Kdor prizanaša šibi, sovraži svojega sina / The one who is easy on the birch hates their son; Komur se šibe škoda zdi, sovraži svojega sina, kdor ga pa ljubi, ta ga vedno strahuje / The one who uses the birch sparingly hates their son, but the one who loves him always intimidates him*. However, punishment should be doled out appropriately: *Šiba naj bo zdravilo, ne vsakdanje kosilo za otroke / For children, a birch should be medicine, not everyday lunch; Ljubi otroka s sarcam in vzgajaj z roko / Love a child with your heart and raise them with your hand; Če z eno roko tepeš otroka, ga z drugo roko stiskaj k srcu / If you beat a child with one hand, hold them close to your heart with the other*. Punishing a child was the exclusive right of the one who supported the child: *Pravico tepsti otroka ima samo tisti, ki mu kruh reže / Only the one slicing a child's bread has the right to beat them*.

5 Tako, kot ti s starši, tako otroci s tabo (As you do to parents, children do to you)

As mentioned before, children learn through experience, and, from birth, through observation, as well: *Otroci so ogledalo staršev / Children are a mirror of their parents*. Adults must have known this even before the first millennium. Folklorist Monika Krupej Telban namely discovered an articulation of a proverb dating back to the Enlightenment in a folk tale titled *Nehvaležni sin (The Ungrateful Son, ATU 980)*: 'Spoštuj starše, da tebi dobro pojde' ('Respect parents, so you will have it good'). This tale or versions

of it are known to be around prior to the first millennium (Kropej Telban 2022: 362). Hidden in the said proverb lies the most vital educational component, that of a role model: *Kakor ti spoštuješ svoje starše, tako bodo tebe spoštovali(a) tvoji otroci (deca tvoja) / As you respect your parents, so your children will respect you; Kakor ti s svojimi starši, tako bodo s teboj tvoji otroci / As you do with your parents, so your children will do with you; Če svojih staršev nisi enkrat ubogal, tebe tvoj sin ne bo desetkrat / If you disobeyed your parents once, your son will disobey you ten times.*

A child (subconsciously) identifies with the role model and repeats actions as an adult, which is why we say that patterns are passed on trans generationally: *Otroci morajo tkati s tem, kar so napredli starši / Children have to weave with what parents spun.* Behavioural patterns are repeated from generation to generation without analysis or deliberation: *Otrok dela, kar očeta delati vidi / A child does what they see their father doing; Dobri starši – dobri otroci / Good parents - good children; Če so starši dobri, so tudi otroci, če so starši slabi, so tudi otroci / If parents are good, the children are also, if parents are bad, the children are also; Kakor mati prede, tako hči tke / As the mother spins the daughter weaves; Napake otrok niso od kdovekod - imajo jih od svojih staršev / Children's flaws don't come from nowhere - they get them from their parents; Kar zleti staršem z jezika, prileti na jezik otroku / Whatever leaves the parents' tongues, lands on the tongue of the child; Žalost staršev zaradi otrok zmeraj na otroke pade / Parents' sorrow around children always falls on children; Kakoršna mati, taka hči (taki otroci), kakoršen oča, taki sin; / Like mother like daughter (like children), like father, like son; Kakšno deblo, takšen klin, kakšen oče, takšen sin; / Like tree trunk, like peg, like father, like son; Kakršna mati, takšen sin / Like mother, like son; Kakršna matka, taka Katka / Like mother, like Katka; Kakoršni stariši, takošni (takovi, taki tudi) otroci / Like parents, like (also) children; Kakršna njiva, takšen pridelek / Like field, like harvest; Poštena mati, poštena hči / Honest mother, honest daughter; Oče sedlar, sin sedlar / Father saddler, son saddler; Če je mati spaka, rada je hči tudi taka / If a mother is a scowl, her daughter tends to be the same; Če je oče ves ušiv, zakaj bi pa še sin ne biv / If the father is lice-ridden, why wouldn't the son be, as well; Če laže mati, bo lagala tudi hči / If the mother lies, the daughter will lie, as well; Če noseča žena kaj vzame, njen otrok krade / If a pregnant woman takes something, her child steals; Kakor poje mati, tako hči / The way the mother sings, the daughter does; Kjer je mati troša, so hčerke potrate / Where the mother spends, the daughters waste ; Glej mater in se poroči s hčerko / Look at the mother and marry the daughter; Oče in sin enacega škvarta / Father and son of the same character. *Otroci se radi po starih zvržejo / Children tend to take after their parents, so: Po otrocih se vse izve / Everything is found out through children.**

Behaviour patterns can become amplified in the next generation: *Če oče zahaja v gostilno, otroci zahajajo v beznico / If a father frequents a pub, the children frequent a shebeen; Ako (če) mož ženo zmerja, jo otroci tepejo (tolčejo) / If a husband insults his wife, the children beat her; Če je oče junak, bode sin gotovo vojščak (vojak) / If the*

father is a hero, the son is sure to become a warrior (soldier), although Čestokrat je očetova suknja sinu prevelika / The father's coat is often too big for the son. They can also oppose the role model; such as Vsaka stvar ima dve plati / Everything has two sides, yet it is the same 'thing' : Če pije rad vino otrok, ne bo maral zanj v starosti / If a child likes to drink wine, they won't care for it in old age; Hči skopuhova je dostikrat žena požeruhova / Daughter of a miser is often wife of a glutton; Lena mati, hitra hči / Lazy mother, quick daughter; Oče skopuh, sin goljuf (razsipnež) / Father miser, son swindler (wasteful); Oče zbira slamo kot zlato, sin razsipa zlato kot slamo / The father collects straw as if it was gold, the son wastes gold as if it was straw; Oče štedljivec, sin zapravljivec / Father a saver, son a squanderer; Kar je oče zbral z motiko, to sin razmetava z lopato / What the father gathered with a hoe, the son is squandering with a shovel.

The realisation of literary tradition is quite on point: *Grehi staršev se maščujejo nad otroci / Parents' sins take revenge on the children ; Grehe staršev plačujejo otroci / Parents' sins are paid for by children; Na otrocih se kaznujejo grehi očetov / Fathers' sins are punished on the children; Če starši greše, otroci trpe / If parents sin, children suffer; Prerado ternje otroke pika, ki so ga nepošteni stariši sadili / All too often children get pricked by thorns, which were planted by dishonest parents. Summed up in one proverb: Greh se prenaša na potomce / A sin is passed on to descendants, although Otroci niso dolžni poravnati grehov svojih staršev / Children are not obligated to settle their parents' sins. This very situation could give rise to the development which worried parents so much; namely the possibility of a child arbitrarily writing off the alleged debt (cf. section *Otroci so večni jarem / Children are an eternal yoke*). Let's revisit the saying 'Spoštuj starše, da tebi dobro pojde' ('Respect parents, so you will have it good'), which verbalises the folklore tale *Nehvaležni sin / The Ungrateful Son* (ATU 980). The missing text reads: *Respect your parents, since your children look up to you. This is why you will have it good, once you become old and helpless, because your children will respect you, as well, and God will reward them: Otroku, če stariše spoštuje, vse dobro Bog podaruje / God gifts all good things to a child who respects their parents.*²³ Caring for elderly family members at home was a sign of respect and a woman's job (*Če so otroci dobri, očeta k možu postavijo; če so pa hudobni, ga na beraško palico pripravijo / If children are good, they make the father a man; if they are wicked, they make him a pauper*), in times when there was no such thing as a network of care facilities.²⁴*

²³ The fourth commandment is about more than just the relationship between children and parents, it is about respect for any authority one is subordinate to.

²⁴ Ethnologist Duška Knežević Hočevar found that 'While either older family members helped care for children, or they were left to care for one another, a woman was expected to provide long-term care to family members in need of help, despite her full engagement at the farm' (Knežević Hočevar 2013: 110). The question of why care for the elderly primarily fell to women, warrants additional analysis.

The so-called writing off debt reveals the precise motif of the ungrateful child: *Oče je dal sinu vinograd, sin pa očetu niti grozda / The father gave his son a vineyard, the son didn't even give his father a grape; Oče je dal sinu grozd, sin očetu niti truge / The father gave his son a grape, the son didn't even give his father a coffin; Težko njemu (onemu otcu), ki se na deco opira / Woe betide him (the father), who relies on his children ; Za devet otrok je prostora na rokah očeta, a za očeta ni prostora v hišah devetih otrok / There's room for nine children on a father's arms, but there's no room for the father in nine children's houses; Ko otroci sklenejo, da bodo skrbeli za starše, teh že ni več med živimi / When children decide to take care of their parents, the parents are no longer among the living.*

It seems the fear of helplessness and abandonment in old age permeated the entire upbringing phase, as *Najstrožji sodniki so učitelju učenci, staršem otroci / The strictest judges for a teacher are students, for parents, children.* With age, parents were forfeiting social power and regressed down the ladder of patriarchal hierarchy. With no organised social services to take care of them, they were left at the mercy of their descendants, which were now in a place where they could realise their vindictive childhood fantasies: *Ne tepi matere svojih otrok, tvoj otrok nekoč te kaznuje, ne Bog / Don't beat the mother of your children, the child, not Got, will punish you one day; Hudobni otroci so žebliji v krsto svojih staršev / Wicked children are nails in their parents' coffin; Hudobija otrok se na starše vračuje / The children's wickedness comes back to the parents.* This is corroborated by Janez Trdina's observations: 'Even children don't willingly get beaten, a grown young man will often take revenge on the one, sometimes even on the parents (by refusing to help), if he was beaten as a boy. Vengeance can turn bloody, it is said: 'Just you wait, I will pay you back when I grow up' (Trdina 1987: 457). A proverb warning parents, in different dialectal and other versions, about their children's vindictive attitude once they grow old, attracts extra attention: *Če vlečeš (hočeš) očeta do praga, te bod(j)o tvoji otroci (otročā) čez prag sunili (sunilā) / If you drag (will) your father to the threshold, your children will push you over the threshold; Ti očeta do praga, sin (otroci) tebe čez prag / You your father to the threshold, the son (children) you over the threshold; Ti očo do praga, sin tebe prejk praga / You father to the threshold, son you over the threshold; Ti ačeta za praga, sin tebe č'z prag / You father to the threshold, son you over the threshold; Kdor očeta do praga privleče, nja bojo otroci čez prag sunili / The one who drags his father to the threshold, he will be pushed over the threshold by his children; Kdor svojega očeta ali mater do praga (potisne) privleče, njega bodo njegovi otroci čez prag sunili (vergli!, potisnili) / The one who (pushes) pulls their father or their mother to the threshold, they will be shoved (thrown!, pushed) over the threshold by their children; Ti starše do praga, otroci tebe čez prag / You your parents to the threshold, children you over the threshold; Kdor stariše svoje do praga pervleče, tiga bodo njegovi otroci čres prag izsuvali / The one who drags their parents to the threshold, they will be hurled over the threshold by*

*their children; Ti sina do praga, sin tebe čez prag / You your father to the threshold, the son you over the threshold.*²⁵

Considering the described cases, the main goal of upbringing from the parent's point of view was to 'create' a diligent person, fearful of authority (father – God), seemed quite sensible. Disciplining in the fear of helplessness in old age did yield results, but not on the long term, because the patriarchal machinery turned as the parents aged.²⁶ Warnings about more democratic relations between parents and children: *Mehkužnost rodi kilave otroke / Faint-heartedness gives birth to crummy children; Kdor otroka ne pouči, ga pogubi / The one who doesn't teach the child, ruins them; Kolikor bolj popuščamo, toliko bolj zahtevni bodo naši otroci / The more we ease off, the more difficult our children will be; Popustljivost staršev zagreni starost, otrokom pa nakoplje bridkost / Parental lenience makes old age bitter, and brings anguish to children*, which could threaten the patriarchal order *Otroka uči, ako ne, bo otrok tebe / Teach your child, if not, the child will teach you; Gorši [boljše] je, da sin očetu streže kakor oče sinu / It is better for the son to serve his father than for a father to serve his son*, have been preserved to this day, albeit in different educational forms, starting with: A child mustn't be indulged, because ...

Fear of unsuitable transfer of cultural norms or ungratefulness (*Nehvaležnost je hčerka ošabnosti / Ungratefulness is the daughter of arrogance*) is revealed as a fear of abandonment, further emphasized by (self)persuasion on part of the adults: *Ni na svetu boljših prijateljev, nego so starši otrokom / There are no better friends in the whole world than parents are to children*, noting that *Starši bolj ljubijo svoje otroke kakor otroci starše, / Parents love their children more than the children love their parents*, as they have perfected the way they love and care for them: *Mati lahko zredi sedem otrok, sedem otrok ne more rediti mame / A mother can raise seven children, seven children can't feed the mother;*²⁷ *Ena mati lahko preživi deset (devet, dvanajst, dvajset) otrok, deset (devet, dvanajst, dvajset) otrok pa ne more ene matere / One mother can support ten (nine, twelve, twenty) children, but ten (nine, twelve, twenty) children can't support one mother; Prej mati preravna deset otrok kot deset otrok eno mater / Sooner a mother cares for ten children than ten children for one mother; Ena mati lahko mnogo otrok vzgoji, vsi otroci pa ene matere ne morejo / One mother can raise many children, but all the children can't raise one mother; Oče bolje skrbi za*

²⁵ The following sentence can also be found in ethnological research of Minot, a French village: 'In the same way parents abused their children, the children never missed a chance to abuse their aging parents' (Zonabend 1993: 138). Fear of revenge could be seen as a principle of agricultural economies in the past centuries.

²⁶ In other words: 'In a family environment, life's transitions of the younger generation are blended with life's transitions of the older generation. The timing of, for instance, a grown child's move away from home and marriage in early adulthood is associated with the timing of older generation's transition to retirement or a child coming into their inheritance. This exposes the interdependence between generations and their operating strategies in mutual interactions throughout life' (Knežević Hočevar 2013: 25).

²⁷ What the proverb in fact says is that a man cannot excell (be successful) at many fields: *Mojster enega poklica hrani ženo in sedmero otrok, mojster sedmih poklicev niti sebe ne more / A master of one profession feeds a wife and seven children, a master of seven professions can't even feed himself.*

deset otrok, kakor deset otrok za enega očeta / A father takes better care of ten children than ten children do of one father; Deset otrok ne more rediti očeta, oče pa vse / Ten children can't feed the father, but the father can feed all of them; Starši lahko preživijo deset otrok, deset otrok pa starše le težko zmore / Parents can support ten children, but ten children can hardly support their parents; Starši lahko vzgojijo ducat otrok, toda ducat otrok ne more dostojno spraviti pod rušo svojih staršev / Parents can raise a dozen children, but a dozen children can't bury their parents properly.

This is the very reason why, as a rule, an old farmer didn't trust his son – the young farmer. Elderly parents generally stayed on the farm which they handed over to their descendant: *Očetu luč, sinu ključ / [Heavenly] light for the father, a key for the son; Bogu di(u)šico, sinu možnjico (zemljico) / A soul to God, a purse (land) to the son*, and were entitled to maintenance. Legal historian Sergij Vilfan writes that 'maintenance often led to arguments between a former and new owner and was one of the most common subject-matter of civil lawsuits [...], but handing over the key was often unavoidable' (Vilfan 1961: 509), in order to ensure the continuation of management. Family farms followed patrilineal inheritance, and considering the then current patrilocality, elderly parents wished: *Boljše je umirati na sinovem pragu kakor na hčerini postelji / It is better to die on your son's threshold than on your daughter's bed*, meaning in one's own house. Complications with transfer of a farm could arise when there was something to inherit, hence the proverb: *Kadar ni dediščine, se otroci na grobu svojih staršev ljubijo, če je je pa kaj, se pa grdo gledajo / When there is no inheritance, children love each other on their parents' grave, when there is some, they glare at each other*, had a somewhat comforting effect.²⁸

6 Conclusion

There is no signifier on Clarin.si for the stage of life we now call 'otročstvo' (childhood). Neither is 'detinstvo' (babyhood), which first appeared in bourgeois culture. Signifiers 'dojenček' or 'dojenec' (baby), 'novorojenček' or 'novorojenec' (newborn), 'malček' (toddler), 'mladostnik' (adolescent) and 'najstnik' (teenager) are also not to be found. The latter is not surprising, as differentiation of childhood came about relatively recently, at the beginning of the 20th century in the scope of consumer clothing culture for children (Cook 2004). Truth be told, the word 'zibel' (cradle) in fact signifies a new-born,²⁹ but only appears three times in the entire collection of proverbs, and only once as relevant

²⁸ For more on agricultural economy prior to the second world war, see, i.a. Makarovič (1978), Knežević Hočevar, Černič Istenič (2010), Lazarevič (2022).

²⁹ This (crib) was actually a metonymy (for a newborn) – replacement of the content with the vessel. New-borns mostly lay swaddled in a crib, which oftentimes had a hole in the bottom to allow the runoff of urine (Šarf 1980: 14).

for this research: *Krsta je sestra zibeli / A coffin is a cradle's sister*. 'Najstnik' (teenager) makes a single appearance, as well: *Najstnik sicer ne sprašuje več, odkod je prišel, pa tudi ne pove, kam gre / While a teenager no longer asks where they have come from, they also don't say where they're going*. Both records are taken from a manuscript by Marjan Podlesnik, the year of the records is unknown. Based on countless findings I can't conclude that for centuries, the majority population in Slovenia wasn't aware of the stage of life directly preceding adulthood, but I can assert that this was a short stage, one that had to be overcome as soon as possible, due to the inability of children to work.

While peasant economy didn't know childhood as we do today, they had ideas about children inhabiting the not-yet-adult stage. A child was a subject to be cared for and raised, which was strenuous for the majority of agricultural holdings. The said subject was incapable of working as they were weak, inexperienced, irrational and at the same time restless and headstrong. Divine in their promise of a future and devilish because they called for sacrifices on part of adults. While their romantic image was innocent and respectable, the enlightenment side showed them as wild and in need of discipline. They had to be brought up as obedient and self-sufficient worker, who will simultaneously care for their parents as they approach old age. Authoritarian upbringing consisted of violent disciplinary measures.

As mentioned above, the weakness of small children in rural areas was their inability to work and contribute to family economy. They presented an expense for self-sacrificing parents, which forced them to a lower status. The longer they remained not working, the larger the incurred 'debt'. They only acquired worth with physical strength and experience, when they were able to produce (for) food on their own. From then on they started to pay back the debt, which was repaid in full through caring for their parents once they reached old age. 'Indebting' was more than just a survival strategy on part of the parents, it also regulated relations in patriarchal rural families. Like women, the elderly, poor, beggars, fools, drunkards and animals, children were marginal members of such patriarchal communities. With authoritarian upbringing, aimed at instilling obedience to the father (authority), a child soon realised their place in the patriarchal structure of the world. If a first-born son came into a wealthier farm, he could hope for more social power, but the patriarchal machinery flipped as he entered old age and became helpless. At that point, his survival depended on successful transfer of cultural norms. The ability to work and be independent were valued as they enabled survival in an unjust and persistent patriarchy.

Quite similarly, Özay Karadağ analysed 785 Turkish proverbs with childhood subject matter and found that the socialisation participants exhibited ambivalence toward the child. On one hand, the child was presented as a precious being, while on the other adults found their characteristics were disruptive (Karadağ 2013). A similar cultural image is painted by 42 (probably) Asante³⁰ proverbs collected by Robert S. Rattray in 1933.

³⁰ The Asante people live in the south of Ghana, in Togo and Ivory Coast.

The analysis of proverbs coming from the Igbo ethnic group from Nigeria, which might still serve as a socialisational metaphorical tool for today's children, reveals that their everyday use maintains social and cultural order in the community (Penfield and Duru 1988). Semantics of these proverbs point to similar socialisational tendencies to those found in the collection of Slovenian proverbs. However, no serious claims can of course be made without an in-depth comparative analysis.

But I can certainly conclude that proverbs regarding children from the Clarin.si collection use authoritarian and authoritative language in regard to different authorities: god, father, mother, parents, etc. They served as a socialisational 'manual' in order to maintain a social and cultural *status quo*, where taking care of a child equated with taking care of their own old age. Authoritativeness and authoritarianism of proverbs revealed parental fear of their children's disobedience, unsuccessful transfer of the system of values and, with that, of their children's arbitrary remission of 'debt' once they grow up. In this, the patriarchal upbringing revealed its soft spot – impotence.

Acknowledgements

I would like to thank Dr. Saša Babič for her assistance with interpreting the Clarin.si collection, proverbs and ethno-linguistic conundrums, as well as for her most valuable comments, and Rok Mrvič for bringing Janez Trdina to my attention. This article was written in the scope of the *Traditional paremiological units in dialogue with contemporary use* (ARIS J6-2579) research project, *Language, culture and values: economic images of everyday life in folklore forms* (ARIS J6-50197) research project, and the *Ethnological, anthropological and folkloristic research of everyday life* (ARRS P6-0088) programme.

References

- Abrahams, Roger D., 1975: A Sociolinguistic Approach to Proverbs. *Midwestern Journal of Language and Folklore* 1/1, 60–64.
- Babič, Saša, 2015: *Beseda ni konj: estetska struktura slovenskih folklornih obrazcev*. Ljubljana: Založba ZRC, ZRC SAZU (Ethnologica – Dissertationes, 6). DOI: <https://doi.org/10.3986/9789612547660>.
- Babič, Saša, 2018: Tudi stara pamet ne vidi vsega vselej prav: konceptualizacija pameti in neumnosti v paremioloških enotah iz zbirke Inštituta za slovensko narodopisje ZRC SAZU. In: Będkowska-Kopczyk, Agnieszka; Pfandl, Heinrich (eds.), *Slavofraz 2016: Phraseologie und (naive) Psychologie = Phraseology and (naïve) psychology = Frazologija i (naivnaja) psihologija*. Hamburg: Dr. Kovač, 81–92.
- Babič, Saša, 2021: Stereotipna podoba ženske v slovenskih folklornih obrazcih. In: Žbogar, Alenka (ed.), *Ustvarjalke v slovenskem jeziku, literaturi in kulturi: 57. seminar slovenskega jezika, literature in kulture: 5.–16. 7. 2021*. Ljubljana: Znanstvena založba Filozofske fakultete, 80–88. DOI: <http://dx.doi.org/10.4312/SSJLK.57.80-88>.
- Bojc, Etbin, 1974: *Pregovori in reki na Slovenskem*. Ljubljana: DZS.
- Brumen, Borut, 1995: *Na robu zgodovine in spomina: urbana kultura Murske Sobote med letoma 1919 in 1941*. Murska Sobota: Pomurska založba.

- Burcar, Lilijana, 2007: *Novi val nedolžnosti v otroški literaturi. Kaj sporočata Harry Potter in Lyra Srebrnosta?* Ljubljana: Sophia (Sodobna družba, 19).
- Cook, Daniel Thomas, 2004: *The comodification of childhood. The children's clothing industry and the rise of the child consumer.* London: Duke University Press.
- Cunningham, Hugh, 1996: The history of childhood. In: Hwang, Philip C.; Lamb, Michael E.; Sigel, Irving E. (eds.), *Images of childhood.* Mahwah, NY: Lawrence Erlbaum Associates, 27–35.
- Dolničar, Alojz, 2014: *Spomini na otroška in partizanska leta. Šmarje pri Jelšah:* [A. Dolničar].
- Dolník, Juraj, 2007: *Lexikológia.* Bratislava: UK Bratislava.
- Fidowicz, Alicja, 2021: Motiv neumnega otroka (tepčka), posebne potrebe in etika v slovenskih in poljskih ljudskih pravljicah – primerjalna perspektiva. In: Jožef Beg, Jožica (ur.), *Večnaslovniška odprtost pravljic: mednarodni znanstveni simpozij: Novo mesto, 17. september 2020: zbornik prispevkov.* Novo mesto: Slavistično društvo Dolenjske in Bele krajine, 39–49.
- Freud, Sigmund, 1977: *Predavanja za uvod v psihoanalizo.* Ljubljana: DZS.
- Freud, Sigmund, 1987: Nezavedno. In: *Metapsihološki spisi.* Ljubljana: ŠKUC, Znanstveni inštitut Filozofske fakultete, 119–176.
- Freud, Sigmund, 2006: O pretvorbah nagonov, posebej pri analni erotiki. In: *Spisi o seksualnosti.* Ljubljana: Društvo za teoretsko psihoanalizo, 97–104.
- Govekar, Fran, 1926: *Bon ton. Knjiga o lepem vedenju, govorjenju in oblačenju v zasebnem in javnem življenju.* Ljubljana: Tiskovna zadruga.
- Grzybek, Peter, 2008: Fundamentals of Slovenian Paremiology. *Traditiones* 37/1, 23–46. DOI: <https://doi.org/10.3986/Traditio2008370102>.
- Grzybek, Peter, 2012: Proverb Variants and Variations: A New Old Problem? In: Lauhakangas, Outi; Soares, Rui J. B. (eds.), *Proceedings of the Fifth Interdisciplinary Colloquium on Proverbs.* Tavira: IAP, 136–152.
- Gutsman, Ožbalt, 1777: *Windifche Sprachlehre.* Klagenfurt: gedruckt bey Ignaz Aloys Kleinmayer. Online: <http://www.dlib.si/details/URN:NBN:SI:doc-3J12UH GK> (11.8.2023).
- Huzjan, Vanja, 2008: Podoba otroka v slovenski ljudski pripovedni pesmi: psihoanalitične vinjete. *Traditiones* 37/1, 113–133.
- Huzjan, Vanja, 2021: Vzgoja mlajših otrok v Ljubljani in okolici v prvi polovici 20. stoletja do druge svetovne vojne. *Glasnik Slovenskega etnološkega društva* 61/1, 54–64. Dostopno na: <https://www.dlib.si/details/URN:NBN:SI:DOC-HE127DLB/?query=%27keywords%3D-Slovensko+etnolo%C5%A1ko+dru%C5%A1tvo%27&pageSize=25&fUDC=Etnologija&fyear=2021> (11.8.2023).
- Huzjan, Vanja, 2022: Slovenian Folk Lullabies: Analysis of the Lullaby Texts and Their Functions. *Folklore* 86, 181–186. DOI: <https://doi.org/10.7592/FEJF2022.86.huzjan>.
- Huzjan, Vanja, 2023: *Materialni svet otroštva.* Ljubljana: Založba ZRC, ZRC SAZU (Ethnologica – Dissertationes, 9). DOI: <https://doi.org/10.3986/9789610507314>.
- Jason, Heda, 1987: Proverbs in Society: The Problem of Meaning and Function. *Proverbium* 16–25, 617–624.
- Karadağ, Özyay, 2013: Child and childhood in Turkey. Turkish Proverbs. *Milli Folklor* 98. Online: <http://acikerisim.duzce.edu.tr/xmlui/handle/20.500.12684/3051> (11.8.2023).
- Keber, Janez, 2015: *Slovar slovenskih frazemov.* Online: <https://fran.si/192> (11.8.2023).
- Kobcek, Fran, 1887: *Pregovori, prilike in reki.* Ljubljana: A. Trstenjak. Online: <http://www.dlib.si/details/URN:NBN:SI:doc-HGXSZ57J> (11.8.2023).

- Kocbek, Fran; Šašel, Ivan, 1934: *Slovenski pregovori, reki in prilike*. Celje: Družba sv. Mohorja (Mohorjeva knjižnica, 72).
- Krikmann, Arvo, 1985: Some Additional Aspects of Semantic Indefiniteness of Proverbs. *Proverbium* 2, 58–85.
- Krofič, Robi, 2011: Podobe otroštva. *Emzin* 21/1–2, 67–69.
- Knežević Hočevar, Duška, 2013: *Etnografija medgeneracijskih odnosov: dom in delo na kmetijah skozi življenjske pripovedi*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610503354>.
- Košir, Diana, 2023: Koncept telesne kazni ali 'šiba novo mašo poje« v izbranih verskih publikacijah 19. stoletja. In: Jež, Andraž (ed.), *Slovenska literatura in umetnost v družbenih kontekstih*. Ljubljana: Založba Univerze, 75–81 (Obdobja, 42).
- Kremenšek, Slavko, 1970: *Ljubljansko naselje Zelena jama kot etnološki problem*. Ljubljana: SAZU (Dela SAZU, Razred za zgodovinske in družbene vede, 16).
- Kropej Telban, Monika, 2022: Pregovori v slovenskih folklornih pripovedih na primeru treh pravljicnih tipov. *Slavistična revija* 70/3, 361–371. DOI: <https://doi.org/10.57589/srl.v70i3.4027>.
- Kržišnik, Erika, 2005: Frazeologija v luči kulture. In: Stabej, Marko (ed.), *Večkulturnost v slovenskem jeziku, literaturi in kulturi: zbornik predavanj / 41. seminar slovenskega jezika, literature in kulture, 27. 6.–15. 7. 2005, Ljubljana*. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete, 67–81.
- Lazarevič, Žarko, 2022: *Delo in zemlja: male študije kmečkega sveta*. Ljubljana: Inštitut za novejšo zgodovino (Razpoznavanja, 47).
- Lipavic Oštir, Alja; Vasiljevič, Anja; Koletnik, Mihaela, 2014: Zasedranost pregovorov v besedišču govorcev slovenskega jezika. *Slavia Centralis* 7/1, 48–72.
- Makarovič, Marija, 1975: *Pregovori – življenjske resnice*. Ljubljana: Kmečki glas.
- Makarovič, Marija, 1978: *Kmečko gospodarstvo na Slovenskem*. Ljubljana: Mladinska knjiga.
- Meterc, Matej, 2013: Antonimija enako motiviranih paremioloških enot (primeri iz slovenščine in slovaščine). *Slavistična revija* 61/2, 361–376.
- Meterc, Matej, 2021: Aktualna raba in pomenska določljivost 200 pregovorov in sorodnih paremioloških izrazov. *Jezikovni zapiski* 27/1, 45–51. DOI: <https://doi.org/10.3986/JZ.27.1.03>.
- Mieder, Wolfgang, 1993: *Proverbs are Never Out of Season: Popular Wisdom in the Modern Age*. New York, Oxford: Oxford University Press.
- Mieder, Wolfgang, 2004: *Proverbs: A Handbook*. Westport, Connecticut: Greenwood Press.
- Penfield, Joyce; Duru, Mary, 1988: Proverbs: Metaphors That Teach. *Anthropological Quarterly* 61/3, 119–128. DOI: <https://doi.org/10.2307/3317788>.
- Prek, Stanko, 1972: *Ljudska modrost trden je most: slovenski ljudski pregovori in reki*. Maribor: Obzorja.
- Puhar, Alenka, 1982: *Prvotno besedilo življenja. Oris zgodovine otroštva na Slovenskem v 19. stoletju*. Zagreb: Globus (Biblioteka Globus).
- Rattray, Robert Sutherland, 1933: The African child in proverb, folklore, and fact. *Africa* 6/4, 456–471.
- Slomšek, Anton Martin, 1855: *Veliko berilo in pogovorilo za vajo učencov drugega odreda*. Beč: C. k. prodavnica šolskih knjig. Online: <http://www.dlib.si/details/URN:NBN:SI:doc-5E6SHK2E> (11.8.2023).

- Šarf, Fanči, 1980: O raziskovanju stanovanjske kulture. In: Kremenšek, Slavko; Bogataj, Janez (eds.), *Poglavja iz metodike etnološkega raziskovanja – 1*. Ljubljana: Filozofska fakulteta, 5–15 (Knjižnica Glasnika SED, 4).
- Terseglav, Marko, 2002: Šalo na stran in šala kot kratka oblika ustnega slovstva. *Traditiones* 31/2, 55–83.
- Trdina, Janez, 1987: 9. Posnetki in opazke začeni 20. Septembra 1871. In: Štabi, Snežana; Kramberger, Igor (eds.): *Podobe prednikov: zapiski Janeza Trdine iz obdobja 1870–1879. 2. knjiga: 'Vsaka svinja naj si riže svoje korenje'*. Ljubljana: Univerzitetna konferenca ZSMS, 453–476 (Krt, 31).
- Turk Niskač, Barbara, 2016: *Med igro in delom. Etnografska raziskava učenja v zgodnjem otroštvu. Doktorsko delo*. Ljubljana: B. Turk Niskač.
- Turk Niskač, Barbara, 2021: *O igri in delu*. Ljubljana: Založba ZRC, ZRC SAZU (Ethnologica – Dissertationes, 7). DOI: <https://doi.org/10.3986/9789610505181>.
- Vilfan, Sergij, 1961: *Pravna zgodovina Slovencev*. Ljubljana: Slovenska matica.
- Zonabend, Françoise, 1993: *Dolgi spomin: časi in zgodovine v vasi*. Ljubljana: ŠKUC, Filozofska fakulteta (Studia humanitatis).





IV

Sodobna raba paremioloških
enot in digitalnih virov

Contemporary use of paremiological
units and digital sources



Dvajset najbolj poznanih slovenskih pregovorov in sorodnih paremioloških izrazov: podatki v paremiološkem slovarju SPP in zbirki Pregovori 1.1

The twenty best-known Slovenian proverbs and related paremiological expressions: Data in the SPP paremiological dictionary and the Proverbs 1.1 collection

V članku predstavljamo vrh slovenskega paremiološkega minimuma – prvih 20 paremij glede na poznavanje med 527 govorce slovenskega jezika. Vsaka od dvajsetih paremij je predstavljena s svojimi variantami v sodobni slovenščini, ki so bile potrjene v jezikovnem korpusu metaFida (približno 4,7 milijarde besed). Veliko število potrjenih variant v sodobni slovenščini primerjamo s stanjem v virih paremiološke zbirke Pregovori, ki vsebuje 37.390 zapisov slovenskih pregovorov in predstavlja hibridni medij med paremiološko zbirko in jezikovnim korpusom. Članek ponuja pilotni poskus paremiološke lematizacije zbirke, ki se zdi ključna, a precej zahtevna naloga v prihodnosti.

• **Ključne besede:** paremiologija, paremija, pregovor, paremiološki minimum, jezikovni korpus, paremiološka lema

In this article, we present the top of the Slovenian paremiological minimum – the first 20 paremias in terms of familiarity among 527 speakers of the Slovenian language. Each of the 20 paremias is presented with its variants in the modern Slovenian language, which have been confirmed in the metaFida language corpus (around 4.7 billion words). We compare the large number of confirmed variants in the modern Slovenian language with the situation in the sources of the Pregovori paremiological collection, which contains 37,390 records of Slovenian proverbs and represents a hybrid medium between a paremiological collection and a language corpus. The article offers a pilot attempt at a paremiological lemmatization of the collection, which seems to be a key, but rather demanding task in the future.

• **Keywords:** paremiology, paremia, proverb, paremiological minimum, linguistic corpus, paremiological lemma

1 Dvajset najbolj poznanih pregovorov in sorodnih paremij (vrh paremiološkega minimuma)

Idejo določanja paremiološkega minimuma – jedra najbolj poznanih pregovorov in sorodnih paremij v določenem jeziku – je predstavil Permjakov (1988: 150–151), ki je prvi skupini anketirancev prikazal seznam 1491 ruskih pregovorov z navodilom, naj izločijo neznane, drugi skupini (sto anketirancev) pa dal za nalogo, da dopolni manjkajočo (drugo) polovico pregovorov, ki so ostali na seznamu. Permjakov je s tem določil seznam 538 paremij, ki so bile znane več kot 90 % anketirancev, kot paremiološki minimum pa v objavah predstavil bodisi 500 bodisi 300 najbolj poznanih izrazov v ruščini. Od takrat pojem paremiološkega minimuma predstavlja enega izmed osrednjih

paremioloških konceptov, njegova določitev pa ena izmed pomembnejših nalog za paremiologijo posameznega jezika (Đurčo 2014: 200–201). Slovenski paremiološki minimum je bil uporabljen za izdelavo frazeodidaktičnega in paremididaktičnega razdelka Frida učnega e-okolja *Slovenščina na dlani* (Ulčnik, Meterc, Jesenšek 2021) ter ob izbiranju in slovaropisni obravnavi iztočnic za *Slovar pregovorov in sorodnih paremioloških izrazov* (v nadaljevanju: SPP) (Meterc 2019), hrvaški minimum (Varga 2022) pa na primer za zasnovo paremiološkega spletnega tezavra.

V anketni raziskavi za določitev slovenskega paremiološkega minimuma smo anketirancem prikazali 918 paremij v celoti po metodi, ki jo je uporabil tudi Đurčo (2002) za slovaški minimum. Anketiranci so odgovarjali ali izraz 1. poznajo in uporabljajo; 2. poznajo, ampak ne uporabljajo; 3. ne poznajo, a razumejo ali 4. ne poznajo in ne razumejo. Imeli so možnost navesti drugačno obliko paremije od prikazane, s katero smo zbrali precej podatkov o variantah (Meterc 2019: 42). Teoretična in metodološka zasnova slovenskega paremiološkega minimuma sta bili podrobno predstavljena v doktorski disertaciji (Meterc 2014) in iz nje izhajajoči monografiji (Meterc 2017), paremiološki minimum v celoti (300 paremij, urejenih glede na poznanost med 316 govorce) pa je bil vključen kot priloga k doktorski disertaciji. V njem so paremije, ki so bile v anketi poznane od 79,7 % do 99,7 % anketirancev (Meterc 2017: 68). V monografiji (Meterc 2017: 237–250) je objavljen v celoti samo paremiološki optimum, v katerem so izrazi razporejeni z upoštevanjem poznanosti (po anketi) in pogostnosti v jezikovnem korpusu.

Posodobljen vrh slovenskega paremiološkega minimuma (prvih 20 pregovorov in sorodnih paremij po poznanosti – primere od (1) do vključno (20)) navajamo z upoštevanjem odgovorov 527 anketirancev. Navajamo jih v najpogostejši obliki glede na korpus metaFida v1.0 in opremljene z odstotki poznanosti ter pomenskimi razlagami (iz SPP):¹

(1) *Kar lahko storiš danes, ne odlašaj na jutri.* 99,2 %

‘izraža, da je treba delo opraviti takoj, ko je to mogoče; izraža spodbudo, da se kaj opravi prej, čeprav bi bilo to mogoče opraviti tudi kasneje’;

(2) *Kdor išče, ta najde.* 99,1 %

‘izraža, da je potrpežljivost ob iskanju česa na koncu poplačana; izraža spodbudo k vztrajnosti ob iskanju’;

(3) *Jabolko ne pade daleč od drevesa.* 98,9 %

‘izraža, da je otrok po kakšni lastnosti, zanimanju, nadarjenosti podoben svojim staršem ali enemu izmed njih’;

(4) *Skozi eno uho noter, skozi drugo ven.* 98,8 %

‘opisuje situacijo, ko kdo kaj navadno namenoma popolnoma presliši, ignorira, ne nameni pozornosti čemu’;

¹ V SPP so bile do decembra 2023 vključene vse paremije s tega seznama razen *Enkrat ni nobenkrat*, ki je bila v slovarski bazi pripravljena za objavo v letu 2024.

(5) *Ljubo doma, kdor ga ima.* 98,7 %

'izraža, da je dom cenjena vrednota; opisuje situacijo, ko je kdo močno navezan na dom, zadovoljen, da je lahko doma';

(6) *Lepa beseda lepo mesto najde.* 98,7 %

'izraža, da prijazno sporazumevanje naleti na enak odziv, ima ugodne učinke';

(7) *Kar se Janezek nauči, to Janez zna.* 98,7 %

'izraža, da so za delovanje v odrasli dobi odločilni vplivi, znanje, izkušnje iz mladosti, otroštva';

(8) *Klin se s klinom izbija.* 98,7 %

'izraža, da je proti čemu najbolje uporabiti enako ali podobno sredstvo; izraža, da je slabo počutje po alkoholu najbolje odpraviti s ponovnim pitjem alkohola';

(9) *Počasi se daleč pride.* 98,6 %

'izraža opozorilo, da se je treba česa lotiti preudarno, vztrajno, počasi; izraža, da se počasnost in vztrajnost obrestujeta';

(10) *Vsi za enega, eden za vse.* 98,5 %

'izraža poziv k solidarnosti, vzajemni pomoči';

(11) *Iz te moke ne bo kruha.* 98,5 %

'opisuje situacijo, v kateri se zdi, da se želja, napoved ne bo uresničila, da prizadevanje ne bo uspešno';

(12) *Kdor se zadnji smeje, se najslajše smeje.* 98,5 %

'izraža, da je bistveno, da se kaj ugodno konča, zaključi, zlasti če je kdo deležen nasprotovanja, škodoželjnosti, posmehljivosti; izraža, da kdo verjame v uspeh, zmago, čeprav trenutne okoliščine temu niso naklonjene';

(13) *Kjer se prepirata dva, tretji dobiček ima.* 98,5 %

'izraža, da se ob sporu dveh strani okoristi kdo tretji';

(14) *Več glav več ve.* 98,5 %

'izraža, da je s sodelovanjem kaj lažje opraviti, doseči, spoznati';

(15) *Zarečenega kruha se največ poje.* 98,5 %

'opisuje situacijo, ko kdo deluje v nasprotju z lastno izjavo, obljubo, zagotovitvilo; izraža, da je težko predvideti, da kdo česa ne bi storil, počel';

(16) *Enkrat ni nobenkrat.* 98,4 %

'izraža, da enkratni pojav česa ne zadošča';

(17) *Po toči zvoniti je prepozno.* 98,4 %

'izraža, da ni smiselno pritoževati se, iskati rešitev po neželjenem dogodku, izpostavljati kaj, kar se je že zgodilo';

(18) *Kdor drugemu jamo koplje, sam vanjo pade.* 98,3 %

'izraža, da je kdo zaradi svojega škodljivega, spornega delovanja prej ali slej ustrezno kaznovan, mora zanj odgovarjati; opisuje situacijo, ko koga doleti to, kar je skušal zlega storiti drugemu';

(19) *Pes, ki laja, ne grize.* 98,3 %

'izraža, da kdo ni tako težaven, grozeč, nevaren kot se zdi; opisuje situacijo, ko kdo hoče napraviti o sebi vtis, ki ne ustreza dejanskosti';

(20) *Kdor čaka, dočaka.* 98,3 %

‘izraža, da je potrpežljivost ob čakanju na kaj na koncu poplačana; izraža spodbudo k vztrajnosti ob čakanju’.

Med navedenimi izrazi je sedemnajst pregovorov, dva reka (primera (4) in (11)) in frazeologizirani slogan (primer (10)). Slogan je sicer »netradicionalen izraz, ustvarjen za promocijo izdelka ali ideje« (Norrick 2014: 8), v paremiološki slovar pa so uvrščeni taki, ki so se začeli uporabljati izven izvornega konteksta (so se frazeologizirali). Tak je slogan *Vsi za enega, eden za vse*, v slovarju SPP pa še na primer *Važno je sodelovati, ne zmagati*, ki je po izvoru slogan olimpijskih iger. Za razlikovanje med pregovorom in rekom uporabljamo Mlackova (1983: 131) merila, po katerih je pregovor enota z zaključenim (pogosto moralističnim) sporočilom, rek pa izraz, ki opisuje določeno vrsto situacije.

2 Vira za vpogled v paremiološko dediščino in sodobno stanje: paremiološki slovar SPP in zbirka (korpus) Pregovori 1.1

Zbirka Pregovori 1.1 vključuje 37390 zapisov slovenskih pregovorov in je nastajala več kot petdeset let na Inštitutu za slovensko narodopisje ZRC SAZU (Babič 2022; Babič, Erjavec 2022). Viri te zbirke obsegajo tiskane knjige, časopise, koledarje, gradivo zbirateljskih akcij v revijah, terenska dela, osebne zapiske itd. V svetovnem merilu gre za edinstven primer paremiološke zbirke, ki je urejena kot jezikovni korpus. Z lematizacijo leksikalnih sestavin pregovorov omogoča najrazličnejše iskalne možnosti – nekaj izmed njih bomo prikazali v nadaljevanju. Poleg ostalih potencialov te zbirke raziskovalni izziv predstavlja natančno skupinjenje zapisov istega pregovora in možnost avtomatičnega prikaza teh zapisov pod dogovorjeno osnovno obliko pregovora (paremiološko lemo). Z drugimi besedami bi to, kar na primeru 20 pregovorov pilotno prikazujemo v tem članku, lahko označili kot paremiološko lematizacijo.

Slovar pregovorov in sorodnih paremioloških izrazov (SPP) je prvi specializirani paremiološki slovar za slovenščino, ki prikazuje pregovore in izraze pregovorom sorodnih (paremioloških) žanrov kot samostojne iztočnice. Je rastoči slovar: od leta 2020 je prirastek slovarja (približno 130 iztočnic) vsako leto objavljen na portalu Fran.si. Od decembra 2023 obsega 520 iztočnic. Vsaka iztočnica je predstavljena z osnovno slovarsko obliko izraza, ki je v rabi najbolj pogosta, tipom izraza (paremiološkim žanrom), v sodobnem gradivu izpričanimi variantami, pomensko razlago in enim ali več zgledi rabe iz sodobnih besedil. Slovarski sestavek prikazuje tudi z iztočnico pomensko povezane izraze, če so ti v slovarju že predstavljeni. Med iztočnicami prevladujejo pregovori, v slovar pa so vključeni tudi vremenski pregovori, reki, verovanja, slogani, antipregovori, velerizmi, nekonvencionalne replike in drugi paremiološki žanri. Slovar pregovorov in sorodnih paremioloških izrazov nastaja s

pomočjo analize jezikovnih korpusov slovenščine, ki vključujejo obsežno število različnih besedil sodobne slovenščine od 90. let prejšnjega stoletja dalje. Za analizo uporabljamo zbirni korpus metaFida v1.0, ki združuje več kot 30 različnih jezikovnih korpusov, v katere je skupno vključenih okoli 4,7 milijarde besed. V analizo so tako na primer vključena novinarska besedila, leposlovje, besedila spletnih objav (tudi z družabnih omrežij in forumov), besedila akademskih zaključnih del in prepisi poslanskih govorov. S korpusno analizo je za vsak v slovarju predstavljen izraz potrjena ustaljenost rabe v sodobni pisni slovenščini.

3 Korpusno iskanje 20 najbolj poznanih paremij in njihovih variant v zbirki Pregovori 1.1 – primer pregovora *Jabolko ne pade daleč od drevesa*

Slovenski jezikovni korpusi sodobnih besedil omogočajo natančne oblikovne in pomske sinhrone analize slovenskih frazemov in paremij (Meterc 2017, 2019; Kržišnik 2022: 303), ki še pred nekaj desetletji niso bile metodološko dosegljive, poleg tega pa se s korpusi in digitalizacijo večajo možnosti za historično obravnavo frazemov in paremij (Trivunović 2019: 58). Za korpusno iskanje 20 pregovorov uporabljamo medsebojno dopolnjujoče se nabore iskalnih pogojev, da bi zajeli kar se da široko paleto možnih variant (sestavinskih, oblikoslovnih, skladijskih in drugih). Podobne postopke uporabljamo pri preverjanju ustaljenosti oblik za SPP in eSSKJ (Slovar slovenskega knjižnega jezika, tretja izdaja) (Meterc 2019: 37–38).

Pregovor (3) *Jabolko ne pade daleč od drevesa* lahko v korpusu *Pregovori* iščemo z enostavnim iskanjem npr. po sestavini *jabolko* ali pa po zvezi *jabolko ne pade*. Z iskanjem zveze *jabolko ne pade* najdemo naslednje oblike tega pregovora (vsaka oblika se je pojavila le enkrat):

- (21) *Jabolko ne pade daleč od drevesa.*
- (22) *Jabolko ne pade daleč od debla.*
- (23) *Jabolko ne pade daleč od jablane.*
- (24) *Jabolko ne pade daleč proč od jablane.*
- (25) *Jabolko ne pade dalječ od drevesa.*
- (26) *Jabolko ne pade daleč od debla, le potoči se.*
- (27) *Jabolko ne pade daleč od drevesa, le potoči se včasih.*
- (28) *Jabolko ne pade daleč od drevesa, le potoči se.*
- (29) *Jabolko ne pade daleč od drevesa, samo potoči se.*
- (30) *Jabolko ne pade daleč od jablane, le potoči se včasih.*

Domnevamo, da gre pri zadnjih petih oblikah s podaljšavo pregovora za tako imenovano razširitev (Permajakov 1970: 125), ki pomena celote pregovora ne spremeni in jo zato uvrstimo med variante – v primeru spremembe pomena pregovora kot npr.

- (31) *Dober glas gre v deveto vas, slab pa še dlje),*

bi šlo za dopolnitev in s tem nastanek novega paremiološkega izraza. Najdeno razširitev lahko primerjamo s slovaško (prav tako zastarelo) varianto

(32) *Jabolko ne pada d'aleko od stromu, a keď aj padne, vždy sa stopkou ku pňu obracia* (dobesedni prevod: Jabolko ne pada daleč od drevesa, če pa že pade, se vedno s pecljem obrača k drevesu).

Tudi ta razširitev prinaša le dodatno podkrepitev istega sporočila.

Bolj natančno je korpusno iskanje pregovorov s pomočjo iskalnega jezika CQL. Predvidevati moramo, da pregovor lahko obstaja tudi v variantah z drugačnimi leksikalnimi sestavinami. Najbolje je, da izbiramo različne dvojice sestavin in se odločimo, koliko prostih mest naj iskalni pogoj omogoči med sestavinama. Če si izberemo iskalni pogoj s sestavinama *jabolko* in *drevo*, ter med njima omogočimo osem možnih prostih mest ([lemma= »jabolko«][0,8]{lemma=«drevo«}), v korpusu *Pregovori* najdemo naslednje oblike (podčrtane so oblike, ki smo jih našli tudi s prejšnjim iskalnim pogojem):

(33) *Jabelko ne pade daleč od drevesa.*

(34) *Jabelko ne pade deleč od drevesa.*

(35) *Jabelko nikoli daleč ne pade od drevesa.*

(36) *Jabolko ne pada dalječ od drevesa.*

(37) *Jabolko ne pade daleč od drevesa, le potoči se včasih.*

(38) *Jabolko ne pade daleč od drevesa, le potoči se.*

(39) *Jabolko ne pade daleč od drevesa, samo potoči se.*

(40) *Jabolko ne pade daleč od drevesa.*

(41) *Jabolko ne pade dalječ od drevesa.*

Poleg sestavinskih (trikrat namesto samostalnika *jabolko*, trikrat namesto prislova *daleč* in enkrat namesto glagola *pasti*) smo našli tudi skladenjsko varianto (primer s sestavino *nikoli*). S tem iskalnim pogojem najdemo sicer še drug pregovor:

(42) *Zrelo jabelko pade samo z drevesa.*

Če uporabimo iskalni pogoj s sestavinama *pasti* in *daleč* in med njima prav tako omogočimo osem prostih mest, najdemo delno drugačen nabor oblik pregovora, saj ta iskalni pogoj zajame tudi oblike s sestavinami *japk*, *jaboko*, *jabuka*, *drieva*, *iver* in *deblo* ter skladenjsko variantnost. Vendar ta iskalni pogoj ne prikaže oblik s sestavino *dalječ*. Za prikaz oblik v nadaljevanju članka je tako smiselno uporabiti tudi iskalne pogoje z drugačnimi kombinacijami polnopomenskih sestavin (npr. *jabolko* in *pasti*).

Posebej zanimiva je oblika

(43) *Štopl ne pade daleč od flaše,*

ki jo najdemo z iskalnim pogojem s sestavinama *pasti* in *daleč*. Zgolj iz zapisa v zbirki ne moremo vedeti, ali je delovala kot varianta pregovora ali (kar se zdi verjetneje) že kot nova paremija: bodisi je šlo za pregovor (s sporočilom, povezanim z alkoholom in dedovanjem nagnjenosti k popivanju) bodisi antipregovor, ki je služil zgolj kot parodija izhodiščnega; tudi v tem primeru pa bi nas zanimalo, ali je šlo za priložnostno prenovitev ali za ustaljen antipregovor. Obliko bi bilo treba najti s sobesedilom, kar bo morda v prihodnosti omogočil kateri izmed historičnih korpusov.

Pozorni moramo biti ne le na ločevanje med (očitnimi in potencialnimi) antipregovori in njihovimi izhodiščnimi pregovori, temveč tudi na pregovore z enakim tematsko-konstruktivnim vzorcem a drugačnim sporočilom. Pregovor (9) *Počasi se daleč pride* si tak vzorec deli še z naslednjimi pregovori:

(44) *Kdor gre s pametjo po svetu, daleč pride.*

(45) *S poštenjem se daleč pride.*

(46) *Po vprašanju se daleč pride.*

(47) *Pravični človek daleč pride ali se ne sme nikjer dolgo muditi.*

(48) *Blj, kakor je človek premišljen, blj daleč pride.*

(49) *Z lažjo se ne pride daleč.*

(50) *Učen medved v gozdu ne pride daleč.*

Primeri od (44) do (50) kažejo tudi na pestrost udejanjanja vzorca in vprašanje razlikovanja med splošnejšim vzorcem (*X (se) daleč pride*) in podvzorcem (npr. *Z X (pozitivna osebna lastnost) se daleč pride* v primeru (45) in vzorec z zanikanjem *Z X (negativna osebna lastnost) se ne pride daleč* v primeru (49)). V primeru (44) je tematsko-konstruktivski vzorec *X daleč pride* združen s splošnejšim konstruktivnim vzorcem *Kdor X, Y*, v primeru (48) pa s konstruktivnim vzorcem *Bolj kot/kakor X, bolj Y*.

4 Analiza oblik paremij v zbirki Pregovori 1.1 in paremiološkem slovarju SPP

Na začetku navajamo osnovno slovarsko obliko iz paremiološkega slovarja SPP. Variante iz slovarja in oblike, najdene v zbirki Pregovori 1.1, so razvrščene po abecedi. Oblike, ki smo jih našli v zbirki Pregovori 1.1 in so obenem potrjene v sodobni rabi ter vključene v paremiološki slovar (skupno presečišče), smo podčrtali. Pri oblikah, ki se v zbirki pojavijo več kot enkrat, število pojavitev navajamo v oklepaju:

(51) *Kar lahko storiš danes, ne odlašaj na jutri.*

- ♦ Variante iz slovarja SPP:

Kar danes lahko storiš, ne odlašaj na jutri. *Kar lahko danes narediš, ne odlašaj na jutri. Kar lahko danes storiš, ne odlašaj na jutri. Kar lahko narediš danes, ne odlagaj na jutri. Kar lahko storiš danes, ne odlašaj na jutri. Kar lahko narediš danes, ne odlašaj na jutri. Kar lahko storiš danes, ne odlagaj na jutri. Kar lahko storiš danes, ne odlašaj do jutri. Kar lahko storiš danes, ne odlašaj za jutri. Kar lahko storiš danes, ne prelagaj na jutri. Ne odlagaj na jutri, kar lahko narediš danes. Ne odlagaj na jutri, kar lahko storiš danes. Ne odlašaj na jutri, kar lahko narediš danes. Ne odlašaj na jutri, kar lahko storiš danes. Ne prelagaj na jutri, kar lahko storiš danes.*

- ♦ Oblike, najdene v zbirki Pregovori:

Dela, ki ga danes lahko storiš, ne odlašaj na jutri. Kar danes lahko storiš, ne odlašaj do jutri. Kar danes lahko storiš, ne odlašaj na jutri. (2) Kar danes zamoreš, ne odlagaj na jutro. Kar imaš storiti danes, ne odlašaj na jutri. Kar lahko danes postoriš, na jutri ne čakaj, kar svet stoji, se zavrtel nikoli ni nazaj. Kar lahko danes storiš, ne odlagaj na jutro. Kar lahko

narediš danes, ne odlašaj na jutri. Kar lahko narediš danes, ne odlašaj za jutri. Kar lahko postoriš danes, ne odlašaj na jutri. Kar lahko storiš danes, ne prelagaj na jutri. Kar moraš danes storiti, ne odlagaj na jutri. Kar moreš storiti danes, ne odlašaj do jutri. Kar moreš storiti danes, ne odlašaj na jutri. Le danes si pomagaj, na jutri ne odlagaj, ne veš če doživiš, kar danes zamudiš. Ne odlašaj do jutri, kar danas sturiti samoreš. Ne odlašaj nič za juter, kar zamoreš danas storiti.

(52) Kdor išče, ta najde.

Variante iz slovarja SPP:

Kdor išče, najde. Kdor išče, ta tudi najde. Kdor išče, tudi najde.

- ♦ Oblike, najdene v zbirki Pregovori:

Kdor išče, najde. Kdor išče, ta najde, kdor trka, se mu odpre, kdor prosi dobi. Kdor išče, ta najde, kdor trka, temu odpro. Kdor išče, ta najde. Kdor išče, vedno najde.

(53) Jabolko ne pade daleč od drevesa.

Variante iz slovarja SPP:

Jabolko ne pada daleč od drevesa. Jabolko ne pade daleč od debla. Jabolko ne pade daleč od jablane. Sadež ne pade daleč od drevesa.

- ♦ Oblike, najdene v zbirki Pregovori:

Iver ne pade daleč od debla. Jabelko daleč od debla ne pade. Jabelko namreč daleč od debla ne pade. Jabelko ne pade daleč od drevesa. Jabelko ne pade daleko od jablane. Jabelko ne pade deleč od drevesa. Jabelko ne rado daleč od jablana pade. Jabelko nikoli daleč ne pade od drevesa. Jabolko ne pada dalječ od drevesa. Jabolko ne pade daleč od debla, le potoči se. Jabolko ne pade daleč od debla. Jabolko ne pade daleč od drevesa, le potoči se včasih. Jabolko ne pade daleč od drevesa, le potoči se. Jabolko ne pade daleč od drevesa, samo potoči se. Jabolko ne pade daleč od drevesa. Jabolko ne pade daleč od jablane, le potoči se včasih. Jabolko ne pade daleč od jablane. Jabolko ne pade daleč proč od jablane. Jabolko ne pade dalječ od drevesa. Jabolko ne spadne deleč od drejva. Jabuk na pade del'č ad dr'vesa. Jabuka ne pade deleč od debla. Jabuka ne pade deleko ud debla. Japk ne pade daleč od drevesa. Sad ne pade daleč od drevesa. Sadež ne pade daleč od drevesa. Sat ne pade daleč uad drieva. Štopl ne pade daleč od flaše. Vsaka jabolka padejo blizu stebila. Vsaka jabuka pade blizu stebila. Vsako jabolko pade blizu stebila.

(54) *Skozi eno uho noter, skozi drugo ven.*

- ♦ Variante iz slovarja SPP:

Pri enem ušesu noter, pri drugem pa ven. Pri enem ušesu noter, pri drugem ven. Pri enem ušesu noter in pri drugem ven. Skozi eno uho noter, skozi drugo pa ven. Skozi eno uho noter in skozi drugo ven.

- ♦ Oblike, najdene v zbirki Pregovori:

Skuzi enu uhu noter skuzi ta drugu vun.

(55) Ljubo doma, kdor ga ima.

- ♦ Variante iz slovarja SPP:

Lepo je doma, kdor ga ima. Ljubo doma, kdor ga ima; kdor ga nima, pa za njim kima. Ljubo doma.

- ♦ Oblike, najdene v zbirki Pregovori:

Blagor doma, kdor ga ima. (2) Dobro doma, kdor ga ima. Lepo je doma, kdor ga ima. Ljub je dom, kdor ga ima, kdor ga nima, za njim kima. Ljubo doma, kdor ga ima. (2) Ljubo doma, kdor ga ima, kdor ga nima, pa za njim kima. Ljubo doma, kdor ga ima, kdor ga nima, pa za njim naj kima. Ljubo doma, kdor ga ima; kdor ga nima, za njim kima. Ljubo doma, kdor ga ima; kdor ga pa nima, pa za njim kima. Ljubo doma, kdor ga pozna. Ljubo doma, kdor ga prav ravna. Ljubo je doma, kdor ga ima. Zlato doma, kdor ga ima.

(56) Lepa beseda lepo mesto najde.

- ♦ Variante iz slovarja SPP:

Dobra beseda dobro mesto najde. Dobra beseda vedno dobro mesto najde. Dobra beseda vedno mesto najde. Lepa beseda pravo mesto najde. Lepa beseda svoje mesto najde. Lepa beseda vedno lepo mesto najde. Lepa beseda vedno pravo mesto najde. Prava beseda mesto najde. Prava beseda pravo mesto najde. Prava beseda vedno pravo mesto najde. Prijazna beseda lepo mesto najde. Prijazna beseda vedno lepo mesto najde. Prijazna beseda vedno pravo mesto najde.

- ♦ Oblike, najdene v zbirki Pregovori:

Dobra beseda – dobro mesto najde. Dobra beseda lepo mesto najde. Dobra beseda vselej dobro mesto najde. Lepa beseda dobro mesto najde. Lepa beseda lep prostor najde. Lepa beseda lepo mesto najde. (2) Lepa beseda najde lepo mesto. (2) Lepa beseda vedno lepo mesto najde. Lepa beseda vselej lepo mesto najde. Liepa beseda lepo mesto najde. Prijazna beseda vedno lepo mesto najde.

(57) Kar se Janezek nauči, to Janez zna.

- ♦ Variante iz slovarja SPP:

Česar se Janezek nauči, to Janez zna. Česar se Janezek nauči, to Janezek zna. Kar se Janezek nauči, Janez zna. Kar se Janezek nauči, to Janezek zna. Kar se Janezek nauči, to Janez ve. Kar se Janezek nauči, to zna.

- ♦ Oblike, najdene v zbirki Pregovori:

Česar se Janezek nauči, bo tudi Janže znal. Česar se Janezek nauči, to Janez zna. Ka se Janček navči, to Janez zna. Kar Anžik se navučil, tudi Anže bo znou. Kar s' Jan'z'k nauči, Jan'z zna. Kar se Janezek nauči, bo Janez znal. Kar se Janezek nauči, Janez zna. (2) Kar se Janezek nauči, to Janez zna. Kar se Janezek nauči, tue Janez zna. Kar se je Janezek učil, to Janez zna. Kar se pa fantič nauči, tudi starček zna. Zde, kar Anžik se navučil, tudi Anže bo znou, inu kar Anžik se ne ahta navučiti, Anže ne bo znou.

(58) Klin se s klinom izbija.

- ♦ Variante iz slovarja SPP:

Klin se izbija s klinom. Klin se s klinom izbije. Klin se s klinom zbija. Klin se s klinom zbije. Klin se zbija s klinom. Klin s klinom.

- ♦ Oblike, najdene v zbirki Pregovori:

Klin je treba izbijati s klinom. Klin klin izbije, sekira pa oba. Klin s klinom izbij! Klin s klinom! Klin s klinom, skrb s skrbjo izbiti. Klin s klinom, vino z vinom. Klin se s klinom zbija,

vino z vinom. Klin se s klinom zbije in še naprej se pije. Klin se s klinom zbije. Klin s'klinam izbijati, vino s'vinam, pijanost s'vinam pregnati.

(59) Počasi se daleč pride.

- ♦ Variante iz slovarja SPP:

Korak za korakom se daleč pride. Počasi se najdlje pride. Počasi se pride daleč. Počasi se zelo daleč pride. S počasnimi koraki se daleč pride. Tudi počasi se daleč pride. Tudi počasi se pride daleč. Tudi z majhnimi koraki se daleč pride. Tudi z malimi koraki se daleč pride. Z majhnimi koraki se daleč pride. Z malimi koraki se daleč pride. Z vztrajnostjo se daleč pride.

- ♦ Oblike, najdene v zbirki Pregovori:

Kdor počasi hodi, daleč pride. Po časi se pride daleč. Po času se daleč pride. Po kapljicah se daleč pride. Počas se daleč pride. Počasi se daleč pride in pripelja.

Počasi se daleč pride. Počasi se tudi daleč pride. Pomalen se daleč pride. Tudi peš se daleč pride. Tudi počasi se daleč pride.

(60) Vsi za enega, eden za vse.

- ♦ Variante iz slovarja SPP:

Eden za vse, vsi za enega. Eden za vse in vsi za enega. Vsi za enega. Vsi za enega in eden za vse.

- ♦ Oblike, najdene v zbirki Pregovori:

Vsi za enega, eden za vse.

(61) *Iz te moke ne bo kruha.*

- ♦ Variante iz slovarja SPP:

Iz te moke kruha ne bo. Iz te moke ne bo veliko kruha. Ne bo kruha iz te moke.

- ♦ Oblike, najdene v zbirki Pregovori:

Iz te moke ni kruha. Iz tiste moke ne bo kruha.

(62) Kdor se zadnji smeje, se najslajše smeje.

- ♦ Variante iz slovarja SPP:

Kdor se zadnji smeje, se najbolj sladko smeje. Kdor se zadnji smeje, se najslajše. Kdor se zadnji smeje, se smeje najslajše. Kdor se zadnji smeji, se najslajše smeji.

- ♦ Oblike, najdene v zbirki Pregovori:

Kateri se smeje zadnji, se smeje najslajše. Kateri se smeje zadnji, se smeje najslajše. Kateri se smeje zadnji, se smeje najslajše. Kdor se zadnji smeje, se najdlje smeje. Kdor se zadnji smeje, se najlepše smeje. Kdor se zadnji smeje, se najslajše smeje. Najbolj se smeje, kdor se zadnji smeje. Najbolje se smeje, kdor se smeje zadnji.

(63) Kjer se prepirata dva, tretji dobiček ima.

- ♦ Variante iz slovarja SPP:

Kadar se prepirata dva, tretji dobiček ima. Kjer se borita dva, tretji dobiček ima. Kjer se kregata dva, tretji dobiček ima. Kjer se prerekata dva, tretji dobiček ima. Kjer se tepeta dva, tretji dobiček ima. Tam, kjer se prepirata dva, tretji dobiček ima.

- ♦ Oblike, najdene v zbirki Pregovori:

Če se koljeta dva, tretji dobiček ima. Če se pravdata dva, tretji dobiček ima. Čer se pr'perata dva, ta tretje pr'fit 'ma. Kad se skaradu dva, tretji dobiček jima. Kder se prepirata dva, tretji

dobiček ima. Kjer dva prepir mori, se tretji veseli. Kjer koli pričkata se dva, kra, kra, dobička tretji tam ima, kra, kra. Kjer se dva tepeta, tretji klobuke pobira. Kjer se jih več prepira, ima vedno eden dobiček. Kjer se kregata dva, tretji dobiček ima. Kjer se prepirata dva, tam tretji dobiček ima. Kjer se prepirata dva, tretji dobiček ima. Kjer se prepirata dva, tretji se smeje. Kjer se prepirata dve, tretja dobiček ima. Kjer se pretepata dva, tretji dobiček ima. Kjer se tepeta dva, tretji dobiček ima. Kjer si skačeta dva v lase, se tretji veseli.

Si skočita dva v lase, se tretji veseli.

(64) Več glav več ve.

- ♦ Variante iz slovarja SPP:

Več ljudi več ve.

- ♦ Oblike, najdene v zbirki Pregovori:

Več glav več ve. (2) Več glav več zna. Več ljudi več ve. (2) Več ljudi več vidi in več ve. Več ljudi več zna.

(65) Zarečenega kruha se največ poje.

- ♦ Variante iz slovarja SPP:

Zarečenega kruha se poje največ. Zarečenega kruha se poje več kot pečenega. Zarečenega kruha se veliko poje.

- ♦ Oblike, najdene v zbirki Pregovori:

Največja tovarna na svetu peče zarečen kruh. Našpotanega kruha se največ poje. Ne zareci se pa nikar ne, ker zarečenega kruha je več pojedenega kakor zapečenega. Pečenega in zaročenega kruha je naj več usnedenega. Preziranega kruha se največ sne. Sarečen kruh se rad rieže. Saročeniga kruha se dosti poje. Ta zarečenega kruha se največ poje. Tega zarečenga kruha je zmeram več sned'nga, ko tega pečenga. Tega zarečenga kruha je zmeram več sned'nga, ko tega pečenga. Urečenega kruha se dovolj poje. Včasih poješ več kruha zarečenega kot pečenega. Več je bilo zarečenega kakor zapečenega. Več se poje zarečenega kruha ko pa pečenega. Zapečenega in zaročenega kruha je največ snedenega. Zarečen kruh se rad reže. Zarečen kruh se rad rieže. Zarečenega in zapečenega kruha je največ snedenega. Zarečenega kruha je več pojedenega kakor zapečenega. Zarečenega kruha je več pojedenega, kakor pa zapečenega. Zarečenega kruha je zmerom več snedeniga ko pečeniga. Zarečenega kruha poješ več kot pečenega. Zarečenega kruha se človek največ naje. Zarečenega kruha se dosti poje, kaj ga še čaka, nobeden ne ve. Zarečenega kruha se dosti sne. Zarečenega kruha se največ poje. (2) Zarečenega kruha se največ sne. Zarečenega kruha se poje več kot pečenega. Zarečenega kruha se več poje kakor pečenega. Zarečenega kruha se več poje kot zapečenega. Zarečenega kruha se več sne kot zapečenega. Zarečenega kruha se veliko poje. Zarečenega kruha se veliko sne. Zarečenega kruha veliko se sne. Zarečeniga kruha veliko se sne. Zarečenoga kruha se puno poji. Zarečenoga kruha se puno poji.

(66) Enkrat ni nobenkrat.

- ♦ Variante za slovar SPP (objava v 2024): /

- ♦ Oblike, najdene v zbirki Pregovori:

Enkrat ni nobenkrat.

(67) Po toči zvoniti je prepozno.

- ♦ Variante iz slovarja SPP:

Po toči je prepozno zvoniti. Prepozno je zvoniti po toči. Zvoniti po toči je prepozno.

- ♦ Oblike, najdene v zbirki Pregovori:

Ne pomaga zvoniti, ko je toča že pobila. Po toči in nesreči zvoniti je prepozno. Po toči je pozno zvoniti. Po toči je prepozno zvoniti. Po toči je škoda zvoniti. Po toči je zastoj zvoniti. Po toči ni zvonjenja. Po toči zastoj zvoniti. Po toči zastoj zvoniti. Po toči zvoniti in po smrti pokoro delati je prepozno. Po toči zvoniti je prekasno. Po toči zvoniti je prepozno. Po toči zvoniti je zastoj. Prepozno po toči je zvoniti, plat zvona nad pogoriščem biti, grivngo po smrti obuditi.

(68) Kdor drugemu jamo koplje, sam vanjo pade.

- ♦ Variante iz slovarja SPP:

Kdor drugemu jamo koplje, sam noter pade. Kdor drugemu jamo koplje, sam notri pade. Kdor drugemu jamo koplje, sam pade vanjo. Kdor drugemu jamo koplje, sam vanjo pada. Kdor drugemu jamo koplje, sam v njo pade. Kdor drugim jamo koplje, sam pade vanjo. Kdor drugim jamo koplje, sam vanjo pade. Kdor jamo drugemu koplje, sam vanjo pade. Kdor jamo koplje, sam vanjo pade.

- ♦ Oblike, najdene v zbirki Pregovori:

Če drugemu jamo koplješ, sam vanjo padeš. Če koplješ jamu drugemu, sam vanjo padeš. Kdo komi jamo kopa, sam v njo padne. (2) Kdor drugemu jamo koplje, sam vanjo pade. Kdor drugemu jamo kople, sam v njo pade. Kdor drugemu jamo koplje, sam not pade. Kdor drugemu jamo koplje, sam sigurno vanjo pade. Kdor drugemu jamo koplje, sam vanjo pade. Kdor drugemu luknjo koplje, naj sam vanjo pade! Kdor drugim jamo koplje, sam vanjo pade. (2) Kdor drugim jamo koplje, se sam vanjo pogrebne. Kdor drugim jamo koplje, se sam vanjo pogrezne. Kdor drugemu jamo koplje, sam vanjo pade. Kdor komur jamo koplje, sam v njo pade. Kduor drugim jamo koplje, sam vanjo pade. Ker drugemu jamo kopa, sam v njo pade. Ki drugon jamu kopa, sam u nju pade. Meni si kopala jamo, zdaj pa si sama padla vanjo. Nevoščljivost drugim kopa jamo, pa sama u njo pade. Što drugomi jamo kopa, san v-njo spadna. V jamo pade, kdor jo drugemu koplje. V jamo pade, kdor jo drugim koplje.

(69) Pes, ki laja, ne grize.

- ♦ Variante iz slovarja SPP:

Kdor laja, ne grize. Kdor veliko laja, malo grize. Kdor veliko laja, ne grize. Ne boj se psa, ki laja. Pes, ki dosti laja, malo grize. Pes, ki laja, ponavadi ne grize. Pes, ki laje, ne grize. Pes, ki veliko laja, bolj malo grize. Pes, ki veliko laja, malo grize. Pes, ki veliko laja, ne grize. Psi, ki lajajo, ne grizejo. Tisti pes, ki laja, ne grize.

- ♦ Oblike, najdene v zbirki Pregovori:

Boječ pes huje laja kakor grize. Kateri pes veliko laja, malo grize. Kdor veliko laja, malo grize. Ne boj se psa, ki glasno laja. Ne boj se psa, ki laja. Ne čuvaj se psa, ki laja, ampak, ki molči. Pes boječ huje laja, kakor grize. (2) Pes boječ huje laje, nego grize. Pes, ki dosti laja, malo grize. Pes, ki laja, ne grize. (2) Pes, ki laja, ne ugrizne.

Pes, ki laja, ni hud. Pes, ki laja, ni nevaren, ampak tisti, ki tiho pride. Pes, ki laja, ni nevaren, ampak tisti, ki tiho pride. Pes, ki laja, ni nevaren. Pes, ki močno laja, ne grize. Pes, ki veliko laja, malo grize. Pes, ki veliko laja, ne grize. Pes, ki veliko laja, ne ugrizne. Tisti pes, ki veliko laja, slabo ugrizne. Uni pas, ki laja, ne grizne. Uni pas, ki laja, ne grizne. Vsak pes, ki veliko laja, malo ugrizne.

(70) Kdor čaka, dočaka.

- ♦ Variante iz slovarja SPP:

Kdor čaka, tudi dočaka. Kdor čaka, ta dočaka. Kdor čaka, ta tudi dočaka.

- ♦ Oblike, najdene v zbirki Pregovori:

Če čakaš, tudi dočakaš. Kdor čaka, dočaka. (2) Kdor čaka, očaka pričaka. Kdor čaka, pričaka. Kdor čaka, ta dočaka. Kdor čaka, včaka! Kdor počaka, dočaka. Ki čeka, dočeka. Ki čeka, tudi dočeka.

4.1 Kaj smo izločili s seznama najdenih oblik

V zgornjem seznamu nismo prikazovali paremij s podobno zgradbo ter sorodno, vendar spremenjeno frazeološko motivacijo (podobo) in spremenjenim sporočilom:

(71) *Dokler se dva za kravo prepirata, jo tretji smeje odžene.*

(72) *Če se moškemu kaj pove, gre pri enem ušesu noter, pri drugem pa ven; če ženski, gre pri obeh ušesih noter, pri ustih pa ven.*

(73) *Nevoščljivost drugim kopa jamo, pa sama u njo pade.*

Pregovor v primeru (72) je nastal iz reka *Skozi eno uho noter, skozi drugo ven*, tako da je lastnost, ki jo opisuje rek, v pregovoru pripisana moškemu, v karakterizaciji žensk pa je izhodiščni rek modificiran. Pregovor v primeru (73) je nastal z zožitvijo pomena izhodiščnega pregovora.

Med težje razločljivimi zapisi v zbirki je

(74) *Po soncu zvoniti ni prepozno.*

Brez primera rabe v besedilu se lahko le sprašujemo, ali gre za napako, pregovor z drugačnim sporočilom ali morda za antipregovor. Podobno se sprašujemo v primeru oblike

(75) *Več glav več zmešnjav,*

ki je potrjena v dveh virih in je tudi nismo vključili, ker domnevamo, da gre za paremijo z drugačnim sporočilom. Le ob iskanju ene izmed vseh 20 paremij z vrha slovenskega paremiološkega minimuma smo naleteli na očitne primere antipregovorov, ki izhajajo iz pregovora *Kdor drugemu jamo koplje, sam vanjo pade*, saj jih poznamo tudi iz sodobne rabe:

(76) *Kdor drugemu jamo koplje, je svoje sreče kovač.*

(77) *Kdor drugim jamo koplje, je grobar.*

(78) *Kdor drugemu koplje jamo – dela na Žalah.*

Nenavaden je zapis (79) *Kdor drugemu jamo koplje*, ki deluje kot napaka zapisovalca. Pregovor z drugim (a dopolnjujočim) pomenom (78) *Pes, ki ne laja, zobe zasaja* je enkrat v zbirki prisoten z napačnim, nezanimanim zapisom:

(79) *Pes, ki laja, zobe zasaja.*

V zgornjem seznamu nismo prikazali tudi očitnega primera pretvorbe v besednozvezni frazem

(80) *Zvoni po toči.*

4.2 Pestrost oblik, najdenih v zbirki Pregovori 1.1, v primerjavi s pestrostjo variant v sodobni rabi

20 v sodobni slovenščini najbolj poznanih paremij je v korpusu Pregovori predstavljениh z veliko množico oblik. Zanje smo iz slovarja SPP izpisali 134 oblik (vključno z osnovnimi slovarskimi oblikami), kar pomeni v povprečju 6,7 ustaljenih variant na eno paremijo, iz zbirke pa skoraj enkrat več – 249 oblik – 12,4 na eno paremijo. Le v primerih (54), (56), (59), (60) in (61) je število v današnji rabi potrjenih variant višje od števila oblik v zbirki Pregovori 1.1.

V gradivu zbirke Pregovori 1.1 je pestra paleta najrazličnejših tipov variant, npr. sestavinskih, oblikoslovnih in skladijskih. Posebej zanimive so številne podaljšave in krajšave, ki pomena paremij ne spremenijo. Med njimi so take, ki jih poznamo tako v sodobni rabi kot tudi iz virov v zbirki, kot v primeru (55) razširitev *Ljubo doma, kdor ga ima, kdor ga nima, pa za njim kima*. V zbirki najdemo tudi razširitve, ki jih v sodobni rabi nismo potrdili, npr.

Zde, kar Anžik se navuči, tudi Anže bo znou, inu kar Anžik se ne ahta navučiti, Anže ne bo znou, Klin s klinom, skrb s skrbjo izbiti, Klin se s klinom zbija, vino z vinom, Klin se s klinom zbije in še naprej se pije in Jabolko ne pade daleč od debela, le potoči se.

Pestrost variant v korpusu Pregovori smo pričakovali, saj je časovni razpon virov te zbirke izredno širok – od 16. stoletja do leta 2010 – večina pregovorov pa ima tudi v sodobni slovenščini visoko število ustaljenih variant, potrjenih v virih za SPP od začetka 90. let 20. stoletja do danes. Le predstavljamo si lahko, kako pester nabor variant bi bil prikazan v zbirki Pregovori 1.1, če bi za podobne časovne intervale (štiri desetletja) imeli enako izčrpne vire, kot jih imamo z jezikovnimi korpusi za zadnjih 40 let. Samo primer (66) (*Enkrat ni nobenkrat*), poleg te osnovne oblike nima potrjene nobene variante tako v sodobni slovenščini kakor tudi v virih korpusa Pregovori. Po manjšem številu oblik izstopajo reki v primerih (54) in (61) in dve paremiji z enostavnim konstrukcijskim vzorcem *Kdor X, ta Y: Kdor išče, ta najde* (52) in *Kdor čaka, dočaka* (70). Največ oblik (kar 36) smo našli za pregovor *Zarečenega kruha se največ poje* (65).

Število variant v zbirki Pregovori 1.1 je visoko zaradi mnogih narečnih variant, narečnih besedil pa v gradivu korpusov, vključenih v metaFido, praviloma ni. Pestrost oblik v zbirki je celo tolikšna, da se ista oblika v zbirki pojavi največ dvakrat – pa še to se zgodi le v primeru 14 oblik izmed skupno 249. Res je, da bi lahko šteli skupaj še variante, ki se razlikujejo le v narečnih dvojnicah leksemov (npr. *Kduor drugim jamo koplje, sam vanjo pade*). V šestih primerih gre za obliko, ki je v današnji rabi tako pogosta, da je bila določena za glavno slovarsko obliko. Pričakovali bi, da bi se nekatere oblike vendarle pojavile v več zgledih. Samo s pomočjo korpusa Pregovori

ne moremo določiti osnovne oblike paremije, paremiološke leme, sicer pa glede na časovni razpon virov za to zbirko, to tudi ni smiselno. Če bi za potrebe paremiološke lematizacije korpusa v prihodnosti izbirali osnovno obliko, bi jo bilo (tudi zaradi vidika sodobnega uporabnika) bolj smiselno določiti glede na obliko, ki je najbolj reprezentativna v sodobni rabi in to celo v primerih, v katerih se taka oblika v zbirki sploh ne bi pojavila (v primeru 20 analiziranih paremij v tem prispevku bi se to zgodilo dvakrat – v primerih (54) in (61)).

4.3 Prekrivnost oblik iz zbirke Pregovori 1.1 z naborom v današnji rabi potrjenih variant

Oblike, ki so obenem potrjene v zbirki Pregovori 1.1 in v sodobni rabi ter vključene v paremiološki slovar (skupno presečišče), smo podčrtali. Tako smo označili le popolnoma enake oblike, pri čemer je treba upoštevati, da se nekatere razlikujejo le v narečni obliki sestavine ali zgolj zapisu (npr. *dalječ* namesto *daleč*). Med 134 oblikami iz slovarja in 249 iz zbirke Pregovori 1.1 je 40 oblik, ki predstavljajo presečišče, kar pomeni, da je bilo 16 % oblik v zbirki potrjenih v sodobni rabi. V povprečju gre za 2 obliki na paremijo, največ pa za 3 oblike. Izmed 20 paremij je v 16 primerih osnovna slovarska oblika prisotna tudi v zbirki Pregovori 1.1. Samo v primerih (51), (54), (58) in (61) njihova osnovna slovarska oblika ni bila potrjena tudi v zbirki. V primerih (51) in (58) presečišče predstavlja oblika, ki je bila v paremiološki slovar uvrščena kot varianta. Le v primerih (54) in (61), ki ju žanrsko lahko uvrstimo med reke, se nobena oblika iz slovarja ne pojavi tudi v zbirki in obratno.

5 Zaključki

Potrdilo se je pričakovanje, da gre pri 20 najbolj poznanih paremijah v sodobni slovenščini za izraze, ki so v tem jeziku prisotni že dlje časa, so tradicionalni in jih lahko (v veliki večini primerov v več variantah) potrdimo tudi v virih korpusa Pregovori. Ker pa se paremije ves čas spreminjajo, je množica v današnji rabi ustaljenih variant le delno prekrivna z oblikami iz starejših virov, ki so vključeni v zbirko Pregovori 1.1. Le 16 % oblik iz zbirke je ustaljena v sodobni rabi in le nekaj manj kot tretjino (29,8 %) v današnji rabi uveljavljenih variant najdemo tudi v virih zbirke. Presečišče sicer izkazuje določeno mero ohranjanja oblik izrazov v slovenskem paremiološkem fondu, vendar pa predvsem pričča o spreminjanju paremiološkega fonda z uveljavljanjem novih oblik paremij, čeprav lahko domnevamo, da so bile nekatere izmed v današnji rabi potrjenih variant prisotne tudi že prej, le da jih viri za zbirko niso uspeli v celoti zajeti. Izrazita kontinuiteta se kaže v osnovnih oblikah paremij (najbolj pogostih v sodobni rabi), saj gre v veliki večini primerov za oblike, ki so bile prisotne že v preteklosti in so tudi izpričane v virih zbirke Pregovori 1.1, kar ni presenetljivo. Presenetljivo pa se nam zdi, da se te osnovne oblike v zbirki Pregovori 1.1 ne pojavljajo z več zgledi iz

različnih zbirk. Tako je le v primeru šestih izmed 20 paremij, pa še tu gre za največ dva navedka v zbirki. Kot smo izpostavili že v komentarju o splošni pestrosti oblik v korpusu Pregovori, je nizka ponovljivost oblik zanimiva lastnost sestave fonda paremij v tem korpusu na splošno.

V članku smo se dotaknili tudi vprašanja skupinjenja različnih oblik istega pregovora v tej obsežni diahroni zbirki oziroma vprašanja paremiološke lematizacije. Hibridni medij paremiološke zbirke v obliki jezikovnega korpusa, ki je trenutno edini takega tipa na svetu, bi lahko v prihodnosti obsegal dva nivoja lematizacije: poleg trenutne leksikalne še paremiološko. Kakšno podobo bi morala posamezna paremiološka lema imeti, je stvar (paremiografske) rešitve oziroma dogovora: pri pregovorih, ki so v rabi prisotni tudi danes, bi bila lahko enaka najbolj pogosti obliki v sodobnih besedilih, za v sodobni rabi nepotrjene oblike pa bi lahko izbrali obliko, ki se pojavi v najmlajšem viru. Drugi del naloge ob paremiološki lematizaciji pa prinaša še obsežnejše delo: vsaka paremiološka lema mora biti povezana z vsako obliko istega pregovora v zbirki. V pričujočem članku nam je verjetno uspelo prikazati večino v zbirki prisotnih oblik posamezne paremije, domnevamo pa, da ne vseh. Težava je v tem, da bi za tako nalogo morali pregledati celotno zbirko »ročno«, saj se oblike z močno sestavinsko in skladenjsko (ter še kakšno) variantnostjo lahko izmaknejo še tako podrobnemu iskanju s kompleksnimi iskalnimi pogoji, posebna težava, ki smo jo prikazali v prispevku, pa je tudi določanje, ali gre zares za paremiološko varianto ali že za nov izraz (npr. nov pregovor, nastal z dopolnitvijo ali pa antipregovor). Namesto ugibanja bo v takih primerih kot v primeru *Štopl ne pade daleč od flaše*, ki smo ga v članku interpretirali kot antipregovor, potrebna potrditev pomena v rabi s sobesedilom – to bo v prihodnosti omogočeno z večanjem historičnih korpusov.

Zahvala

Raziskava je nastala v okviru projekta *Tradicionalne paremiološke enote v dialogu s sodobno rabo* (ARIS: J6-2579) in raziskovalnega programa *Slovenski jezik v sinhronem in diahronem razvoju* (ARIS: P6-0038).

Viri in literatura

- Babič, Saša, 2022: Raziskovanje paremioloških enot in nastajanje znanstvene zbirke pregovorov v okviru Inštituta za slovensko narodopisje ZRC SAZU. *Slavistična revija*, 70/3, 289–298. DOI: <https://doi.org/10.57589/srl.v70i3.4019>.
- Babič, Saša; idr., 2023, Collection of Slovenian paremiological units Pregovori 1.1, Slovenian language resource repository CLARIN.SI. Na spletu: <http://hdl.handle.net/11356/1853>.
- Babič, Saša; Erjavec, Tomaž, 2022: Izdelava in analiza digitalizirane zbirke paremioloških enot. Izdelava in analiza digitalizirane zbirke paremioloških enot. V: Fišer, Darja; Erjavec, Tomaž (ur.), *Jezikovne tehnologije in digitalna humanistika: zbornik konference = Proceedings*

- of the Conference on Language Technologies and Digital Humanities*. Ljubljana: Inštitut za novejšo zgodovino, 17–22.
- Đurčo, Peter, 2002: K výskumu súčasnej živej slovenskej paremiológie. V: Mlacek, Jozef (ur.), *Studia academica slovacca 31. Prednášky XXXVIII. letnej školy a kultúry*. Bratislava: Stimul, 51–60.
- Đurčo, Peter, 2014: Empirical Research and Paremiological Minimum. V: Hrisztova-Gotthardt, Hrisztalina; Varga, Melita Aleksa (ur.), *Introduction to Paremiology: A Comprehensive Guide to Proverb Studies*. Varšava, Berlin: De Gruyter Open Ltd, 183–205. DOI: <https://doi.org/10.2478/9783110410167>.
- Kržišnik, Erika, 2022: Paremiološke enote – variante, variacije in novi frazemi. *Slavistična revija* 70/3, 301–314. DOI: <https://doi.org/10.57589/srl.v70i3.4038>.
- Mlacek, Jozef, 1983: Problémy komplexného rozboru prísloví a porekadiel. *Slovenská reč* 48/2, 129–140.
- Meterc, Matej, 2014: *Primerjava paremiologije v slovenskem in slovaškem jeziku na osnovi paremiološkega optimuma: doktorska disertacija*. Ljubljana: Filozofska fakulteta.
- Meterc, Matej, 2017: *Paremiološki optimum: najbolj poznani in pogosti pregovori ter sorodne paremije v slovenščini*. Ljubljana: Založba ZRC, ZRC SAZU (Lingua Slovenica, 10). DOI: <https://doi.org/10.3986/9789610504153>.
- Meterc, Matej, 2019: Analiza frazeološke variantnosti za slovarski prikaz v eSSKJ-ju in SPP-ju. *Jezikoslovni zapiski* 25/2, 33–45. DOI: <https://doi.org/10.3986/JZ.25.2.2>.
- Meterc, Matej, 2020–: Slovar pregovorov in sorodnih paremioloških izrazov. Na spletu: <https://fran.si/216>.
- Meterc, Matej, 2022: Razlikovanje med dopolnitvijo in razširitvijo pri podaljšavi pregovora v frazeološki teoriji in paremiografiji. *Slavistična revija* 70/3, 315–331. DOI: <https://doi.org/10.57589/srl.v70i3.3996>.
- Norrick, Neal, 2014: Subject Area, Terminology, Proverb Definitions, Proverb Features. V: Hrisztova-Gotthardt, Hrisztalina; Varga, Melita Aleksa (ur.), *Introduction to Paremiology: A Comprehensive Guide to Proverb Studies*. Varšava, Berlin: De Gruyter Open Ltd. DOI: <https://doi.org/10.2478/9783110410167>.
- Permjakov, Grigorij, 1970: *Ot pogovorki do skazki. Zametki po obščej teorii kliše*. Moskva: Nauka.
- Permjakov, Grigorij, 1988: *Osnovy strukturnoj paremiologii*. Moskva: Nauka.
- Trivunović, Eva, 2019: Diahrono raziskovanje biblijskih in izbiblijskih frazemov. *Jezikoslovni zapiski* 25/2, 47–61. DOI: <https://doi.org/10.3986/JZ.25.2.3>.
- Ulčnik, Natalija; Meterc, Matej; Jesenšek, Vida, 2021: Frazemi in pregovori v učnem e-okolju Slovenščina na dlani [Phrasemes and proverbs in the e-learning environment »Slovenščina na dlani«]. V: Ulčnik, Natalija; Antloga, Špela (ur.), *Slovenščina na dlani 4*. Maribor: Univerza v Mariboru, Univerzitetna založba, 75–92. DOI: <https://doi.org/10.18690/978-961-286-521-4>.
- Varga, Melita Aleksa; Feldvari, Kristina, 2022: Constructing the Digital Proverbial Thesaurus: Theoretical and Methodological Implications. V: Corpas Pastor, Gloria; Mítkov, Ruslan (ur.), *Computational and Corpus-Based Phraseology - 4th International Conference, Europhras 2022*. Cham: Springer, 1–15. DOI: https://doi.org/10.1007/978-3-031-15925-1_1.



Proverbs about wealth and poverty on social media: Usage and functions

Pregovori o bogastvu in revščini na družbenih omrežjih: raba in funkcije

This paper discusses the under-researched area of the use and functions of proverbs in social media communication. The study focuses on how proverbs about wealth and poverty are used on social media, taking into account the content of messages, user profiles and features, and the characteristics of Slovene proverbs. The corpus-based analysis and interpretation of their use in social media texts reveals that proverbs express mostly negative – critical, moralizing, judgemental – attitudes to both wealth and poverty in different types of social media texts. They serve various pragmatic functions like giving advice, warnings, criticism and expressing emotions. The research contributes to understanding the use of proverbs on social media as a form of social commentary.

• **Keywords:** corpus of texts from Slovene social media sites Janes V1.0, figurative language, paremiology, phraseology, proverbs 1.0 collection, proverbs from the field of human economy, social media, social networks

Članek obravnava slabo raziskano področje rabe in funkcij pregovorov v komunikaciji na družbenih omrežjih. Študija se osredotoča na to, kako se pregovori, povezani z bogastvom in revščino, uporabljajo na družbenih omrežjih, pri čemer upošteva vsebino sporočil, profil in lastnosti uporabnikov ter značilnosti slovenskih pregovorov tega področja. Analiza in interpretacija rabe pregovorov na podlagi korpusa razkriva, da pregovori večinoma izražajo negativen – kritičen, moralizirajoč – odnos do bogastva in revščine v različnih vrstah besedil na družbenih omrežjih. Opravljajo različne pragmatične funkcije, kot so nasveti, opozorila, kritika in izražanje čustev. Z raziskavo želimo prispevati k razumevanju rabe pregovorov na družbenih omrežjih kot oblike družbenega komentarja.

• **Ključne besede:** korpus besedil s slovenskih družbenih omrežij Janes v1.0, figurativni jezik, paremiologija, frazeologija, zbirka pregovorov 1.0, pregovori s področja gospodarstva, družbeni mediji, družbena omrežja.

1 Introduction

Proverbs are a linguistic phenomenon and constitute a treasure trove of a nation's intangible cultural heritage. From a historical perspective they enshrine the wisdom and beliefs of previous generations, and from a synchronic perspective the use of proverbs reflects how their content and messages are perceived and valued in modern society. Although proverbs originated a long time ago, they are still used in modern languages and form a noticeable part of our everyday communication in conversations, newspapers, magazine articles, horoscopes, etc. However, in the 21st century social networks such as Twitter, Facebook, Instagram, YouTube, Tiktok and other

social media such as forums and blogs also seem to be a popular medium for the use of proverbs.¹

The factors influencing the use of proverbs on social media are neither well researched nor explained. This study therefore attempts to fill a gap in the literature by answering the following questions: Does the inclusion of proverbs in dictionary resources affect their use on social media? Is the use of proverbs on social media deliberate and conscious? What types of modifications of proverbs appear on social media? What is the role of proverbs on social media? Is it different from the role of proverbs in other media? Are there differences in the use of proverbs depending on the type of social media?

1.1 Characteristics of proverbs

Although there is no universal definition of proverbs, the essential characteristics of proverbs can be inferred from the following two definitions by Mieder.

Proverbs are short, generally known sentences of the folk that contain wisdom, truths, morals and traditional views in a metaphorical, fixed and memorizable form and that are handed down from generation to generation. (Mieder 1996: 597)

Proverbs contain everyday experiences and common observations in succinct and formulaic language, making them easy to remember and ready to be used instantly as effective rhetoric in oral or written communication. (Mieder 2004: xi)

When discussing their utility and if we aim to comprehend their use and functions on social media, it is important to understand the basic characteristics that make these fixed, mostly metaphorical sentences or microtexts that have been passed down from generation to generation in more or less unchanged form, such a useful means of communication. When it comes to the characteristics of proverbs, various elements and contents are highlighted in linguistic and folkloristic literature (Mieder 2004). For the purpose of our linguistic study, Ulčnik's (2015) discernment of proverb characteristics proves beneficial, as it categorizes them into multiple thematic groups with individual features naturally exhibiting some degree of overlap. They can be separated according to different dimensions or aspects into six levels:

a) On a structural level, proverbs exhibit features such as sentence or microtext structure, relative structural stability with the possibility of variations and modification, and the use of ellipsis.

b) On the semantic level, proverbs exhibit metaphorical meaning, they contain idiomatic expressions and metaphoric elements, and demonstrate the possibility of dual reading (literal and metaphorical). Features like expressiveness, picturesqueness and connotativity put them in the domain of figurative language.

¹ Cf. Mieder (2005) who shows that the frequent use of traditional and new proverbs and also their innovative variations are noticeable in the mass media.

c) On the pragmatic or functional level, proverbs serve as expressions of folk wisdom, convey general truths derived from life experiences, exhibit multifunctionality, offer instructive insights, demonstrate didacticism and perform speech acts.

d) On the stylistic level, proverbs demonstrate a skilful integration of rhyme, alliteration, assonance, personification, hyperbole, parallelism and other rhetorical tools, showcasing not only creativity but also rhetorical efficiency.

e) On the psycholinguistic level, their key features include general knowledge, the retention of units in memory and memorability.

f) On the cultural level, features encompass intergenerational tradition, cultural heritage, unknown authorship and untranslatability.

Based on the basic characteristics, we define proverbs for the purposes of our research as pithy expressions that typically take the form of sentences expressing folk wisdom and universal truths. These expressions use metaphorical language and reflect various life experiences and the culturally relevant values of a particular language community in which they serve communicative functions and perform speech acts, such as reinforcement, disagreement, irony, warning, threat, advice, guidance, apology, encouragement, consolation, conclusion and argument. The secret of their longevity in the language lies in their high and universal communicative value. Thanks to their ability to convey many universal messages that are formally summarised in short sentences, proverbs are a popular language tool for expressing, emphasizing, elucidating, summarizing or arguing in favour of our thoughts, attitudes, opinions and values. Proverbs are stored in our vocabulary, i.e. they are reproduced rather than created in communication. Therefore, both the form and content of proverbs make their use in social media practical and expected.

2 Methods

In our research we used both quantitative and qualitative methods. The analysis was based on data from the Proverbs 1.0 collection, available on the Clarin repository (clarin.si) via the KonText and NoSketch search engines (Babič and Erjavec 2022). The collection contains over 36,200 proverbs extracted from 2,515 bibliographic units spanning more than 400 years, from 1578 to 2010 (Babič et al. 2022).

In this study, we have focused on the proverbs from the field of human economy that are related to wealth and poverty. Although we are aware of the value of discussing the cultural and linguistic interpretations of wealth and poverty, our research is limited to their dictionary definitions in Slovene, where *wealth* primarily denotes the condition of a rich person or a large quantity of material goods, while *poverty* is defined as a severe lack of material goods, a small amount of property or just refers to poor people. Therefore, we conducted an extensive search and selected proverbs

containing the following components: *denar* (money), *kruh* (bread), *bogat* (rich), *reven* (poor), *bogastvo* (wealth), *revščina* (poverty), *bogataš* (rich man), *revež* (poor man), *siromak* (pauper), *sirota* (orphan), *zlato* (gold), *srebro* (silver) and *berač* (beggar). The collection contains 2,051 proverbs and their variants (the distribution of components is illustrated in Figure 1).

We validated the relevance of the proverbs for contemporary use by consulting contemporary dictionaries available on the *Fran portal* (Fran, 2014-)² and by verifying their concordances in the Slovenian reference corpus *Gigafida*³ and the *Corpus of Texts from Slovene Social Media Sites Janes*.⁴ We then conducted a corpus-based linguistic analysis to investigate the frequency of their use and analyse the structure and meaning of the selected proverbs. To gain insight into the specific ways proverbs are used in social media, we compared the frequency of the selected proverbs in two corpora, *Janes* and *Gigafida*. As the corpora were of different sizes, we relied on a statistical calculation that took into account the occurrence of proverbs per million words and their normalized and relative frequency. When analysing the proverbs in the *Janes* corpus, we considered various parameters including the type and characteristics of the social media, the attributes of social network users (such as public profile, gender, etc.), the sentiment of the context (Fišer et al. 2016; Liu 2012) and the correlation between frequency of use and familiarity with proverbs. We interpreted their pragmatic functions in the social media context, categorized according to different types of speech acts, such as advising, warning, criticizing, encouraging and expressing feelings. In the final phase of our analysis, we concentrated on the use of the selected proverbs in contemporary social media, with a particular focus on the discursive interpretation of their functions and any potential socially or politically relevant modifications.

3 Results and interpretation

3.1 Proverbs on wealth and poverty

More than 52 % of Slovenian proverbs that deal with wealth or poverty have the components *denar* (money) or *kruh* (bread) in their structure (Figure 1). This suggests that proverbs often do not speak directly about wealth and poverty, but through the metaphorical use of the words bread and money (Jakop 2022). A comparison between

² The Fran portal, version 11.0, combines 45 dictionaries, a linguistic atlas, language and terminological counselling, 14 collections, a total of 764,931 dictionary entries.

³ Corpus Gigafida 1.0 includes more than one billion words from various text types such as newspapers (47.8%), magazines (16.5%), internet (28%), non-fiction (3.8%), and fiction (3.5%) (Krek et al. 2019).

⁴ The Janes Corpus contains around 200 million words from different types of social media, such as tweets (60%), forum posts (20%), user comments from online news portals (6%), blog posts with their user comments (14%), and conversations and user pages from Wikipedia (2%) (Fišer et al. 2016, Fišer et al. 2020).

the components of wealth (e.g. *bogataš, zlato*) and poverty (e.g. *berač, revščina*) shows that the proportion of proverbs dealing with poverty is only slightly higher than the proportion of proverbs dealing with wealth (Figure 2). In this context, it should be emphasized that individual proverbs are frequently composed of contrasting elements. For example, *revščina – bogastvo: Revščina ne sili h kraji kakor bogastvo ne odvrača od nje*, (lit. Poverty does not compel theft, just as wealth does not deter from it). Semantically and functionally this intends to establish a moral and didactic opposition between wealth and poverty, as well as between being rich or poor. Slovenian proverbs indicate that people are about as affected, annoyed and intrigued by poverty as they are by wealth. Slovenian proverbs moralise, judge, evaluate and teach both sides to about the same extent.

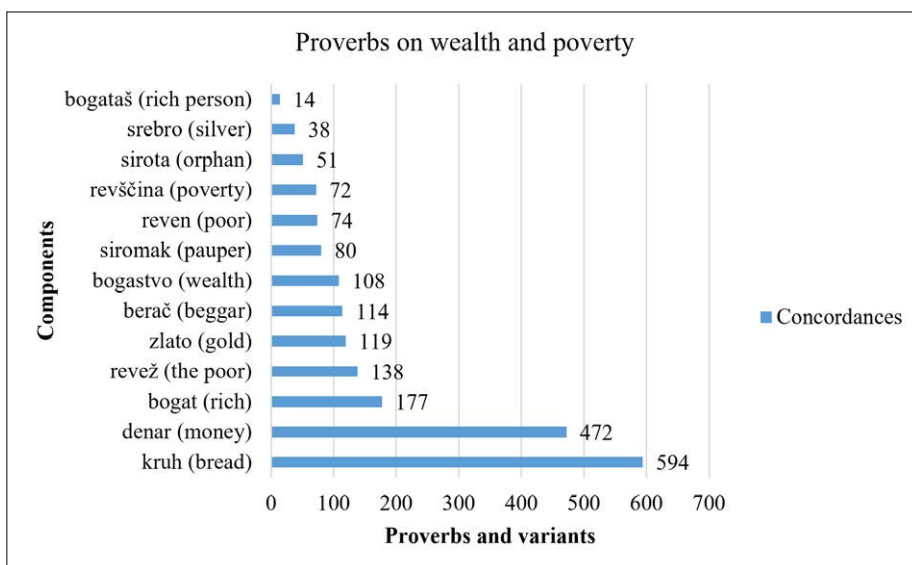


Figure 1: Components and concordances of proverbs related to wealth and poverty in the Proverbs collection.

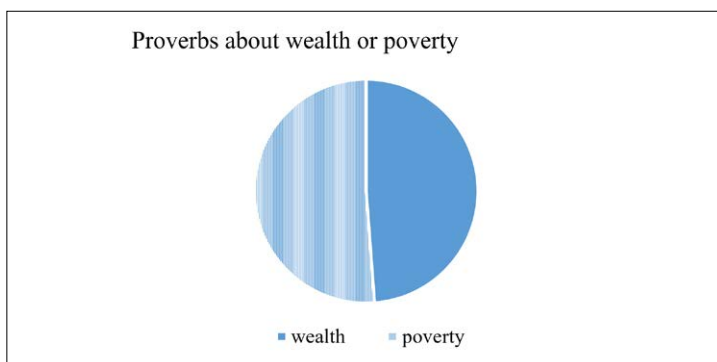


Figure 2: Components of proverbs related to wealth or poverty in the Proverbs collection.

Less than 1% of the proverbs from the Proverbs collection are included in the Dictionary of Standard Slovenian (SSKJ), and they contain seven out of the thirteen components. There are five proverbs with the component *kruh* (bread), two proverbs with either the component *denar* (money) or *zlato* (gold), and one proverb each with the components *revež* (poor person), *berač* (beggar), *sirota* (orphan) and *srebro* (silver). None of the 52 proverbs with the component *revščina* (poverty) that are in the Proverbs collection are included in the SSKJ. Restricting the research to dictionary sources would therefore give the false impression that there are not many Slovenian proverbs related to wealth and poverty, which is by no means true. The results show that collections of proverbs generally offer richer proverbial material than dictionaries, which are primarily focused on defining lexical units. It is therefore essential to promote and maintain the ongoing collection of this folklore genre in all languages, especially in languages spoken by fewer people, to ensure that their intangible cultural heritage is preserved. These findings underline the importance of preserving these cultural expressions for future generations.

3.2 Meanings and functions

The results of the semantic analysis of proverbs indicate that proverbs about wealth and poverty not only address the material aspects of human life but also discuss aspects of spiritual wealth or poverty, moralise on human qualities and the human condition (e.g. generosity, stinginess, frugality, extravagance, laziness, health, happiness, love and satisfaction), interpersonal relationships (e.g. friendship, family and children) and impart teachings about certain external circumstances such as weather phenomena, e.g. *Če sije sonce na božični dan jasno, bo novo leto bogato* (lit. If the sun shines clearly on Christmas Day, the new year will be rich).

Remarkably, the analysis revealed that Slovenian proverbs predominantly cast both extremes – wealth and poverty – in a negative light. Proverbs point to moderation in relation to material goods, for example, *Bog nas obvaruj prevelikega bogastva, prevelike brhkote in revščine* (lit. May God protect us from excessive wealth, excessive beauty and poverty). Very few Slovenian proverbs present great wealth or poverty positively, e.g. *Revščina rodi genije* (lit. Poverty produces geniuses). However, this may vary culturally from language to language. Kholmatova (2022: 172) examined the conceptual and axiological components of wealth and poverty in the Tajik language picture of the world and found that wealth is largely viewed positively, signifying both material and spiritual good. However, she also reports that in the Dictionary of Tajik Proverbs and Sayings the frequency of proverbs with a positive connotation is significantly lower than those with a negative connotation when it comes to the concept of wealth.

Slovenian proverbs about wealth teach us that values such as health, knowledge, freedom, friendship, contentment, generosity, honesty, diligence, modesty, thrift,

prudence, diligence, patience and being loved are worth more than any material wealth, e.g. *Lenoba je mati revščine* (lit. Laziness is the mother of poverty), *Zdravje je največje bogastvo* (lit. Health is the greatest wealth), *Dober prijatelj je boljši kot celo bogastvo* (lit. A good friend is better than great wealth). They warn that wealth is fleeting, while poverty can be long-lasting or even permanent, e.g. *Kdor se danes s bogastvom baha, bo že jutro kruha prosil* (One who boasts with wealth today will be begging for bread tomorrow), *Revščina je tako lepljiva, da se je ne moreš znebiti* (Poverty is so sticky that you can't get rid of it).

Unlike wealth, poverty cannot be hidden, it can be inherited and persist, and it hinders people's quality of life and progress, e.g. *Revščine in kašlja ni moč skriti* (lit. Poverty and a cough cannot be hidden). However, both the rich and the poor are prone to illness and death, e.g. *Revni in bogati smo pred smrtjo enaki* (lit. The rich and the poor are equal before death). Slovenian proverbs suggest that inheritance and large families have a negative impact on economic stability, as inheritance often leads to the dispersion of wealth and impoverishment, e.g. *Bogastvo do tretjega rodu ne pride rado* (lit. Wealth does not often last to the third generation). And also that neither wealth nor poverty makes a person happy, money corrupts both the rich and the poor, e.g. *Težko je biti reven in pošten* (lit. It's hard to be poor and honest).⁵

Proverbs warn against the negative consequences of wealth, such as greed, gluttony, arrogance, pride, covetousness, avarice, corruption and fear. As the proverbs teach us, all these are bad qualities in an individual. Proverbs point out that poverty is caused by laziness, idleness, ignorance, wealth in childhood, litigation, indebtedness and habits such as gambling, drinking and promiscuity, e.g. *Karte in steklenica so prapori revščine* (lit. Cards and a bottle are the banners of poverty). Proverbs warn that love is tested in poverty and that poverty leads to theft and crime, but it also strengthens people, makes them more creative and shows them who their true friends are, something a rich person never knows. In line with Christian teaching, the proverbs moralise that those who have sufficient material means should help the poor, e.g. *Kar daš siromaku, se povrne v nebesih* (lit. What you give to the poor will be returned in heaven), but also convey realistically that good deeds and actions are not usually rewarded with gratitude and do not bring material means, e.g. *Dobrota je sirota* (lit. Kindness is an orphan).

The function of these proverbs is usually to express a statement, a general truth concerning wealth or poverty, e.g. *Denar je sveta vladar* (lit. Money is the ruler of the world). But they can also have the function of weather sayings, as in the case of *Če je april deževen, kmet ne bo reven* (lit. If April is rainy, the farmer will not be poor). Moreover, proverbs can compare, evaluate or hierarchise the phenomenon of wealth or poverty within the system of social, moral and cultural values, e.g. *Denar človeka*

⁵ Cf. Thompson's (2009) analysis of Amharic (a language spoken in Ethiopia) proverbs about wealth. They show the presence of some similar concepts: wealth is unfair, dignity matters more than wealth, etc.

kvari (lit. Money corrupts people), *Bolje je biti berač kot škrt bogataš* (lit. It is better to be a beggar than a stingy rich man), *Poštenje več velja kakor kup zlata* (lit. Honesty is worth more than a pile of gold) or *Kdor zgodaj vstaja, mu kruha ostaja* (lit. He who gets up early will have bread to spare).

3.3 Frequency of occurrence

The data indicates a higher number of proverbs or their variants and modifications in the social media compared to other media, whether online or in print. Among the proverbs that deal with wealth, the proverb *Denar je sveta vladar* (lit. Money is the ruler of the world) is used most frequently in social media. The frequency of the proverb *Denar je sveta vladar* and its variants is three times higher in the Janes corpus than in Gigafida – almost three times higher than in printed texts and almost six times higher than in internet texts within the reference corpus (Figure 3). The proverb is included and explained in the SSKJ. According to paremiological research, this proverb also ranks first in terms of familiarity according to the paremiological optimum. Meterc reports a recognition rate of 99% for this proverb (Meterc 2017). The recent survey showed that familiarity with the proverb seems to be an important factor influencing its use in social media.

We have also confirmed with other examples that the most frequently used proverbs in the social media are those that belong to commonly known proverbs in the Slovene language, although not all of them were included in the dictionary. Although the proverb *Kdor zgodaj vstaja, mu kruha ostaja* (lit. He who gets up early will have bread to spare) is included and described in the SSKJ, we found only one confirmation in the Janes corpus and none in the Gigafida corpus which is surprising since the proverb promotes good work habits. The reason for the proverb not being used much in contemporary Slovene may be due to some socio-cultural changes regarding the value of getting up early, as shown by the recent anti-proverb *Rana ura, slovenskih fantov grob* (lit. Early hours, the grave of Slovenian lads), which comments humorously on the belief that getting up early pays off and reflects diligence by conveying the idea that getting up early is a challenge (Meterc 2020). The use of this anti-proverb in social media is more widespread than that of the original proverb *Rana ura, zlata ura* (Early bird catches the worm) (Justin et. al 2015).

The frequency distribution of proverbs may vary when examining individual examples. For example, the proverb *Ni vse zlato, kar se sveti* (lit. Not everything that glitters is gold) occurs with roughly equal frequency in the social media and in general usage (Figure 4). Upon closer examination of the concordances, we found that the frequency of use in the reference corpus is related to the fact that this proverb was used to translate the title of the movie *Fool's Gold*. This result underlines the importance of taking context into account when analysing proverbs in corpora. However, it is impressive that almost all the proverbs analysed show remarkable differences in usage

frequency between social media and internet texts. Proverbs are far more prevalent on social media platforms, with occurrences ranging from 3 to 9 times higher than in other internet contexts (Figure 4). This suggests that social media features, rather than the internet per se, are the key factor influencing the frequency of use. We have to be careful when interpreting this data, as it depends on which internet texts were included in the corpus. In Gigafida, these are mainly texts from web domains. In corpora that contain more texts from social media, we can expect a higher proportion of commonly known proverbs. It is important to take this into account if the corpus is the source for creating dictionaries or teaching materials.

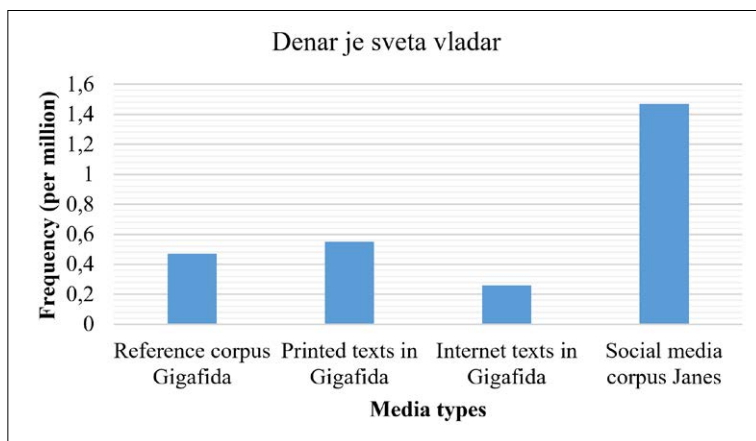


Figure 3: Frequency of use (per million words) of the proverb *Denar je sveta vladar* in different types of media.

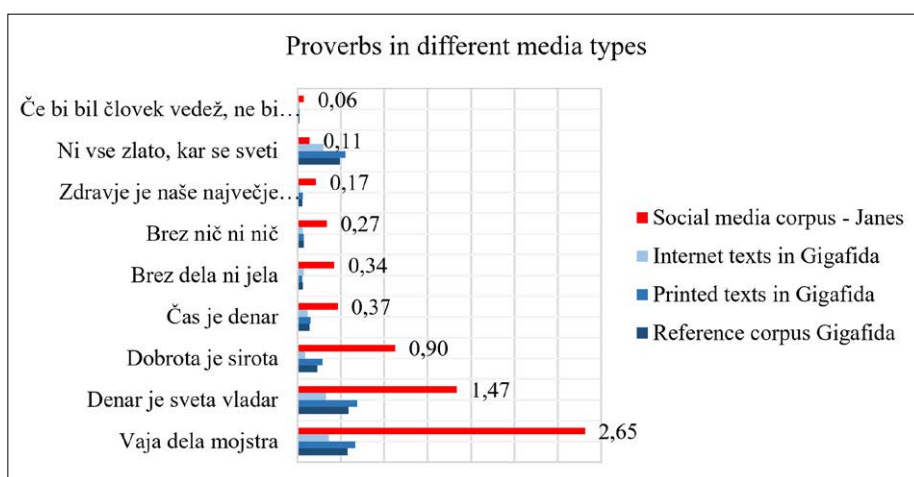


Figure 4: Frequency of proverb use (per million words) in different types of media.

3.3.1 Frequency of occurrence depending on social media type

Social media platforms have drastically changed the way people communicate, share content, interact and collaborate over the last forty years (Ahmed et al. 2019, Biały 2017). Social networks enable global connectivity and international communication flows (Eleta and Golbeck 2014). The nature and purpose of social media is to share knowledge and to communicate and this has a significant impact on the use of proverbs. Social media platforms are a medium that has a direct impact on the rapid spread of a particular proverb across different types of social networks, such as blogs, Facebook, YouTube or Twitter, through retweets, sharing content among Facebook users and so on. Proverbs usually spread very fast (within hours), which means that their use in social media can have a rapid impact on the formation of public opinion on a given topic as proverbs are a means of spreading so-called common knowledge. An interesting discovery (Ebenso et al. 2012) shows that modern technology and social networks such as Facebook provide a new forum for the dissemination and preservation of proverbs, and that traditional proverbs are not just a relic of an unchanging past, but rather an integral part of the contemporary understanding of the world. From a linguistic perspective, social media can also facilitate the rapid actualization of proverbs or their fragments that may have been long forgotten.

In our analysis, we followed the proverb *Kdor jezika špara, kruha strada* (lit. Whoever uses their tongue sparingly, goes short of bread), which a blogger used in the title of his blog post. Users of other social networks shared it on Twitter the same day and at least twice as many did in the following two days. This finding also has implications for marketing strategies, as the expressiveness of proverbs attracts social media users' attention. Users share content, including proverbs, to attract the attention of their social environment or to influence the opinion of others on a particular topic.

Our analysis found that proverbs are commonly used in various types of social media platforms, including popular social networking sites such as Facebook, microblogging platforms such as Twitter, as well as blogs, discussion forums, media sharing sites such as YouTube and comments on news articles and websites. Using the Janes corpus, we empirically examined the use of proverbs in tweets, forum posts, comments on online news portals, blogs and user comments. The relative frequency of use shows that proverbs are most frequently used in comments on web news, followed by blogs and tweets, while they are least frequently used in forums. If we consider the main purpose of online forums, which is for users to exchange opinions on the same topic, it seems that proverbs are more appropriate for comments on static content such as news articles or blog posts, and titles. These types of content, unlike forums, are not meant for dialogue with other users. They are mainly intended for providing one-way responses to static content (Dobrovoljc 2018).

We confirmed the use of all six tested proverbs on all three social media platforms: Twitter, Facebook and YouTube. This indicates that the different types of social media

are equally suitable for the use of proverbs. One might assume that Twitter is less suitable for longer proverbs due to its character limit. However, our results suggest the opposite. Our analysis found that the length of a proverb has no significant impact on its use in social media. Both long and short proverbs are used across different types of social media. The longest proverb in our analysis, which comes from the Bible (Matthew 19:24), *Lažje pride kamela skozi šivankino uho kot bogataš v božje kraljestvo* (lit. It is easier for a camel to go through the eye of a needle than for a rich man to enter the kingdom of God), was used across all types of social media, including Twitter. Proverbs are chosen for their contextual multi-functionality and their ability to convey a message quickly and effectively (collective memory). The communicative power of proverbs is so significant that a tweet or an online comment can effectively convey a message by using a single proverb, a fragment of a proverb, or an innovative modification. Despite being brief, proverbs have the ability to effectively convey a complete idea or message. In addition, by using a proverb the person distances themselves from what is being said. Distance and the authority of ‘general opinion’ are very important elements that influence the use of proverbs on social networks. In this respect, proverbs in social media adhere to two cooperative principles (Grice 1975): the maxim of quantity, which mandates conveying enough information without redundancy (‘Do not make your contribution more informative than is required’), and the maxim of relevance, which requires the message to be contextually applicable (‘Make your contribution relevant’). Due to their conciseness, proverbs serve as a functional linguistic tool in the social media, allowing users to express and affirm their thoughts, opinions, beliefs and values in just one sentence. This is also supported by the findings concerning the situation of the Uyghurs, showing that blogs and comments often used veiled language such as metaphors, sarcasm and humour, or references to traditional Uyghur proverbs and cultural practices, to express subversive political communication in an indirect way (Clothey et al. 2016). Our results are also in line with some other findings on the use of proverbs in Arabic tweets, which show how Twitter enables individuals everywhere to spread their ideas around the world (Al-Wehaibi and Khan 2015).

3.4 Gender and sentiment

Both men and women use proverbs on social media. Although the number of male users in the Janes corpus is larger than the number of female users (about 60% men), the proverb *Denar je sveta vladar* was used slightly more often by women (based on its normalised frequency). (Figure 5).

Based on our corpus data analysis, we found no evidence to support the claim that women talk more about money on social media than men. It would therefore be wrong to make such an assumption. However, we found that women tend to use certain proverbs more often than men, as in the case of the proverb *Vaja dela mojstra* (lit. Practice makes perfect), which is not related to the topic of wealth and poverty and which we

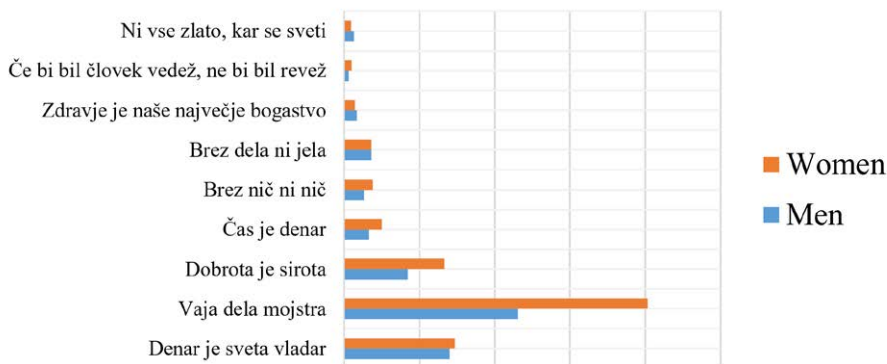


Figure 5: Use of proverbs according to gender.

only use as a control in our analysis. Our analysis of nine proverbs showed that six of them had a slightly higher percentage of female users. Further research would be needed to confirm and interpret this phenomenon. The study by Hussein (Hussein 2009) has shown that gendered ideology is discursively formulated in some Ethiopian, Kenyan and Sudanese sexist proverbs.

The sentiment in which a particular proverb is used could be an indicator of the contexts in which users employ proverbs, whether in positive or negative contexts. We have noticed differences among proverbs. For example, the proverb *Denar je sveta vladar* (lit. Money is the ruler of the world) is predominantly used in texts with a negative sentiment, where users comment critically on phenomena such as bribery, greed, capitalism, power, profit, development and even the winner of the Eurovision Song Contest, implying that money can buy victory, which is judged as immoral. The proverb appears in social media in the context of current issues related to financial aspects in politics, healthcare, justice, education, business, sports (football, etc.), banking, the military, the Church, love, the automotive and pharmaceutical industries, and so on. When the proverb is used in these contexts, it serves to highlight corruption and to evaluate it negatively. Using a proverb emphasises the moral and ethical implications of such actions. This suggests that proverbs play an important role in conveying social values and reinforcing ethical behaviour.

All the examined proverbs from the realm of wealth-poverty predominantly carry negative sentiments when used in social media. An exception is the proverb *Dobrota je sirota* (lit. Kindness is a beggar), which is predominantly used in texts that carry a positive sentiment. However, deeper analysis of the corpus material has shown that the proverb is mainly used to express irony or sarcasm, i.e. as a negative comment on a topic. Sentiment recognition is particularly difficult in tweets and therefore not entirely

reliable because they contain insufficient context. The user often suggests a positive or negative sentiment in social media texts with the help of emoticons, punctuation, modifications, and so on.

We can generally say that proverbs referring to wealth and poverty appear in social media primarily to support a negative textual atmosphere, as an outlet for expressing anger, frustration, grievances and other negative emotions. This maintains the primary rhetoric of wealth and poverty in proverbs, which as Sandoval (Sandoval 2006) argues, ‘serves particular moral goals and functions to order, justify, and even prevailing social arrangements that embrace matters of economic practice and distribution’ (Holloway 2008).

3.5 Variations and modifications

Among the proverbs analysed, we observed a greater number of variations and modifications in their use on social media than in their general use in traditional media discourse. Compositional modifications are often used to update proverbs with modern lifestyle elements. For example, the proverb *Kdor jezika špara, kruha strada* (lit. Whoever uses their tongue sparingly, goes short of bread), which means that ‘if you are too careful with your words and don’t speak up when necessary, you may miss opportunities to earn a living or improve your situation’ undergoes contemporary modifications in the context of social media: *Kdor lajka/retvita špara, kruha strada* (lit. Whoever fails to give likes or make retweets, goes short of bread). The components *lajk* (like) and *retvit* (retweet) are employed instead of *jezik* (which means both ‘language’ and ‘tongue’ in Slovene). What is more, the use of an emoticon with a tongue can replace the word component: *Kdor 😛 špara, kruha strada*. Such modifications of proverbs in the social media can serve to express a socially relevant issue in an expressive way. The above findings highlight the adaptability of proverbs to changing social and cultural circumstances, ensuring their continued relevance and efficacy in imparting wisdom, teaching and transmitting social values.

One of the important contextual features we found in our material, both in social media texts and in general language use, is the employment of textual introducers (Čermák 2005). This is where proverbs in texts are introduced with phrases such as *kot pravi pregovor, rek* (lit. as the proverb says, as the saying goes), *pravijo* (lit. they say, it is said). In social media, however, many examples exist where statements are not actually established proverbs. Instead, users want them to function as such, serving as expressions of wisdom or as thoughts that are socially endorsed within a broader framework, as a source of collective wisdom and authority. References to other paremiological systems are also common, e.g. *kot pravi kitajski/latinski/nemški/nizozemski pregovor* (lit. as the Chinese/Latin/German/Dutch proverb says). This finding suggests that the use of proverbs by social media users is a conscious choice.

4 Conclusion

Proverbs are ready-made mini-texts carrying powerful, universal messages that enable social media users to effectively express their opinions and beliefs. According to their inherent characteristics, proverbs offer users the opportunity to quickly verbalise their thoughts and share them with others. This makes proverbs a valuable tool for communicating on social media, where brevity is crucial for the successful communication of ideas and opinions. Social media has certain characteristics that influence the use of proverbs, including interactivity, creativity, content sharing, virtual community building, socialisation, global reach and internationality.

Our research has shown that proverbs are used more frequently in social media than in other media. Put simply, you are more likely to come across a proverb or its modification when using a social network than when reading a printed or online newspaper or magazine.⁶ However, the use of proverbs in social media is influenced by several intertwined factors, such as general familiarity with the proverb, the context and the current social circumstances. The inclusion of a proverb in a dictionary is not a guarantee of its use in social media and even the absence of a proverb in a dictionary does not necessarily exclude its use in social media.

The use of introductory devices shows that users consciously choose proverbs to add an element of collective behaviour, approval and support to their public message. Current social media issues influence the modification of proverbs. Proverbs on social media are often accompanied by additional para-linguistic elements, such as emoticons to increase the clarity of positive or negative sentiment or just as creative language play. The role of proverbs in social media is to attract attention with expressed general opinions and thus influence the public opinion of other social media users. It is also noticeable that socio-political and culturally relevant messages are highlighted and spread with the help of proverbs.

Acknowledgements

This paper is a result of the scientific projects *Traditional Paremiological Units in Dialogue with Contemporary Use* (2020-2023, J6-2579), *Language, culture and values: economic images of everyday life in folklore forms* (2023-2026, J6-50197), and the programme *Slovenian language in synchronous and diachronic development* (2022-2027, P6-0038).

⁶ The study (Granbom-Herranen, Babič, Voolaid 2015) which examines the use of proverbs in contemporary Estonian, Finnish, and Slovenian newspaper contexts and in everyday use found out that proverbs are used there as rhetorical tools, but that the Slovenian newspaper included many fewer proverbs than the Estonian and Finnish ones.

References

- Ahmed, Yunis Ali; Ahmad, Mohammad Nazir; Ahmad, Norasnita; Zakaria, Nor Hidayati, 2019: Social Media for Knowledge Sharing: A Systematic Literature Review. *Telematics and Informatics* 37, 72–112. DOI: <https://doi.org/10.1016/j.tele.2018.01.015>.
- Al-Wehaibi, Rehab Nasser; Khan, Muhammad Badruddin, 2015: Investigate the Context Usage of Arabic Proverbs in Twitter. *International Conference on Cloud Computing*, 1–8. DOI: <https://doi.org/10.1109/CLOUDCOMP.2015.7149645>.
- Babič, Saša; Erjavec, Tomaž, 2022: Izdelava in analiza digitalizirane zbirke paremioloških enot. *Conference on Language Technologies & Digital Humanities Ljubljana, Slovenia*, 17–22.
- Babič, Saša; Peče, Miha; Erjavec, Tomaž; Ivančič Kutin, Barbara; Šrimpf Vendramin, Katarina; Krojež Telban, Monika; Jakop, Nataša; Stanonik, Marija, 2022: Collection of Slovenian paremiological units Pregovori 1.0. Slovenian language resource repository CLARIN. SI. Online: <http://hdl.handle.net/11356/1455>.
- Biały, Beata, 2017: Social media—from social exchange to battlefield. *The Cyber Defense Review* 2/2, 69–90.
- Clothey, Rebecca A.; Koku, Emmanuel F.; Erkin, Erfan; Emat, Husenjan, 2016: A voice for the voiceless: online social activism in Uyghur language blogs and state control of the Internet in China. *Information, Communication & Society* 19/6, 858–874. DOI: <https://doi.org/10.1080/1369118X.2015.106157>.
- Čermák, František, 2005: Text introducers of proverbs and other idioms. *Jezikoslovje* 6/1, 57–77.
- Dobrovoljc, Kaja, 2018: Raba tipično govornjenih diskurzivnih označevalcev na spletu. *Slavistična revija* 66/4, 497–513.
- Ebenso, Basse; Gbemiga, Adeyemi; Adegoke, Adegboyega O.; Emmel, Nick, 2012: Using indigenous proverbs to understand social knowledge and attitudes to leprosy among the Yoruba of southwest Nigeria. *Journal of African Cultural Studies* 24/2, 208–222. DOI: <https://doi.org/10.1080/13696815.2012.704263>.
- Eleta, Irene; Golbeck, Jennifer, 2014: Multilingual use of Twitter: Social networks at the language frontier. *Computers in Human Behavior* 41, 424–432.
- Fišer, Darja; Erjavec, Tomaž; Ljubešič, Nikola, 2016: JANES v0.4: Korpus slovenskih spletnih uporabniških vsebin. *Slovenščina 2.0: empirical, applied and interdisciplinary research* 4/2, 67–99. DOI: <https://doi.org/10.4312/slo2.0.2016.2.67-99>.
- Fišer, Darja; Ljubešič, Nikola; Erjavec, Tomaž, 2020: The Janes project: language resources and tools for Slovene user generated content. *Lang Resources & Evaluation* 54, 223–246.
- Fišer, Darja; Smailović, Jasmina; Erjavec, Tomaž; Mozetič, Igor; Grčar, Miha, 2016: Sentiment annotation of Slovene user-generated content. *Proceedings of the 2016 conference language technologies and digital humanities*, 65–70.
- Fran, dictionaries of the Fran Ramovš Institute of the Slovenian Language ZRC SAZU, 2014- (version 8.0). Online: <https://fran.si/>.
- Granbom-Herranen, Liisa; Babič, Saša; Voilaid, Piret, 2015: Proverbial expressions in newspapers. *Traditiones* 44/3, 5–32.
- Grice, Paul H., 1975: Logic and Conversation. In: Cole, Peter; Morgan, Jerry L. (eds.), *Syntax and Semantics 3: Speech Acts*. New York: Academic Press, 41–58.
- Holloway, Jeph, 2008: The Discourse of Wealth and Poverty in the Book of Proverbs. *Journal of the Society of Christian Ethics* 28/1, 257–259. DOI: <https://doi.org/10.5840/jsce200828141>.

- Hussein, Jeylan Wolyie, 2009: A discursive representation of women in sample proverbs from Ethiopia, Sudan, and Kenya. *Research in African Literatures*, 96–108.
- Jakop, Nataša, 2022: Vpliv družbenih omrežij na rabo paremioloških enot: na primeru enot s sestavino *kruh*. *Slavistična revija* 70/3, 349–360. DOI: <https://doi.org/10.57589/srl.v70i3.4042>.
- Justin, Martin; Hirci, Nejc; Gantar, Polona, 2015: Rana ura, slovenskih fantov grob: analiza frazeoloških prenovitev v spletni slovenščini. Fišer, Darja (ur.), *Zbornik konference Slovenščina na spletu in v novih medijih* 25. Ljubljana: Založba Univerze v Ljubljani, 33–37.
- Kholmatova, Khursheda Ayubjonovna, 2022: About of the Concept ‘Wealth and Poverty’ in the Paremiological Picture of the World, *Research and Education* 1/8, 172–175.
- Krek, Simon et al., 2019: Corpus of Written Standard Slovene Gigafida 2.0, Slovenian language resource repository CLARIN.SI. Online: <http://hdl.handle.net/11356/1320>.
- Liu, Bing, 2012: Sentiment analysis and opinion mining, *Synthesis lectures on human language technologies* 5/1, 1–167. DOI: <https://doi.org/10.2200/S00416ED1V01Y201204HLT016>.
- Meterc, Matej, 2017: *Paremiološki optimum: najbolj poznani in pogosti pregovori ter sorodne paremije v slovenščini*. Ljubljana: Založba ZRC, ZRC SAZU. DOI: <https://doi.org/10.3986/9789610504153>.
- Meterc, Matej, 2020–: *Slovar pregovorov in sorodnih paremioloških izrazov* [lit. Dictionary of proverbs and similar paremiological expressions]. Online: <https://fran.si/216>.
- Mieder, Wolfgang, 1996. *American folklore: An encyclopaedia*. New York: Garland Publishing.
- Mieder, Wolfgang, 2004: *Proverbs: A handbook*. Greenwood Publishing Group.
- Mieder, Wolfgang, 2005: “A proverb is worth a thousand words”. Folk wisdom in the modern mass media. *Proverbium* 22, 167–232.
- Sandoval, Timothy J., 2006: *The discourse of wealth and poverty in the book of Proverbs* 77. Leiden: Brill.
- Thompson, Tok, 2009: Getting Ahead in Ethiopia: Amharic Proverbs about Wealth. *Proverbium* 26, 367–386.
- Ulčnik, Natalija, 2015: Razvoj slovenske paremiologije in paremiološke terminologije. *Slavia Centralis* 1, 105–118. DOI: <https://doi.org/10.18690/scn.8.1.105%E2%80%9393118.2015>.

'You are where you post': Changing identities through online social networks

»Si tisto, kjer objavljaš«: spreminjanje identitet na spletnih družbenih omrežjih

This paper examines identity change through online social networks, using personal experiences as the research lens. The author discusses the transition from observer to active participant, highlighting in particular the challenges he faced when deleting his Twitter profile. Through a detailed analysis of the process, the chapter explores the preparations for 'digital suicide', the reactions of the online community and beyond to this act, and the lasting effects of adopting a new digital persona on alternative platforms. This narrative of the paper highlights the complex relationship between 'online' and 'offline' identities and offers insights into self-renewal and changing identities in the digital age.

• **Keywords:** online social networks, digital age, changing identities, Twitter, Mastodon, Threads

Članek obravnava spreminjanje identitete prek spletnih družabnih omrežij, pri čemer kot raziskovalni objektiv uporablja osebne izkušnje. Avtor obravnava prehod od opazovalca do aktivnega udeleženca, pri čemer izpostavlja zlasti izzive, s katerimi se je soočil pri brisanju svojega profila na Twitterju. S podrobno analizo procesa poglavje raziskuje priprave na „digitalni samomor“, odzive spletne skupnosti in širše na to dejanje ter trajne učinke sprejetja nove digitalne osebe na alternativnih platformah. Ta pripovedni del članka poudarja zapleteno razmerje med „spletnimi“ in „nespletnimi“ identitetami ter ponuja vpogled v samoobnovo in spreminjanje identitet v digitalni dobi.

• **Ključne besede:** spletna družbena omrežja, digitalna doba, spreminjanje identitet, Twitter, Mastodon, Threads

1 Introduction

In 2007, I created a Twitter account and posted a few experimental texts, which didn't garner any engagement. This was followed by several years of inactivity. I occasionally tried to revive my profile by posting from time to time, to no effect, until 2020, which brought a turning point. I started to post more and my posts subsequently led to more engagement. I realised they were becoming more and more resounding, read and shared by an increasing number of users, so I started posting even more – which led to even more engagement and followers. Simultaneously, and lacking an in-depth qualitative analysis along the lines of the one my colleague Ajda Pretnar Žagar and I carried out on the Instagram profile of the former Slovenian president, Borut Pahor (Podjed, Pretnar 2018; Podjed 2019: 81–88), I was learning about which posts attracted the most attention, were the most interesting, most shared and most wanted. Appeasing and gentle comments rarely fit into this category, unlike harsher and more combative,

cynical and sarcastic ones. Since such posts were met with the most digital ‘applause’ in the form of shares and likes, I continued the practice of posting captivating content while concurrently realising that the point was not only in their quality, but also in their quantity. More posts meant more comments, which I attempted to answer regularly, rejected the ones contradicting my statements, and supported the ones defending them.

After two and a half years of rigorous posting, I found Twitter increasingly captivating and repulsive at the same time. I removed a growing number of irritating commentators, who were corroding my daily life with insults and sarcastic remarks, from my ‘filter bubble’ (Pariser 2011), by either silencing them, that’s to say using the *mute* option to insure I could no longer see what others were writing about me, or by blocking them, thus preventing them from seeing my own posts, while I stopped following their posts, as well.

This paper focuses on the process of leaving the Twitter network, renamed to X in 2023, explains the community’s reactions both online and offline, and identifies long-term consequences of terminating a digital persona and moving to other platforms, with special emphasis on the take-over of Twitter by billionaire Elon Musk in the period from 2022 to 2023, and also describes subsequent changes of the network itself and the people constituting it.

2 The beginning of technofeudalism

The middle of 2022 saw the start of a change in the shareholders of Twitter. In April of that year, Elon Musk, one of the richest people on Earth, announced the take-over of the network. The *Twittersphere*, i.e. Twitter’s active users, burst with unease, much more so in October of that year, when the announcement was realised. For 44 billion dollars, he had bought a network with approximately 400 million ‘souls’ at the time, although many of them were – in the words of Gogol’s novel – ‘dead souls’, i.e. fake accounts and so-called *bots*, user profiles managed by programmes instead of real people. At that moment, the actual value of an individual user of the network became clear: it was around one hundred Euros. That was the amount of money the magnate had bought his own kingdom for and transformed into the ‘king of our digital world’, as I commented in an article published by the *Delo* newspaper at the time of Musk’s take-over (Be. B. 2022). I was alluding to Yanis Varoufakis’s writing (2021, 2023) on ‘technofeudalism’, where he explained that in neo-feudal economy – unlike in capitalism – all social, political and, most importantly, economic power falls into the hands of a small, yet extraordinarily influential elite, which has total power and control over the markets. Analysing the operations of online giants like Meta, Amazon and Google, Varoufakis finds that the digital ‘scenes’ created by these elites will become even more concerned with the needs and wants of individuals or groups of selected few. This will

de facto end any semblance of (neo)liberal or free trade markets, pushing them toward feudal markets with strict executive control over all activities and transactions (cf. Podjed 2022a). In one of his articles, Varoufakis articulated this transition from one system to the next as follows: 'This is how capitalism ends: not with a revolutionary bang, but with an evolutionary whimper. Just as it displaced feudalism gradually, surreptitiously, until one day the bulk of human relations were market-based and feudalism was swept away, so capitalism is today being toppled by a new economic mode: techno-feudalism' (Varoufakis 2021).¹

Musk's over-taking of Twitter was a clear sign that the new social-economic system was actually on the horizon – or already underway, without a bang. For it was one man who, almost overnight, took the reign over the entire digital 'kingdom', i.e. an online social network, and assumed control over every decision concerning the way this community cohabited. He even decided on renaming an entire digital world, from Twitter to X, which happened in June 2023, and simultaneously changed the typical colour scheme, replacing the blue and white one with black and white, while a black letter X forced out the logo of a blue tweeting bird. Musk introduced numerous changes to the network, in order to better align it with his value system, style of thinking and managing companies, as well as his general way of living and reigning (for more on this see Isaacson 2023). He began, among other things, stressing the importance of 'freedom of speech', using it as grounds for reinstating profiles of individuals formerly blocked by Twitter, including the controversial rapper Kanye West, who officially assumed the name Ye in 2021, and the former American president Donald Trump, whose posting was disabled after inciting rioters in front of the Capitol building in January 2021. Prior to the 'reviving' of Trump's profile, Musk posted a poll on his Twitter account, asking the community whether the former American president should be given another chance. As it turned out, the majority, specifically 51.8 percent of approximately 15 million voters, supported the idea, and Elon Musk tweeted: 'The people have spoken. Trump will be reinstated.' He added a weighty Latin saying: '*Vox Populi, Vox Dei.*' One of the first written mentions of this saying is related to English scholar and clergyman Alcuin (735–804), who mentioned the following in one of his letters: '*Nec audiendi qui solent dicere, Vox populi, vox Dei, quum tumultuositas vulgi semper insaniae proxima sit.*' In short, he was of the opinion that the people who keep saying the voice of the

¹ The term 'technofeudalism' by Varoufakis might be exciting, but disputable in many aspects and calls for further questioning and analysis of the current situation (see e.g. Krašovec 2021) and past socio-economic systems. One of the problems is that the new system really doesn't include some of the basic elements of 'old' feudalism, among them compulsory labour and handing over tributes, although – as I discussed elsewhere (e.g. Podjed 2022a) – one could see the use of smart phones and other devices as far from voluntary. For this reason, it might be a good idea to further discuss the role of data-serfs. Constant surrendering of data with the help of 'addictive' technological solutions (Alter 2017; Eyal 2019; Hari 2022) facilitating connections and communication, which are in fact comparable with tracking and listening devices (Podjed 2019), as these people labour daily on digital platforms and, with the help of technological solutions, keep expanding the wealth of a select few, who are getting rich through collecting and managing data.

people is the voice of God should not be listened to, since the riotousness of the crowd is always very close to madness (Yeo 2022). Interestingly, when used in contemporary contexts, the phrase highlights the opposite of the original thought. Of late, the people are always right, so the ruler must follow. The voice of the people is supposedly the ‘voice of God’, but in the case of Musk apparently only as long as ‘god’ concurs with this ‘voice’. This naturally begs the question of who has actually taken over god’s role in the case of today’s X network, previously known as Twitter? Has the role of god in an online form of quasi-democracy been taken over by the people, who even have the power to revive ‘dead souls’? Or has god’s and therefore all-knowing and all-seeing position been occupied by the network’s owner, i.e. Elon Musk? Judging from actions that have followed, one could conclude that the latter is the correct answer. Musk, a self-proclaimed ‘free speech absolutist’ in the past, removed rapper Kanye West’s profile from his network in December 2022, following a series of anti-Semitic posts and statements. However, West’s profile had been blocked several weeks prior, in October 2022, but was later once again allowed to post. The final straw was a post of the Star of David, the symbol of Judaism, merged with a Nazi symbol – the swastika. At that point, Kanye West was deleted. Eight months later, the rapper under a new name of Ye was ‘reincarnated’ on the also newly named network, X, the difference being that his new digital existence no longer depended on the will of the ‘people’, i.e. ‘*populous*’, but the final say was left to *Deus* himself, i.e. Elon Musk in the role of god and arbiter.

3 Schismogenesis online

The arrival of a new ‘king’, Elon Musk, who bought his own kingdom, subjects included, certainly contributed to my thinking about deleting my Twitter account. I was also annoyed with ceaseless harassment and taunting by those who obviously shared neither my ideologies nor my ideas online. But these two reasons alone were not enough for deletion and self-abolition. I was fundamentally more bothered by something that remained hidden, yet was extremely harmful to our society: division among people, who could perhaps coexist in the physical world, but seemed to hate each other’s guts in the digital one. I articulated this in an article published by the *Večer* newspaper:

Division of people into groups, yelping at one another like rabid dogs through wire fencing, is not a mere coincidence but rather the essence of a business model behind online social networks, based on tribalisation, i.e. division into ours and yours, good and bad, left and right. Remaining moderate and neutral on Twitter is a nearly impossible task, as others put you in a position, where you have to defend and uphold your own statements, no matter how benign. In short, one is guilty of existing. (Podjed 2022)

As I attempted to explain later on in the abovementioned commentary, separated and isolated communities, which had been forming on Twitter, were an excellent example

of *schismogenesis*, a phenomenon first defined by anthropologist Gregory Bateson in the 1930s while analysing social behaviour of an ethnic group in New Guinea, whose members strived to differentiate as much as possible from members within the same group, perceived as inferior or as opponents. According to Bateson (1935, 1999), such focusing on differences and emphasizing one's own superiority both in words and in actions, even in style of clothing, leads to a rift in a community. Such a rift had been – by the looks of it – also happening on Twitter, where Slovenian users started to transform into 'Slevenci' (left-leaning Slovenians) and 'Sdesnenci' (right-leaning Slovenians), based on their alleged political positions and preferences, as I stated in my commentary for the *Večer* newspaper, or even into more tribal-sounding 'Levuharji' (for left-leaning individuals) and 'Desnosučne' (for right-leaning individuals) (Podjed 2022). As I added, they 'were standing on opposite sides, targeting one another with verbal spears, occasionally penetrating the digital fabric and breaking through into the physical world' (ibid.). This war, waged mostly verbally on the former Twitter, currently X, and other social networks, of course, validated social psychologist Jonathan Haidt's hypothesis, presented in his book *The Righteous Mind*, namely that we're increasingly turning away from looking for our opponents outside of our community, finding them within instead. A liberal citizen's worst enemy is thus one who defends conservative values in their own country – and vice versa. The fact that we are part of the same ideational community, the same nation, is thus slowly falling into oblivion. Which side of the ideology spectrum we are on has become more important (Haidt 2012).

Dividing people into groups which get angry with each other and bombard each other with insults is more than just a coincidence, it is actually the essence of the business model behind online social media, based on tribalisation. Schismogenesis is at the heart of tribalisation, as is the rule saying that 'we are not them'. Schismogenesis, or deepening of divisions among different groups, has been accelerated by algorithms and ensured broader exposure to the most contentious posts, i.e. the searing, offensive and hatred stimulating ones. In this way, an online network merely imitates and bolsters 'algorithms' already in people's minds. Every one of us notices posts which stand out from the crowd first, and these are usually shocking, surprising and, naturally, bursting with hatred towards the other side. When people come across such posts, they more often than not write a vexed comment, like or share the post themselves. And the vicious circle of mutual hatred continues, gaps in a community widen, and people demonise each other further, at the same time turning into demons in their own 'echo chamber', which gradually transforms into an 'epistemic bubble' (Nguyen 2020).

In the aforementioned article for the *Večer* newspaper, I stressed that the three aforementioned reasons – behaviour of others, arrival of a new 'king' and schismogenesis – might not be enough for me to delete the former Twitter account. The last straw was the realisation that I was becoming someone who tweeted not only online, but live, as well. I had been gradually changing in line of the principle 'you are what you post'.

By saying this, I'm paraphrasing a famous saying 'you are what you eat', and by the way also touching on a lesser known one, 'you are what you drive'. As Saša Babič and I have found – in the wake of Tim Dant (2004) – from research on transportation habits, a car can be a key identity marker, 'merging' with the owner into a temporary formation, an assemblage which could be called 'driver-car' (Podjed, Babič 2015; Babič, Podjed 2016). We have established that people in vehicles change their identity and behave differently in traffic when they are walking on pavement, driving a bicycle or sitting behind a wheel. And it doesn't stop there: they are characterised by the type of their car. Therefore, it is not insignificant whether they are sitting in a large sedan, a SUV or a smaller vehicle with hardly enough space for four people. As we noted in our research, the car type influenced even the forms of expressing anger and other emotions inside the vehicle, as well as through the vehicle, e.g. flashing lights, honking and taking off with screeching tires.

I witnessed something similar to the findings of the abovementioned research on traffic habits and transformation of drivers during bouts of road rage in 2022, in the streets and other public places, in increasingly specific instances, though the people's identity didn't depend on the vehicle, but instead on the networks and media which transformed the society. A watershed moment for this realisation was a seemingly insignificant one, but it stayed with me for a long time. I articulated it in the aforementioned article for the *Večer* newspaper as follows:

The one drop too many foretelling the flood was a fleeting moment, which happened approximately a month ago, as I was waiting in front of a car park for a friend, a public figure because of his professional work, media appearances, as well as for his Twitter posts, which had been – in my opinion – legitimate and reconcilable, never offensive. I was standing next to a fence and waved hello from afar as he was getting out and taking his parking ticket. At that moment, his face fell as he heard a harsh remark of a stranger rushing by and snapped at him along the way. It was not about content it was about the way the words were spoken. Harshly, offensively, thoughtlessly. We went for a walk through the city and it took both of us a few minutes to regain composure. Once he calmed down I asked him if this happened a lot. 'Almost every day,' he confirmed, adding that people had obviously started tweeting live (Podjed 2022b).

While such incidences left me in shock over where society was heading, I was still more or less unscathed by the 'digital transformation', which started leaving traces in the physical world, as well. The one drop too many was, as written above, of an entirely personal nature:

I noticed some time ago that tweeting had made me start to transform in line with the principle 'you are what you post'. Just like people are characterised by the vehicle they drive, they are similarly characterised by the network they primarily use. If someone drives a German brand black sedan, they will behave differently than if they were driving a bicycle. And if someone predominantly uses Twitter, they will eventually start thinking in 'twitter' and forming short, powerful sentences in everyday life, as well. (Podjed 2022b)

I continued that my increasingly obsessive thinking about my next tweet and checking online reactions to my posts led to my having less and less time for myself and others. I caught myself ever more often in a similar situation than Nir Eyal. In his book *Indistractable* he writes about his pre-school daughter who, during play-time, wanted to answer a question, which was important to her. He asked her to wait a moment, as he just received a push notification and started to answer. Before he was done, his daughter left, and he missed an unrepeatable moment – spending uninterrupted time with someone who was supposed to be the centre of his attention (Eyal 2019).

When I, too, realised that more and more important things and people slipped out of my focus, I decided to delete my profile. First, I wrote the aforementioned text for the *Večer* newspaper, as a sort of an 'obituary' at the occasion of parting with my own digital double, and then, on Wednesday, 9 November 2022, I posted my last tweet, announcing that I would be terminating my Twitter profile on Sunday – and providing a link to the longer farewell text in the form of the aforementioned article. What followed was astonishing. My last post, which many had shared, as well, was generating numerous comments, likening me to a 'snowflake', who apparently got scared of posting and threw in the towel instead of continuing to participate in the endless 'game' of who was going to surprise, insult, hurt or taunt someone else with their posts and comments the most. It struck me as especially interesting that my leaving the social network was the most upsetting to those who had called on me the most to delete my account. It seems that my presence wasn't as disturbing for them as they tried to make it appear with their comments. Perhaps it was altogether the other way around: my online presence gave meaning to their digital existence, and by my bidding farewell their existence lost a little bit of their purpose.

The next surprise regarding my farewell from Twitter was how technically challenging the process was. Considering that creating a profile was relatively easy, the 'digital suicide', as deleting one's existence on the network could be called, was far more difficult. On the Twitter version of the time it was not enough to merely click on the 'deactivate account' option, which should erase all past activity, it was necessary to delete every digital trace 'by hand', that is delete older posts and reposts. Because doing all these 'by hand' would have been a very lengthy process, I turned to software tools to do that for me, but quickly realised that the free version had an upper limit of posts per day it could delete. This led me to buying the paid version, which enabled me to erase old posts.

When I – presumably, at least – erased all my tracks, as I didn't want anyone to register using my username in the future and be able to manage my posts, I was facing a new dilemma. After deactivating my account, Twitter let me know that there was a limited time period within which I could reactivate my profile. I expected the process of erasure to take a day, but it was thus extended to several weeks of insecurity, during which my digital double – akin to *Schrödinger's cat* – was alive and existing, while at

the same time it had already left the digital world and was dead; somewhat like being stuck in some sort of purgatory. To be or not to be – that was the question in this case, and like me, other users of the network were perplexed by it, as they were leaving Twitter at the time and later X in droves, searching for a new haven, a new community that would allow them to come together and socialise. To be alone, isolated, expelled from the community, albeit on one's own initiative, was becoming increasingly unbearable in the digital world, as well.

4 Arrival to the new digital haven

I was not alone in my search for a new digital shelter. In November 2022, right after Musk's official Twitter takeover, large numbers of people started to leave the network and searched for another place elsewhere. Mastodon was one of the possible havens, but it's organised differently than the former Twitter, now X. Instead of a centralised network, Mastodon is actually a federation of different separate networks, which are at the same time loosely connected, intertwined. A vast number of people joined this 'network of networks' in a short amount of time. As Eugen Rochko, the creator of Mastodon and the only full time employee at the time explained, the network hosted 2.5 million users per month on 8,600 servers. During the 'exodus', the number rose quickly; the network registered around four thousand downloads of application per day, and as many as 149 thousand per day for Android operating system, and 235 thousand for iOS system during a short peak. Rochko explained that the leap coincided with the time Twitter announced mass layoffs, eliminating entire departments in the company, including those responsible for public relations, and those working on post-moderation and safety (Lunden 2022).

On 9 September 2022 I set up a profile on Mastodon, or rather on a federal 'national' subnetwork called Toot.si, which appealed to me and was also used the most during the 'move'. I posted my first series of short texts on 18 December. They mostly covered my feelings at entering this new 'network space' (cf. Podjed 2010, where I wrote about the formation and development of social networks). The first three posts were Twitter-like in that they were short, but also somewhat sarcastic and (self)ironic, which had been the norm on the network I left behind. First I wrote, self-ironically referring to my baldness: 'Here, too, I have to answer the question, 'What's on your mind?' It's a good thing I don't get asked 'What's on your head?'' That's what I wrote in the comment window on Toot.si, in answer to the question to users of the network: 'What's on your mind?' This first post already garnered response from several 'deserters', greeting me in the comments and advising me to set my interface language to Slovenian. Soon after that, I wrote: 'Meeting on a new network is somewhat reminiscent of a high-school reunion: same discussions, older profile pictures.' As I recollect, I used these words to greet my colleagues who migrated to the new digital realm from Twitter, and

let them know we were still peers, albeit aged a bit. Again this time a user I used to correspond with on Twitter chimed in, commenting that I must have substituted one drug for another when I switched to Toot.si. I responded by saying that I was actually on other drugs, as well, including Facebook, Instagram and LinkedIn, making Toot.si a 'recreational drug', at best.

My third text, which I posted that same day, already caused the first conflict, but not between me and 'the people', also known as everymen who used to bicker on Twitter about everything. I received a mild warning from the network admin, the 'first among equals'. He warned me that I should have lowered my expectations regarding the network I had just moved to. This was in response to something I wrote in a post because the interface seemed clumsy to me: 'I mean, I'm not going to complain about the interface exactly, but I would still like to know, where we are going from here. To IRC? For a beer?' The first responses came from 'everymen', not because my remark annoyed them; they had different drinks in mind for us to hang out over. But the admin wrote, albeit calmly, and also being a tiny bit patronising, that Mastodon is an open source network, and to check the issue tracker for possible previous mentions of 'this inconvenience that has been bothering you'. If nothing along these lines has been written there yet, I should start a new issue, i.e. a new topic or problem and present my opinion in detail. 'This is the most constructive manner of complaining,' the admin ended. I was somewhat offended by this message, as I was not used to such comments from admins at the network I had left. On the contrary: there was actually a complete absence of comments or warnings from the admin's position. Despite this introductory misunderstanding, those first days on the new network seemed romantic, something that I wrote and posted the following day, that is on 19 November 2022. I mentioned that I liked the 'romantic side of revamped relationships,' adding that, 'Words of encouragement, mutual assistance, kindness in every step of the way ... Not unlike a married couple, occasionally on bad terms and profoundly tired of each other, going away for the weekend with no real expectations, and then extending it because things start 'rolling'.' Newcomers, tired of constant bickering, actually seemed extremely congruous, with a positive attitude – unlike Twitter, where conflicts were quick to emerge. One last thing I would like to add is that I find this 'long weekend' is also a lesson to all 'migrants', allowing us to once again realise the importance of relationships with those we hold close to heart. I ended my post with 'May the romance last,' adding three heart emojis.

But the 'romance' didn't last long, not so much because of my posts, but because of other's posts, which admins found controversial or inappropriate and supposedly included too many triggers for people with psychological and physical trauma. Troublesome posts included, for instance, pictures of naked bodies and even cold cuts on a plate, which supposedly triggered people with eating disorders. The network explained how these triggers supposedly affected people:

Be mindful of the people you communicate with, as their experience might differ drastically from yours, including severe psychological and physical trauma. People who lived through a traumatic experience and developed triggers can experience a panic attack when faced with a trigger, especially if it's unexpected. Content warnings don't present a major inconvenience to you as a user, and at the same time ensure a much better well-being for individuals who would like to read your posts, but find it difficult to confront these topics for a variety of reasons (Toot.si 2023).

A list of triggers published alongside this guideline was exhaustive. They included, for instance, food, eye contact, mental health, indecency (including nudity, erotica and pornography), self-harm, violence, rape, ableism (i.e. discrimination towards people with special needs and prejudice against them), ageism (i.e. discrimination and stereotyping based on age), specism (i.e. discrimination based on species membership), racism, supremacy, xenophobia, sexism, fat-shaming, transphobia, homophobia, antisemitism, holocaust, Nazism and neo-Nazism, guns, war, politics, current and controversial news, illicit drugs, alcohol, sarcasm, flashy images, inaccessible content with many emojis or uncommon font and even all caps, which indicate yelling in online texts, on forums and networks (Toot.si 2023).

In short, almost everything that was most desirable and most shared on Twitter, was banned or controversial on the new network. Stricter rules should even prompt new users of Mastodon, or its sub-network Toot.si, to describe posted photos and other images in order to facilitate access for the blind and visually impaired.

5 Intercultural clashes

New rules led to many 'intercultural' clashes and communicational slips, reminiscent of *Close Encounters of the Third Kind*, a 1977 film directed by Steven Spielberg and featuring Earthlings' attempts at communication with visitors from space. An even better metaphor than this film would be colonisation of newly discovered continents, when supposedly 'well-mannered' and 'sustainable' natives were stormed by a horde of gruesome colonisers, unfamiliar with the language and oblivious of local customs – and also unprepared to learn, accept or integrate them (for stereotyping natives, see Graeber and Wengrow 2022). This led to a growing number of disputes, at least at the beginning, as exodus from Twitter continued, it took a few weeks for some sort of temporary armistice and coexistence to be established among the 'natives', some of whom had been on the network for years, and the 'newcomers', rushing into the unknown in the hope of a better world and new possibilities suggesting themselves on the horizon.

I kept following developments on this network up to 4 January 2023, all the while finding that my posts were often missed, unnoticed. What I once wrote on Twitter, garnering a huge response, from likes to shares, was overlooked on Toot.si. The line

between too much and too little attention is obviously a thin one, even when it comes to networks, while expectations about 'local norms' and customs are great.

I observed similar intercultural clashes on the network Threads, which was presented to the public on 5 July 2023 by Meta, a company led by Mark Zuckerberg, co-founder of Facebook and (co)owner of numerous other digital communication solutions and networks, including WhatsApp and Instagram. When introducing this network, Meta took advantage of the very same exodus from Twitter, or X network, and created a platform which almost perfectly matched Twitter in terms of appearance: short texts, allowing users to like, comment and share. The timing of the launch could hardly be more appropriate, which reflected in an exceptionally rapid growth of the number of users; in just five days Threads was used by 100 million users, which surpassed the previous record held by ChatGPT, which is based on artificial intelligence. Access in the European Union was initially disabled due to legal challenges, but on 14 December 2023, EU citizens, including citizens of Slovenia, were given the option to post on this network. While user experience was better than the one on Mastodon's Toot.si, the experience of 'intercultural clash' was similar, even if for a somewhat different reason. The Meta Company namely made it possible for Instagram users to open accounts on Threads and automatically follow all Instagram accounts they followed beforehand. This led to an unusual twist, in which Instagram influencers, who spent years solidifying their positions on the network, mainly by posting images and short videos, capitalised on their pre-existing Instagram profiles to also accumulate the most social capital, i.e. connections, on Threads network, while newcomers joining the network over their disillusionment with Twitter had to start weaving their personal networks and 'threads' (as a series of messages and comments is called, which gave a name to this new solution for making connections and exchanging opinions) from scratch.

I summed up my impressions on the first few days on this 'substitute' network on 3 January 2024 with the following text:

Overnight, the most influential users of Threads have become the ones who accumulated the most social capital on Instagram and have now transformed it into a different environment, which is primarily based on words, not images. This leads to an intercultural clash among those who are used to put on a show of strength, only using words. And the latter, who have predominantly never worked on their visual identity, have been reduced in size and can barely get a word in edgeways.

Reactions from other users of the network for the most part corroborated my findings, while one of the Instagram influencers posted a comment saying that 'it's still too early to draw conclusions', also adding that: 'Now is the right moment for everyone who wish to build a new base of people.'

6 Conclusion

I find the first mention of this ‘new base’ extremely significant for the subject at hand, that is changing identities during constant changes in the era of online social networks, or in the ‘network society’ (Podjed 2010)². This is because networks are constantly changing and transforming, and their structure is susceptible to both social and economic changes, as well as to technological breakthroughs. Musk’s acquisition of Twitter is perhaps the most indicative of these networks’ fragility, which became apparent when he transformed a community of approximately 400 million people, more or less actively participating on the network, almost overnight. A single business move, that is the acquisition of Twitter, rocked the community with – if the network were a country – more citizens than the United States of America. The transformation of the network to X coincided with the exodus, which I presented in this paper and which led to its users’ search for new havens on Mastodon, Threads and other networks, where they attempted to establish their ‘own base’, which would allow them to feel comfortable enough to continue their digital existence.

Such events seem to have made online networks turn into ‘permanent temporary dwellings’, onto which we transfer a good amount of our identities, with the awareness that the comfortable bubble could burst at any moment – and also realisation (or at least subconscious understanding) that we will never be the owners of our dwellings and consequently our identities, that the owner will always be someone else. It depends on the good or bad mood of this individual – or a group of people –, whether our lives are leisurely and exciting or stressful and bitter. But there is really no alternative to sticking to the networks which the digital transformation has made ‘hyperreal’ (Baudrillard 1999), that is, more real than reality itself. For there is not only truth in the saying ‘You are what you post’, which I have mentioned earlier, there is another maxim coming to the foreground: ‘I post, therefore I am.’

Acknowledgements

This paper is part of the research project *Isolated People and Communities in Slovenia and Croatia* (J6–4610) and the research programme *Ethnological, Anthropological and Folklore Studies Research on Everyday Life* (P6–0088). Editors of this collection of papers agreed to the publication of its contents to be presented in a chapter of Dan Podjed’s monograph titled *Krizolacija / Crisolation* (co-published in Slovenian language by Založba ZRC and Cankarjeva založba, 2024).

² This awakens past discussions on identification processes and identities (e.g. Muršič 1997), which are presented as fluid, temporary and negotiated regularly. This is perhaps applies even more to digital environments than to physical space.

References

- Babič, Saša; Podjed, Dan, 2016: Vozila in stereotipi: Primerjava Ljubljane in Beograda. *Glasnik Slovenskega etnološkega društva* 56/1–2, 74–84.
- Bateson, Gregory, 1935: Culture Contact and Schismogenesis. *Man* 35, 178–183.
- Bateson, Gregory, 1999 [1936]: *Naven: A Survey of the Problems Suggested by a Composite Picture of the Culture of a New Guinea Tribe Drawn from Three Points of View*. Stanford: Stanford University Press.
- Baudrillard, Jean, 2010 [1981]: *Simulacra and Simulation*. Ann Arbor: University of Michigan Press.
- Be. B., 2022: Dan Podjed: Musk je zdaj kralj veselja in digitalnega sveta. *Delo* (online edition), 28.10.2022. Online: <https://www.delo.si/novice/slovenija/dan-podjed-musk-je-zdaj-kralj-vesolja-in-digitalnega-sveta> (15.3.2024).
- Dant, Tim, 2004: The Driver-car. *Theory, Culture & Society* 21/4–5, 61–79. DOI: <https://doi.org/10.1177/0263276404046061>.
- Eyal, Nir, 2019: *Indistractable: How to Control Your Attention and Choose Your Life*. Dallas: BenBella Books.
- Graeber, David; Wengrow, David, 2021: *The Dawn of Everything: A New History of Humanity*. London: Penguin Books.
- Haidt, Jonathan, 2012: *The Righteous Mind: Why Good People Are Divided by Politics and Religion*. London: Penguin Books.
- Isaacson, Walter, 2023: *Elon Musk*. London etc.: Simon & Schuster.
- Lunden, Ingrid, 2022: How Mastodon Is Scaling Amid the Twitter Exodus. *TechCrunch*, 23. 12. 2022. Online: <https://techcrunch.com/2022/12/23/how-mastodon-is-scaling-amid-the-twitter-exodus> (15.3.2024).
- Nguyen, C. Thi., 2020: Echo Chambers and Epistemic Bubbles. *Episteme* 17/2, 141–161. DOI: <https://doi.org/10.1017/epi.2018.32>.
- Pariser, Eli, 2011: *The Filter Bubble: What the Internet Is Hiding from You*. London, New York: Viking.
- Podjed, Dan, 2010: Omreženi prostori: Facebook kot znanilec vzpona omrežij in zatona prostorske paradigme. In: Mencej, Mirjam; Podjed, Dan (eds.), *Ustvarjanje prostorov*. Ljubljana, Praga, Bratislava and Krakov: Filozofska fakulteta Univerze v Ljubljani, Filozofická fakulta Univerzita Karlova, Filozofická fakulta Univerzita Komenského, Wydział Filologiczny Uniwersytetu Jagiellońskiego, 133–161.
- Podjed, Dan, 2022a: A Righteous Path to the Digital Future. In: Kampl, Tin (ed.), *Justice in the 21st Century: Discussions about Justice in the Current Situation in Our Society*. Ljubljana: ENoP, 16–23.
- Podjed, Dan, 2022b. Nekrolog dr. Dana Podjeda: Zakaj zapušcam twitter? *Večer* (online edition), 9.11.2022. Online: <https://vecer.com/pogledi/pogled-nekrolog-dr-dana-podjeda-zakaj-zapuscam-twitter-10320654> (15.3.2024).
- Podjed, Dan; Babič, Saša, 2015: Crossroads of Anger: Tensions and Conflicts in Traffic. *Ethnologia Europaea* 45/2, 17–33. DOI: <https://doi.org/10.16995/ee.1163>.
- Podjed, Dan; Pretnar, Ajda, 2018: Samopromocija na Instagramu: primer predsednikovega profila. *Zbornik konference Jezikovne tehnologije in digitalna humanistika, 20. september – 21. september 2018, Ljubljana, Slovenija [Proceedings of the Conference on Language Technologies & Digital Humanities, September 20th – 21st, 2018, Ljubljana, Slovenia]*, 221–226.

- Podjed, Dan, 2019: *Vidni: Zakaj se vse več opazujemo in razkazujemo*. Ljubljana: Založba ZRC, ZRC SAZU.
- Toot.si, 2023: Pogosti sprožilci in vsebinska opozorila. Online: <https://toot.si/about#triggers>
- Varoufakis, Yanis, 2021: Techno-feudalism Is Taking Over. *DiEM25*, 9.7.2023. Online: <https://diem25.org/techno-feudalism-taking-over/> (15.3.2024).
- Varoufakis, Yanis, 2023: *Technofeudalism: What Killed Capitalism*. London: The Bodley Head.
- Yeo, Amanda. 2022. Elon Musk says Twitter Will Give “Amnesty” to Suspended Accounts from Next Week. *Mashable*, 25.11.2022. Online: <https://mashable.com/article/twitter-amnesty-elon-musk-suspended-unban> (15.3.2024).

Avtorji / Authors

Dr. Saša Babič

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
sasa.babic@zrc-sazu.si
ORCID: 0000-0002-3050-757X

Nikola Bakarić

Univerza uporabnih znanosti Velika Gorica, Velika Gorica (Hrvaška)
bakaric@vvg.hr
ORCID: 0000-0001-5453-8883

Dr. Tomaž Erjavec

Institut Jožef Stefan, Oddelek za tehnologije znanja, ZRC SAZU, Inštitut za slovenski jezik
Frana Ramovša Ljubljana (Slovenija)
tomaz.erjavec@ijs.si
ORCID: 0000-0002-1560-4099

Dr. Vanja Huzjan

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
vanja.huzjan@zrc-sazu.si
ORCID: 0000-0002-2674-8880

Dr. Barbara Ivančič Kutin

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
barbara.ivancic@zrc-sazu.si
ORCID: 0000-0001-7581-6328

Dr. Nataša Jakop

ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša, Podiplomska šola ZRC SAZU,
Ljubljana (Slovenija)
natasajakop@zrc-sazu.si
ORCID: 0000-0002-4575-272X

Dr. Monika Kropelj Telban

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
monika.kropej@zrc-sazu.si
ORCID: 0000-0001-6459-3683

Dr. Outi Lauhakangas

Mednarodno združenje za paremiologijo, Tavira (Portugalska)
outi.lauhakangas@sci.fi
ORCID: 0000-0003-1963-2756

Dr. Matej Meterc

ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša, Ljubljana (Slovenija)
matej.meterc@zrc-sazu.si
ORCID: 0000-0002-0198-3745

Rok Mrvič, mag.

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
rok.mrvic@zrc-sazu.si
ORCID: 0000-0002-7401-5999

Dr. Davor Nikolić

Univerza v Zagrebu, Fakulteta za humanistiko in družbene vede, Oddelek za hrvaški jezik,
Zagreb (Hrvaška)
dnikoli@ffzg.hr
ORCID: 0000-0003-4962-7798

Dr. Dan Podjed

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
dan.podjed@zrc-sazu.si
ORCID: 0000-0003-1914-6053

Dr. Ajda Pretnar Žagar

Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Inštitut za novejšo
zgodovino, Ljubljana (Slovenija)
ajda.pretnar@inz.si
ORCID: 0000-0002-5927-4538

Dr. Katarina Šrimpf Vendramin

ZRC SAZU, Inštitut za slovensko narodopisje, Ljubljana (Slovenija)
katarina.srimpf@zrc-sazu.si
ORCID: 0000-0001-8890-2110





TRADICIONALNE PĀREMILOŠKE
ENOTE V DIALOGU S SODOBNO RABO

ISBN 978-961-05-0885-4 23 €



9 789610 508854