



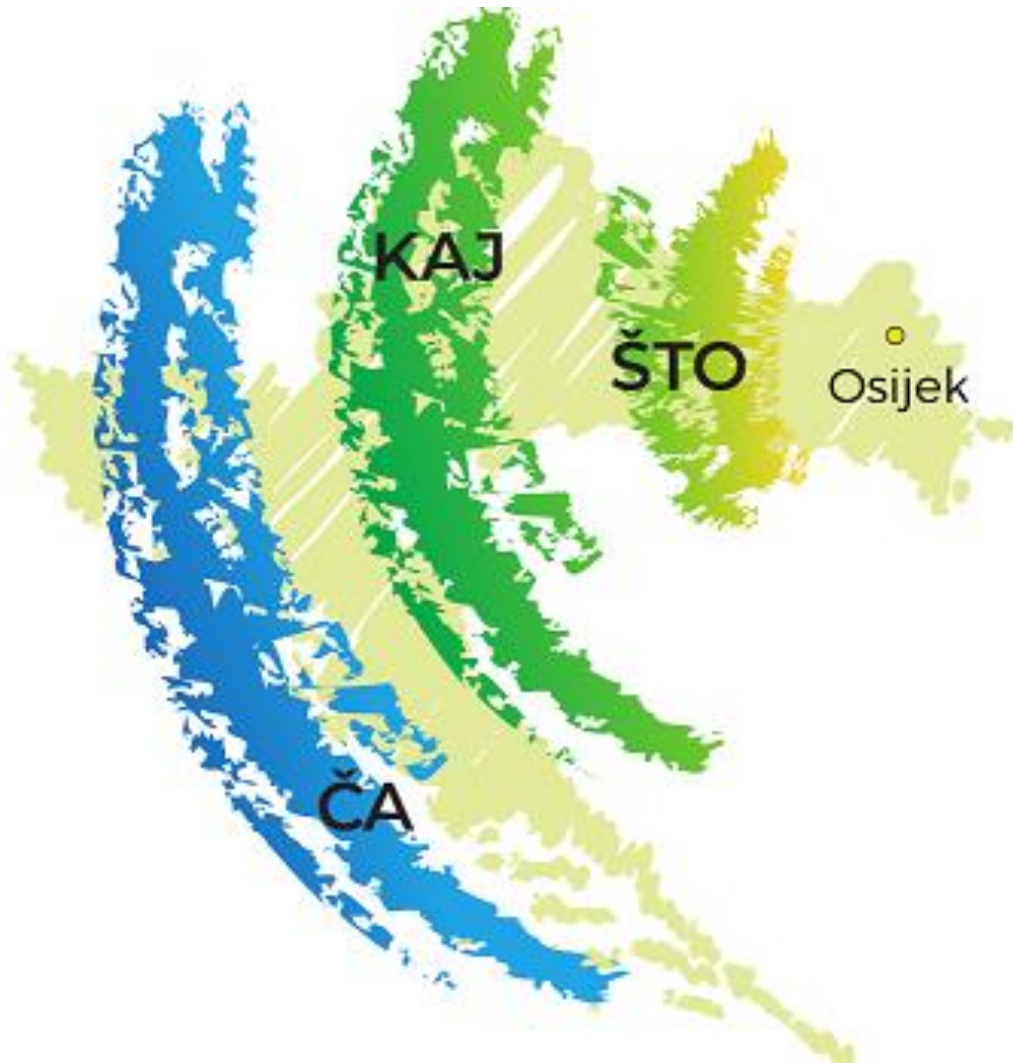
Sveučilište Josipa Jurja  
Strossmayera u Osijeku



Sveučilište J. J. Strossmayera u  
Osijeku - Odjel za matematiku



Filozofski fakultet  
Sveučilišta u Osijeku



Knjiga sažetaka

Book of abstracts

Osijek, 25. i 26. rujna 2020.

Osijek, September 25-26, 2020.



*Izdavač / Publisher*

Sveučilište J. J. Strossmayera u Osijeku – Odjel za matematiku  
Trg Ljudevita Gaja 6, Osijek

*Odgovorna osoba izdavača:*

Prof. dr. sc. Kristian Sabo

*Skup / Conference*

Računalno jezikoslovlje i Zlatna formula hrvatskoga jezika  
Computational Linguistics and Golden Formula of Croatian Language

*Recenzenti*

Dr. sc. Josip Lasić, v. lektor, Sveučilište u Splitu, Filozofski fakultet  
Izv. prof. dr. sc. Domagoj Matijević, Sveučilište Josipa Jurja Strossmayera u Osijeku, Odjel  
za matematiku

*Urednici*

Ivana Kuzmanović Ivičić, Domagoj Ševerdija

*Lektorica*

Ana Maria Ban, prof

*Grafičko oblikovanje i strojna obradba:*

Mario Essert, Mario Lončarić

*Mjesec i godina objavljivanja publikacije*

studen, 2021.

**ISBN 978-953-8154-18-8**

Sveučilište Josipa Jurja Strossmayera u Osijeku



Suglasnost za izdavanje ove knjige sažetaka donio je Senat Sveučilišta Josipa Jurja Strossmayera u Osijeku na 2. sjednici u akademskoj godini 2021./2022. održanoj 24. studenoga 2021. godine pod brojem 37/21.

*Tisak*

e-knjiga sažetaka

*Znanstveno-stručni skup: (\*)*

**Računalno jezikoslovlje i Zlatna formula hrvatskoga jezika**



*Scientific and Professional Conference:*

**Computational Linguistics and Golden Formula of Croatian Language**

<http://m-brojevi.com/en/skup-2020>

(\*) Skup je organiziran u sklopu istraživačkoga projekta koji je 2019. financirala Zaklada Adris (Područje Baština): Mrežni okvir ča-kaj-što sastavnice hrvatskoga identiteta (Sveučilište Josipa Jurja Strossmayera u Osijeku – Odjel za matematiku)

# Organizacija skupa

## Conference organisation

### Organizacijski odbor / Organisation Board

- [doc. dr. sc. Ivana Kuzmanović Ivičić](#) (MathOS), predsjednica Odbora
- [prof. dr. sc. Milica Lukić](#) (FFOS)
- [dr. sc. Mario Essert, prof. u miru](#)
- [doc. dr. sc. Domagoj Ševerdija](#) (MathOS)
- [doc. dr. sc. Silvija Ćurak](#) (FFOS)
- [izv. prof. dr. sc. Mihaela Ribičić Penava](#) (MathOS)
- [dr. sc. Vera Blažević Krezić](#) (FFOS)

### Programski odbor / Editorial Board

- [doc. dr. sc. Silvija Ćurak](#) (FFOS), predsjednica Odbora
- [dr. sc. Ivana Kurtović Budja](#) (IHJJ)
- [prof. dr. sc. Ljiljana Kolenić](#) (FFOS)
- [prof. dr. sc. Milica Lukić](#) (FFOS)
- [doc. dr. sc. Domagoj Ševerdija](#) (MathOS)
- [doc. dr. sc. Marija Malnar Jurišić](#) (FFZG)
- [prof. dr. sc. Andjela Frančić](#) (FFZG)
- [dr. sc. Anita Celinić](#) (IHJJ)
- [izv. prof. dr. sc. Boris Bosančić](#) (FFOS)
- [dr. sc. Vera Blažević Krezić](#) (FFOS)
- [dr. sc. Vlatka Štimac Ljubas](#) (LZMK)

### Naši podupiratelji i donatori:



Adris grupa, Rovinj



Superfluo, Zagreb



Končar, Zagreb



ASC, Zagreb

## Cilj skupa

Cilj je Skupa okupiti dijalektologe, istraživače povijesti hrvatskoga jezika, metodičare i računalne stručnjake koji se bave istraživanjem međusobne povezanosti hrvatskih narječja na svim jezičnim razinama te kontrastivnim pristupom njihovu poučavanju u školama. Namjera je predstaviti **Zlatnu formulu hrvatskoga jezika (ča-kaj-što)** preko sintaktičko-semantičkoga mrežnog okvira koji već postoji za standardni hrvatski jezik.

Iz emisije „Dobro jutro, Hrvatska“ [2. ožujka 2020.](#)

Skupom su obuhvaćena sljedeća tematska područja:

- fonologija, morfologija i leksik triju hrvatskih narječja
- povijest hrvatskoga jezika
- dijalektološke teme u nastavi hrvatskoga jezika
- korpusna leksikografija i leksikologija
- poteškoće digitalizacije i mrežnoga održavanja korpusa i rječnika
- računalno jezikoslovlje u jezičnoj integraciji hrvatskih narječja

*Pozvani predavači:*

- [prof. dr. sc. Sanja Zubčić](#) (FFRI) – čakavština
- [prof. dr. sc. Đuro Blažeka](#) (UFZG) – kajkavština
- [doc. dr. sc. Željko Jozić](#) (IHJJ) – štokavština

## Aim of the conference

The aim of this conference is to bring together dialectologists, Croatian language history researchers, methodologists and computer experts who are researching the interconnectedness of Croatian dialects at all language levels and the contrasting approaches to teaching them in schools. The wish is to present the Golden Formula of the Croatian language dialects (*ča-kaj-što*) via a syntactic-semantic web framework that already exists for standard Croatian.

From the TV show *Good Morning Croatia* from [2 March 2020](#).

The conference will examine:

- the phonology, morphology and vocabulary of the three Croatian dialects
- the history of the Croatian language
- dialectological topics in Croatian language teaching
- corpus lexicography and lexicology
- difficulties in digitizing and maintaining the corpus and vocabulary online
- computer linguistics in linguistic integration

## Zlatna formula

Zlatna formula hrvatskoga jezika ča-kaj-što hrvatsko je nematerijalno kulturno dobro upisano na Listu nematerijalnih kulturnih dobara Republike Hrvatske u studenome 2019. godine. Autor toga naziva, predlagatelj te autor koncepcije očuvanja znanja i praksi koje pruža tijesna povezanost govora hrvatskoga jezika jest dr. Drago Štambuk.

Sam naziv izražava međupovezanost govora triju narječja hrvatskoga jezika, čakavskoga, kajkavskoga i štokavskoga, koja je veća nego što se dosad tvrdilo. Na Skupu će se pokazati da Formula pokriva svu dijalektnu raznolikost, ali i povijest hrvatskoga jezika sa svima trima stilizacijama počevši od fonologije, preko leksika do sintakse (npr. tronaglasni sustav; glagoli govorenja; upitne rečenice). Najbolji način pokazivanja te ujedinjujuće snage Zlatne formule jest stvaranje relevantnih računalnih korpusa dijalektalne i povijesnojezične građe te brza i precizna obrada podataka, a konferencija Računalno jezikoslovlje i Zlatna formula hrvatskoga jezika ča-kaj-što nastojat će odgovoriti na pitanje kako to postići.

## Golden formula

The Golden Formula of the Croatian language, ča-kaj-što, has been recognized as an intangible Croatian cultural asset and first appeared on the List of Intangible Cultural Heritage of Croatia in November 2019. Its creator, Dr Drago Štambuk, is known as a proponent and author of the preservation of knowledge and practices found within the close connection between Croatian dialects, including their local speeches.

The Formula's name itself expresses the interconnectedness of the three Croatian dialects – Chakavian, Kajkavian and Stokavian – which is greater than previously claimed. The purpose of this conference is to demonstrate how the Formula covers all dialectal diversity, as well as the history of the Croatian language and all three of its stylizations, starting with phonology and continuing with vocabulary and syntax (e.g. the three-accentual system; speaking verbs/verba dicendi; interrogative sentences). The best way to demonstrate the unifying strength of the Golden Formula is to create relevant electronic databases of Croatian dialectal and language history corpus materials, as well as encourage fast and efficient data processing. This conference seeks to answer the question of how to achieve this goal

## Rektorovo pismo / govor

Poštovane sudionice i sudionici Skupa!

Kao što u našoj viziji Sveučilišta Josipa Jurja Strossmayera u Osijeku piše, ono „svoj razvoj temelji na bogatom kulturnom i povijesnom nasljeđu u užem regionalnom kontekstu i širem (srednjo)europskom kontekstu. Isto tako nastoji očuvati tradicijske vrijednosti i posebnosti visokog obrazovanja koje baštini više od tri stoljeća.“ Vaš Skup nalazi se čvrsto u temelju te vizije i višestoljetnog rada hrvatskih jezikoslovaca na očuvanju hrvatskog identiteta i jezika.

Skup Računalno jezikoslovlje i Zlatna formula hrvatskoga jezika povezuje dvije naše sastavnice: Odjel za matematiku i Filozofski fakultet Sveučilišta u Osijeku, što predstavlja najviše traženo interdisciplinarno područje zajedničkog djelovanja i rada. Ponosni smo što se to događa u našoj sredini. Već smo 2019. godine bili domaćini skupa Hrvatski jezik u računalnom jezikoslovlju gdje su se okupili predavači iz različitih područja i raspravljali o ulozi hrvatskoga jezika u hrvatskom identitetu u računalnoj integraciji.

Ovogodišnji skup Računalno jezikoslovlje i Zlatna formula hrvatskoga jezika nastavak je tog rada s novim trolistom ča-kaj-što sastavnica hrvatskoga jezika, našega dijalektnog blaga koje Hrvatska ima. Na vama je da u znanstvenom, stručnom, a potom pedagoškom radu to blago sačuvate za generacije koje dolaze. Želim vam djelotvoran i uspješan skup!

Prof. dr. sc. Vlado Guberac,

rektor Sveučilišta Josipa J. Strossmayera u Osijeku

## Pozdravna riječ veleposlanika dr. Drage Štambuka

Poštovani sudionici prvoga znanstveno-stručnog skupa o Zlatnoj formuli hrvatskoga jezika ča-kaj-što,

velika mi je čast i radost obratiti vam se uoči otvaranja osječke Konferencije, prve koja se posvećuje Zlatnoj formuli hrvatskoga jezika, nakon njezina proglašenja, 29. studenoga 2019., kulturnim dobrom Republike Hrvatske i stavljanja na Nacionalnu listu šticipene nematerijalne baštine.

Cijeloga života u njezinom sam posjedu, od trenutka kada sam kao dječak iz čakavskoga mjesta Selca na otoku Braču počeo pohađati osnovnu školu, u kojoj su štokavci iz susjednoga Sumartina odmah došli na svoje, a mi čakavci, iako većina, trebali smo se prilagoditi službenom jezičnom standardu. Očutio sam tada snažan rez u hrvatskome jeziku i moja tadašnja, intuitivna, ne do kraja osviještena odluka, bila je – raditi na svladavanju razlika i integraciji hrvatskih triju idioma ili stilizacija; ne želim ih zvati narječjima, a još manje dijalektima, jer ih se time, držim, donekle unižava i bilježi manjom vrijednošću.

Zbog čega je Zlatna formula hrvatskoga jezika ča-kaj-što kulturno dobro od hrvatskoga nacionalnog značaja?

Prvo – zato što preslaguje odnos prema hrvatskome jeziku i upućuje na put kojim je, povijesno, bilo u interesu našega naroda razvijati vlastiti jezik i poticati mu rast u stablo ujedinjujuće trojedne ča-kaj-što hrvatske koiné.

Drugo – jer je Formula stožerni koncept oko kojega sabiremo duboko korijenje i sve tri važne jezične grane u moćnu krošnju hrvatskoga jezika koja nam daje hlad odmora i spokoja – spaja i ucjelovljuje, stišava i liječi povijesne rane i prokazuje krive odluke – ne samo u jeziku već i u politici, a u jezičnoj ponajviše.

Treće – Formula nas književno okuplja, stavljaajući djela iz svih jezičnih dionica na jedan svehrvatski stol; uz glavno štokavsko jelo: stilizaciju koja se u međuvremenu sjajno potvrdila – ne manje važne su čakavsko-kajkavske stilizacije i jestvine, a koje sve tri zajedno, u duhu punine i potpunosti, odvajkada hrane naša usta i našu dušu.

Živodajna molekula vode idealni je poredbeni pars pro toto u predstavljanju ča-kaj-što formule preko osmotskih sila između dvaju manjih atoma vodika (ča i kaj) i većeg atoma kisika (što). Tek sva tri atoma zajedno čine molekulu vode s valentnim kemijskim vezama, a jezično su one predstavljene crticama unutar oznaka „ča-kaj-što“; jer Formula se piše s crticama između upitnih zamjenica: ča, kaj i što; crtice označavaju odvojenost, ali i spojenost, fluidno jezično zajedništvo kakvo nam je bilo potrebno nekad, ali i danas, u teško izborenoj slobodi još više.

Ča-kaj-što zaista jest poput životvorne vode – H<sub>2</sub>O.

Možda se najidealnije povijesno opravdanje za ideju trojednosti hrvatskoga jezika nalazi u Ozaljskome književnom krugu u kojemu se u 17. stoljeću njegovala i stvarala alkemijska slatina sljublivanja i prožimanja trojstvenoga nam jezika; no bečkim zatiranjem zrinško-frankopanske elite koja je prirodno i samosvojno promicala rečeni koncept hrvatske koiné iliti mješavine, propao je organski pristup kojim je hrvatski



jezik mogao i trebao krenuti posve drugim razvojnim putem i smjerom te tako ujediniti jezične dionice hrvatskoga naroda i razlomljeno mu ozemlje.

Četvrto – Formula, što je osobito važno, upućuje na bliskost i povezanost unutar triju idioma, doživljavajući ih jednako svojim i učvršćujući nam svijest o jedinstvenosti i uzajamnoj prožetosti hrvatskoga jezika bez obzira na njegovu blistavu leksičku pričuvu i golemu razvedenost. I hrvatska je jadranska obala iznimno razvedena, no to smo oduvijek smatrali njezinom prednošću. I ovaj će skup računalnim postupcima dokazivati međusobno prožimanje unutar trojednoga jezika kao neupitnu datost pa u tom nastojanju želim časnim sudionicima mnogo uspjeha.

Budimo ponosni i dičimo se blagom koje u jeziku „Višnji nam Bog je do“; sretni – da u krvavo stečenoj državnoj neovisnosti smijemo uživati u svojem bogatstvu, prepuštati mu se i rabiti ga na načine koji odgovaraju posebnosti hrvatskoga narodnog bića i osobitoj mu sudbini.

Hvala vam, dragi jezikoslovci i književnici, na ljubavi koju iskazujete prema svojem složenome jeziku jer on nije samo znameni nositelj tradicija već i temelj sadašnjosti; a kroz cjelokupnost Zlatne formule snažno je usmjeren i na hrvatsku budućnost.

Budimo, stoga, blagodarni prema cjelovitome ča-kaj-što jeziku jer će nam uzvratiti kreativnošću koja podiže primjernu i primjerenu kuću bitka i Riječi, jača nacionalni identitet i brani opstanak doma nam hrvatskoga.

Neka oznak ča-kaj-što zvoni u vašim srcima kano živodajna H<sub>2</sub>O (HA-DVA-O), kano presveto trojstvo: Otac-Sin-Duh, a poput zlatne strijele naperen bude u hrvatsku snivanu i dozivanu budućnost; u njoj da se potvrđuje i svijetli ukazujući na izlazak iz starog nerazumijevanja i dugovječne zbuđenosti.

Jer kako zapisa pop Martinac nakon Krbavske bitke: „Turci nalegoše na jzik hrva'ski“; gdje jezik označava narod svjedočeci istovjetnost „pûka i hûka“; radeći na cjelovitom jeziku, brinući se za nj – svi mi, njegovi zaljubljenici, istovremeno podižemo dom opstojnosti hrvatskome narodu i širimo slavu njegova imena.

Ča-kaj-što!

Svekoliko vam dobro i hvala na pozornosti!

Drago Štambuk

Zagreb, 4. rujna 2020.

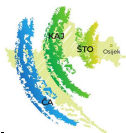
(Obraćanje sudionicima konferencije Računalno jezikoslovlje i Zlatna formula hrvatskoga jezika, Osijek, 25. – 26. rujna 2020.)

## Sadržaj

AIM OF THE CONFERENCE .....	V
GOLDEN FORMULA .....	VI
ČAKAVŠTINA: OD RURALNOGA FENOMENA DO UMJETNE INTELIGENCIJE	1
CHAKAVIAN: FROM A RURAL PHENOMENON TO ARTIFICIAL INTELLIGENCE .....	2
PRIJEDLOG NOVIH PRAVACA ISTRAŽIVANJA LEKSIKA MJESNIH GOVORA KAJKAVSKOG NARJEČJA	3
PROPOSAL FOR NEW DIRECTIONS IN THE RESEARCH OF THE VOCABULARY OF THE LOCAL SPEECHES OF THE KAJKAVIAN DIALECT .....	4
NEZAMIJENJENI JAT U SLAVONSKOME DIJALEKTU	6
UNREPLACED OR PRESERVED JAT' IN THE SLAVONIAN DIALECT .....	7
POVIJESNOJEZIČNE TEME U NASTAVI HRVATSKOGA JEZIKA	8
HISTORICAL LINGUISTIC TOPICS IN CROATIAN LANGUAGE TEACHING .....	9
ČRNKOV ZAPIS IZ DRŽAVNOG ARHIVA U BEČU O OBRANI SIGETA OD TURAKA I POGIBIJI NIKOLE ZRINSKOG 1566. G.	10
ČRNKO'S RECORD ABOUT THE DEFENSE OF SIGET FROM THE TURKS AND THE DEATH OF NIKOLA ZRINSKI IN 1566, KEPT IN THE STATE ARCHIVE IN VIENNA .....	11
ODGOJITELJI I UČITELJI PRED IZAZOVIMA HRVATSKE DIJALEKTNE RAZNOLIKOSTI	12
TEACHERS AND PRE-SCHOOL TEACHERS FACING THE CHALLENGE OF THE CROATIAN DIALECTAL DIVERSITY .....	13
UPITNICI ZA HRVATSKI JEZIČNI ATLAS I ZLATNA FORMULA HRVATSKOGA JEZIKA ČA-KAJ-ŠTO NA PRIMJERU GOVORA MAKARSKOGA PRIMORJA	14
QUESTIONNAIRES FOR THE CROATIAN LANGUAGE ATLAS AND THE GOLDEN FORMULA OF THE CROATIAN LANGUAGE ČA-KAJ-ŠTO – DIALECT OF MAKARSKA COAST .....	15
MOLIŠKOHRVATSKI IDIOM: KVANTITATIVNA ANALIZA DIJAFAZNIH VARIJACIJA U DVOJEZIČNOM KONTEKSTU	16
U RIČ SAKRITA ŠOKADIJA, MED ĐAKOVA I VINKOVCI POLOŽITA	18
BETWEEN ĐAKOVO AND VINKOVCI PLACED ŠOKADIJA – HIDDEN IN THE WORD .....	19
SMUTICA ZA ZDRAV(L)JE HRVATSKOG JEZIKA	20
SMOOTHIE FOR THE HEALTH OF THE CROATIAN LANGUAGE .....	21
ZLATNA FORMULA HRVATSKOGA JEZIKA U HRVATSKOME ZVUČNOM ATLASU	24
THE GOLDEN FORMULA OF THE CROATIAN LANGUAGE IN CROATIAN SOUND ATLAS .....	25
MJESNI GOVORI KAO BAŠTINA U JEZIČNOJ POLITICI I ODGOJNO-OBRAZOVNOM PROCESU	27
LOCAL SPEECHES AS A HERITAGE IN LANGUAGE POLICY AND EDUCATIONAL PROCESS	28
MJESTO VIŠERJEČNIH FARMACEUTSKIH NAZIVA U OPĆEJEZIČNIM RJEČNICIMA HRVATSKOGA JEZIKA .....	29
THE LEXICOGRAPHICAL LEVEL OF MULTI-WORD PHARMACEUTICAL TERMS IN GENERAL-LANGUAGE CROATIAN DICTIONARIES .....	30
KATOLIČKI KALENDAR U ZRCALU GACKE ČAKAVŠTINE	31

DER KATHOLISCHE KALENDER IM SPIEGEL DES ČAKAVISCHEN DIALEKTS DER REGION GACKA .....	32
KAJKAVSKI DONJOSUTLANSKI IKAVSKI DIJALEKT U PROJEKTU MREŽNI OKVIR "ČA-KAJ-ŠTO" SASTAVNICE HRVATSKOGA IDENTITETA .....	33
KAJKAVIAN LOWER SUTLA IKAVIAN DIALECT INCLUDED IN THE PROJECT NETWORK FRAMEWORK "ČA-KAJ-ŠTO" COMPONENTS OF CROATIAN IDENTITY .....	34
„BEĆARAC – DESETERAC S PORUKOM“ .....	35
„BEĆARAC – TEN-PART WITH MESSAGE“ .....	36
TRI HRVATSKA E-LEKSIKONA U MREŽNOME OKVIRU .....	37
THE THREE CROATIAN E-LEXICONS IN THE NETWORK FRAMEWORK .....	38
UREĐIVANJE MREŽNIH E-RJEČNIKA U SINTAKTIČKO-SEMANTIČKOM OKVIRU .....	39
EDITING ONLINE E-DICTIONARIES IN THE SYNTACTIC-SEMANTIC FRAMEWORK .....	40
OBRADA LINGVISTIČKIH PODATAKA KORISTEĆI API SSF-A .....	41
PROCESSING OF LINGUISTIC DATA USING SSF API .....	42
HRVATSKI MREŽNI SUBČESTIČNI (STRUKTURNI) RJEČNIK .....	43
CROATIAN NETWORK SUBPARTICLE (STRUCTURAL) DICTIONARY .....	44
HRVATSKI MREŽNI E-RJEČNIK VIŠERJEČNICA .....	45
CROATIAN NETWORK E-DICTIONARY OF MULTIWORD EXPRESSIONS .....	46





Sanja Zubčić, Filozofski fakultet u Rijeci, [sanja.zubcic@ffri.uniri.hr](mailto:sanja.zubcic@ffri.uniri.hr)

## ČAKAVŠTINA: OD RURALNOGA FENOMENA DO UMJETNE INTELIGENCIJE

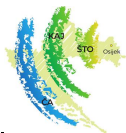
Svaki je jezik podložan promjenama, a one mogu biti motivirane unutarjezičnim i/ili izvanjezičnim čimbenicima. Središnji pojam ovoga predavanja *promjene* su koje je čakavština kao autonomna, ali i kao subordinirana jezična zajednica prolazila ili prolazi.

Prikazat će se temeljne promjene u dijalekatskoj strukturi u povijesnom razvoju čakavštine te će se usporediti njezina predmigracijska i postmigracijska karta, ne zanemarujući i recentne promjene u stratifikaciji.

Središnji će dio predavanja biti posvećen izvanjezičnim čimbenicima koji su uvjetovali promjene u strukturi jezične zajednice i njezinoj vitalnosti. Glavnina tih čimbenika uvjetovana je društvenim promjenama i najvećma je vodila prema postupnom smanjenju izvornih govornika pojedinih mjesnih govora, pa čak i nestanku pojedinih govora. Taj će se proces promatrati kroz prizmu *teorije jezične vitalnosti* (Giles, Bourhis i Taylor 1977), koja svoj zamah doživljava tek posljednjih petnaestak godina i bavi se istraživanjem čimbenika koji utječu na snagu i sposobnost jezične zajednice da očuva svoj jezik. Prema Pauwelsu (2016) tri su skupine čimbenika koji utječu na očuvanost nekoga jezika: biografski (primjerice dob govornika, spol, bračni status, stupanj obrazovanja, govornikov stav o svojem jeziku i dr.), značajke manjinske zajednice (brojnost zajednice, gustoća njihove te jezična i kulturološka sličnost s većinskom zajednicom), značajke većinske zajednice (većinske zajednice prema manjinskom jeziku i kulturi te postojanje zakona i propisa koji podržavaju jezičnu raznolikost i višejezičnost). U izlaganju ćemo se osvrnuti na najvažnije od njih.

S obzirom na prirodu ovoga skupa, posebna će se pozornost posvetiti uporabi čakavštine u novim medijima i društvenim mrežama te ulozi koju ti mediji mogu imati u promicanju i očuvanju čakavštine.

**Ključne riječi:** čakavština, jezične promjene, vitalnost jezične zajednice



Sanja Zubčić, Faculty of Humanities and Social Sciences in Rijeka, [sanja.zubcic@ffri.uniri.hr](mailto:sanja.zubcic@ffri.uniri.hr)

## CHAKAVIAN: FROM A RURAL PHENOMENON TO ARTIFICIAL INTELLIGENCE

Every language is susceptible to change, which can be intralinguistically and/or extralinguistically motivated. The key topic of this lecture is thus the changes that Chakavian as an autonomous yet subordinate Croatian linguistic community has been or is still undergoing. This lecture highlights basic changes in the dialectological structure and historical development of Chakavian and compares its pre-migration and post-migration maps, while taking into account recent changes in stratification.

The central part of the lecture is dedicated to examining the extralinguistic factors that have caused changes in the structure of the Chakavian linguistic community and its vitality. Most of these factors stem from changes in society, usually leading to a gradual decline in the number of native speakers or even disappearance of certain idioms. This process is observed through ethnolinguistic vitality theory (Giles, Bourhis and Taylor 1977), which has gained momentum only in the past fifteen years, to investigate factors that influence the strength and ability of a linguistic community to preserve its language. According to Pauwels (2016), there are three groups of factors affecting the preservation of a language: biographical (e.g. speakers' age, gender, marital status, education level, attitude towards their language, etc.), characteristics of the minority community (e.g. size, density, linguistic and cultural similarity to the majority community) and characteristics of the majority community (e.g. attitude towards the minority language and culture, existence of laws and regulations supporting linguistic diversity and multilingualism).

In accordance with the conference outline, the lecture dedicates special attention to the use of Chakavian in new media and social networking sites, as well as the role that they may have in its preservation.

**Keywords:** Chakavian, linguistic changes, vitality of a linguistic community

Đuro Blažeka, Učiteljski fakultet u Zagrebu, [djuro.blazeka@ufzg.hr](mailto:djuro.blazeka@ufzg.hr)

## PRIJEDLOG NOVIH PRAVACA ISTRAŽIVANJA LEKSIKA MJESNIH GOVORA KAJKAVSKOG NARJEČJA

Autor govori o nekim spoznajama u svojem dugogodišnjem istraživanju leksika kajkavskog narječja i apostrofira četiri važna pravca u kojima se ta istraživanja mogu razvijati i koja bi bila važna za istraživanje leksika hrvatskog jezika u cjelini i ukupnim spoznajama o njemu. Prvi je izrada razlikovnih rječnika između bliskih mjesnih govora jer bi izrada posebnoga velikog rječnika za svaki govor bila neekonomična ako već postoji znanstveni rječnik za neki manje ili više bliski govor. Predlažu se kategorije razlikovnosti na temelju kojih bi se riječi uvrštavale u takve rječnike. Drugi je novi pristup rečeničnim potvrdoma u dijalektalnim rječnicima, a koji bi razlikovao „obične rečenične potvrde“ (egzemplifikacije) i „ekspresivne rečenične potvrde“ koje bi bile iznimna leksikološka vrijednost same po sebi. Treći je istraživanje pseudoanalogonimije, i to u tri pravca: a) pseudoanalogonimija između nekoga hrvatskog mjesnog govora (skupine govora / poddijalekta / dijalekta) i standardnog jezika, b) pseudoanalogonimija između dvaju mjesnih govora (skupine govora / poddijalekta) nekoga hrvatskog dijalekta, c) pseudoanalogonimija između dvaju mjesnih govora (skupine govora / poddijalekta) različitih hrvatskih dijalekata. Četvrti je predloženi pravac istraživanje suvremenih semantičkih adaptacija arhaičnoga kajkavskog leksika na primjeru hungarizama i germanizama. Predlaže se sedam tipova takvih adaptacija pomoću kojih arhaični leksik opstaje u uvjetima nestanka realija koje je označavao i utjecaja leksika iz hrvatskoga standardnog jezika koji ga zamjenjuje.

**Ključne riječi:** razlikovni rječnici, rečenične potvrde, pseudoanalogonimija, suvremene semantičke adaptacije

### Literatura:

1. Đuro Blažeka: Hungarizmi u govoru Goričana, *Suvremena lingvistika*, godina 32, svezak 1, Zagreb 2006, str. 1-27.
2. Adaptacije germanizama u međimurskom dijalektu, *Studia Slavica Hungaricae, Academiae scientiarum Hungaricae*; editor-in-chief István Nyomárkay; 54/1 (2009) 47–76. Budapest 2009., str. 39-76.
3. Temeljna načela i postupci u pretvaranju sirove dijalektološke građe u znanstveni rječnik, *Slavia Centralis*, številka 2 / 2010 / letnik III., Oddelek za slovanske jezike in književnosti - Filozofska fakulteta Univerze v Mariboru, str. 54-73.
4. Đuro Blažeka: "Lažni prijatelji" između hrvatskoga standardnoga jezika i međimurskoga dijalekta, *Filologija* 57, Hrvatska akademija znanosti i umjetnosti, Razred za filološke znanosti, Zagreb 2012., str. 1- 33.
5. Đuro Blažeka: "Razlikovni rječnici" između 2 bliska mjesna govora, *Studia Slavica Savarensia* 1-2., Szombathely 2016., str. 57-71.

Đuro Blažeka, Faculty of Teacher Education University of Zagreb, [djuro.blazeka@ufzg.hr](mailto:djuro.blazeka@ufzg.hr)

## PROPOSAL FOR NEW DIRECTIONS IN THE RESEARCH OF THE VOCABULARY OF THE LOCAL SPEECHES OF THE KAJKAVIAN DIALECT

The author discusses some findings from his long-term research on the lexicon of the Kajkavian dialect, and emphasizes four important directions in which new studies could be developed to further research the lexicon of the whole Croatian language and overall knowledge about it. The first direction is to compile dictionaries of distinctive features between Croatian's closely related local dialects, as a special dictionary for each speech would be uneconomical if a scientific dictionary already existed for closely related speech. The presentations suggests categories of distinctiveness for this purpose, on the basis of which words would be included in such dictionaries. The second direction is a new approach to sentence confirmations in dialect dictionaries, which would distinguish between "ordinary sentence confirmations" (exemplifications) and "expressive sentence confirmations", which would themselves be of exceptional lexical value. The third is the research of pseudoanalogonymy, which has three prongs: a) the pseudoanalogonymy between a local Croatian dialect (groups of speech/subdialects/dialects) and the standard language, b) that between two local Croatian speeches (groups of speech/subdialects) and c) that between two local dialects (groups of speech/subdialects) of different greater Croatian dialects.

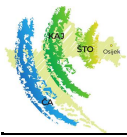
The fourth proposed direction is the research of contemporary semantic adaptations of the archaic Kajkavian lexicon with examples of Hungarianisms and Germanisms. Seven types of these adaptations are proposed by which the archaic lexicon survives in the present conditions of the currently disappearing realia (concrete objects) that the lexicon once signified. These adaptations also exist in the standard Croatian's influence, which the dialect's lexemes used to signify, and the influence of the lexicon on standard Croatian itself, which is replacing it.

**Keywords:** dictionaries of distinctive features, sentence confirmations, pseudoanalogonymy, modern semantic adaptations

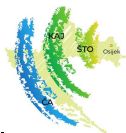
### References:

1. Đuro Blažeka: Hungarizmi u govoru Goričana, *Suvremena lingvistika*, godina 32, svezak 1, Zagreb 2006, str. 1-27.
2. Adaptacije germanizama u međimurskom dijalektu, *Studia Slavica Hungaricae, Academiae scientiarum Hungaricae*; editor-in-chief István Nyomárkay; 54/1 (2009) 47–76. Budapest 2009., str. 39-76.
3. Temeljna načela i postupci u pretvaranju sirove dijalektološke građe u znanstveni rječnik, *Slavia Centralis*, številka 2 / 2010 / letnik III., Oddelek za slovanske jezike in književnosti - Filozofska fakulteta Univerze v Mariboru, str. 54-73.





4. Đuro Blažeka: "Lažni prijatelji" između hrvatskoga standardnoga jezika i međimurskoga dijalekta, *Filologija* 57, Hrvatska akademija znanosti i umjetnosti, Razred za filološke znanosti, Zagreb 2012., str. 1- 33.
5. Đuro Blažeka: "Razlikovni rječnici" između 2 bliska mjesna govora, *Studia Slavica Savarensia* 1-2., Szombathely 2016., str. 57-71.



Željko Jozić, Institut za hrvatski jezik i jezikoslovlje, Zagreb, [zjozic@ihj.hr](mailto:zjozic@ihj.hr)

## NEZAMIJENJENI JAT U SLAVONSKOME DIJALEKTU

Nezamijenjeni ili očuvani jat u dijalektološku je literaturu kao pojam uveo Pavle Ivić u svojoj disertaciji iz 1957. godine pišući o govoru galipoljskih Srba u Pehčevu. Ivić se pozvao na istraživanje rumunjskoga dijalektologa Emila Petrovicija „O govoru Krašovana: studija iz južnoslavenske dijalektologije” (Bukurešt, 1937.), u kojem je zabilježena vrlo slična realizacija praslavenskoga glasa jat u govoru „karaševskih katolika”. No, prije Ivića i Petrovicija u hrvatskoj je dijalektologiji zabilježen nezamijenjeni jat (doduše, nije tako nazvan) još 1900. godine te nešto kasnije, 1913. godine. Naime, Šime Varnica prvi je prije 120 godina zabilježio „neobičnu” realizaciju jata kao zatvorenoga *ę* ili kao *ei* u govoru sela Gradišta u slavonskome dijalektu, a to je potvrdio i Stjepan Ivšić u „Današnjem posavskom govoru” 1913. godine. Tema izlaganja potvrđivanje je teze da je u govoru sela Gradišta i danas prisutan nezamijenjeni praslavenski jat, potvrđen i snimljen u nedavnome istraživanju.

**Ključne riječi:** nezamijenjeni jat, slavonski dijalekt, Gradište kod Županje

### Literatura

1. Ivšić, Stjepan. 1913. *Današnji posavski govor*. Rad JAZU 196 i 197. Zagreb
2. Finka, Božidar; Šojat, Antun. 1973. *O slavonskom dijalektu ekavskoga izgovora u okolici Vinkovaca*. Rasprave Instituta za jezik, 2. Zagreb
3. Varnica, Šime. 1900. *Iz Gradišta*. Zbornik za narodni život i običaje južnih Slavena, V. Zagreb.

Željko Jozić, Institute for Croatian language and linguistics, Zagreb, [zjozic@ihj.hr](mailto:zjozic@ihj.hr)

## UNREPLACED OR PRESERVED JAT' IN THE SLAVONIAN DIALECT

Unreplaced or preserved jat' was introduced into the dialectological literature by Pavle Ivić in his 1957 dissertation, which concerned the speech of the Gallipoli Serbs in Pehčevo, Northern Macedonia. Ivić referred to the research of the Romanian dialectologist Emil Petrovici, "On the speech of Krasovan: A study from South Slavic dialectology" (Bucharest, 1937), which recorded a very similar realization of the Proto-Slavic voice jat' in the speech of the "Karaševio Catholics". However, before Ivić and Petrovici, unreplaced or preserved jat' (although not so named) was recorded in Croatian dialectology as early as 1900 and, a bit later, in 1913. Namely, Šime Varnica was the first to record an "unusual" realization 120 years ago of jat' as a closed ɛ or as ei in the speech of the villagers of Gradište in the Slavonian dialect, as confirmed by Stjepan Ivšić in "Današnji posavski govor" in 1913. The topic of this presentation is thus the confirmation of Ivić's thesis that in the speech of Gradište, an unreplaced Proto-Slavic jat' is still present today, as confirmed and recorded in recent research.

**Keywords:** unreplaced Proto-Slavic jat', Slavonian dialect, Gradište near Županja

### References:

1. Ivšić, Stjepan. 1913. Današnji posavski govor. Rad JAZU 196 i 197. Zagreb
2. Finka, Božidar; Šojat, Antun. 1973. O slavonskom dijalektu ekavskoga izgovora u okolici Vinkovaca. Rasprave Instituta za jezik, 2. Zagreb
3. Varnica, Šime. 1900. Iz Gradišta. Zbornik za narodni život i običaje južnih Slavena, V. Zagreb.

Diana Stolac, Filozofski fakultet Sveučilišta u Rijeci, [diana.stolac@ri.t-com.hr](mailto:diana.stolac@ri.t-com.hr)

## Povijesnojezične teme u nastavi hrvatskoga jezika

Na Filozofskom fakultetu u Rijeci u akademskoj 2019./2020. godini pokrenut je u okviru Nastavničkoga modula novi izborni kolegij za studente diplomskoga studija Hrvatski jezik i književnost – *Teme iz povijesti hrvatskoga književnog jezika u nastavi hrvatskoga jezika*. Polazište u sadržaju kolegija pregled je povijesnojezičnih jedinica u nastavi hrvatskoga jezika u osnovnoj i srednjoj školi (prema važećem nastavnom planu i programu za osnovnu i srednju školu). Studenti stječu nove ishode koji ih cjelovitije pripremaju za nastavnički poziv i posebno za nastavne jedinice vezane za povijest hrvatskoga jezika i povijest hrvatskoga književnog jezika. Iz skupine ishoda u izlaganju izdvajaju se dva: (1) kompetentno odabrati i primijeniti različite vrste podataka i važne pojedinosti iz povijesti hrvatskoga jezika za motivaciju učenika za čitanje tekstova starije pisane baštine i (2) steći zanimanje i pozitivan odnos prema pisanoj baštini u didaktičkoj situaciji i izvan nje. Komentiraju se rezultati rada prve generacije studenata, posebice njihovi timski uradci (pregledi povijesnojezičnih tema u aktualnim udžbenicima, pitanja na natjecanjima i maturi te prijedlozi zadataka za ispitivanje učenikova znanja o povijesnojezičnim temama). Predstavlja se i koncepcija priručnika koji se kao digitalni sadržaj priprema za nastavnike u osnovnim školama, gimnazijama i strukovnim školama. Naglasak je na neknjiževnim tekstovima objavljenima između 17. i 19. stoljeća, koji će pomoći nastavnicima u motivaciji učenika (od zadataka u kajkavskim i štokavskim udžbenicima matematike preko pravila ponašanja za stolom u starokajkavskom bontonu do recepata u ljekarušama). Spoznaje nakon prve provedbe kolegija unose se u pripremu kolegija za sljedeću generaciju studenata.

**Ključne riječi:** hrvatski jezik, povijest jezika, nastava, udžbenik

Literatura:

1. Frančić, A. – Kuzmić, B., Jazik hrvatski – jezične raščlambe starih hrvatskih tekstova, Zagreb 2009.
2. Bratulić, J. i dr., Povijest hrvatskoga jezika, knj. 1., 2., 3., 4., 5., 6., Croatica, Zagreb 2009.-2019. (antologije djela)
3. Stolac, D., Ekološke teme u hrvatskim srednjovjekovnim statutima, u: Ekologija u odgoju i obrazovanju, Gospić, 2004, str. 79-89.
4. Visinko, K., Čitanje: poučavanje i učenje, Zagreb, 2014. (izbor)
5. osnovnoškolski i srednjoškolski udžbenici hrvatskoga jezika i čitanke;  
<http://www.azoo.hr/>; <http://www.ncvvo.hr>

Diana Stolac, Faculty of Humanities and Social Sciences University of Rijeka, [diana.stolac@ri.t-com.hr](mailto:diana.stolac@ri.t-com.hr)

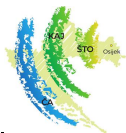
## Historical linguistic topics in Croatian language teaching

In the 2019/2020 academic year, the Faculty of Humanities and Social Sciences in Rijeka launched a new elective course for graduate students of the Croatian language and literature, named Topics from the history of the Croatian literary language in the teaching of the Croatian language. The starting point in the course content is an overview of historical language units in the teaching of the Croatian language in primary and secondary schools (according to their current curriculum). Students acquire new outcomes that prepare them more comprehensively for the teaching profession, and especially for teaching units related to the history of the Croatian language and Croatian literary language. From the group of outcomes in the presentation, two stand out: (1) competently selecting and applying different types of data and important details from the history of the Croatian language to motivate students to read texts from older written heritage; and (2) acquiring interest in and a positive attitude towards the written heritage of the Croatian didactic situation and beyond. This presentation comments on the results of the first generation of students' work, especially their teamwork (reviews of historical language topics in current textbooks, questions at competitions and graduation exams, as well as suggested tasks to test students' knowledge of historical language topics). The concept of a related manual, which is being prepared in digital form for teachers in primary, secondary and vocational schools, is also presented. Its emphasis lies on non-literary texts published between the 17th and 19th centuries, which will help teachers motivate students (from assignments in Kajkavian and Štokavian mathematics textbooks, to table manners in Old Kaikavian etiquette, to prescriptions in medical treatment books). The findings from the first implementation of the course are introduced into the course's preparation for the next generation of students.

**Keywords:** Croatian language, language history, teaching, textbook

References:

1. Frančić, A. – Kuzmić, B., Jazik horvatski – jezične raščlambe starih hrvatskih tekstova, Zagreb 2009.
2. Bratulić, J. i dr., Povijest hrvatskoga jezika, knj. 1., 2., 3., 4., 5., 6., Croatica, Zagreb 2009.-2019.
3. Stolac, D., Ekološke teme u hrvatskim srednjovjekovnim statutima, u: Ekologija u odgoju i obrazovanju, Gospić, 2004, str. 79-89.
4. Visinko, K., Čitanje: poučavanje i učenje, Zagreb, 2014. (izbor)
5. osnovnoškolski i srednjoškolski udžbenici hrvatskoga jezika i čitanke;  
<http://www.azoo.hr/>; <http://www.ncvvo.hr>



Ivan Mance, Hrvatska pošta d.d., Velika Gorica, [ivan.mance@gmail.com](mailto:ivan.mance@gmail.com)

Darko Žubrinić, Fakultet elektrotehnike i računarstva, Zagreb, [darko.zubrinic@gmail.com](mailto:darko.zubrinic@gmail.com)

## Črnkov zapis iz Državnog Arhiva u Beču o obrani Sigeta od Turaka i pogibiji Nikole Zrinskog 1566. g.

Znameniti zapis Ferenc (Franje) Črnka, sačuvan u prijepisu anonimnoga hrvatskog glagoljaša na ozaljskom području, pisan brzopisnom glagoljicom 1566. ili 1567. godine, sadržava točno 40 stranica. Čuva se u Državnom arhivu u Beču, u privatnoj zbirci obitelji Heinricha Auersperga. Dosad su objavljene samo dvije stranice izvornika, u poznatoj studiji Stjepana Ivšića iz 1918. godine. Tekst zapisa pisan je mješavinom triju temeljnih hrvatskih narječja – čakavskog, kajkavskog i štokavskog govora, s povremenim natruhama iz mađarskog, njemačkog, turskog i talijanskog jezika. Tijekom predavanja bit će prvi put prikazano svih četrdeset stranica toga iznimno važnog jezičnog i povijesnog dokumenta, koji su 2019. godine nabavili Ivan Mance i Darko Žubrinić. Radi se o spisu visoke međunarodne vrijednosti, od interesa za poznavanje povijesti ne samo Europe (osobito Hrvatske, Mađarske i Austrije) nego i Turske.

**Ključne riječi:** Franjo Črnko, Nikola Zrinski, Siget, hrvatska glagoljica, Heinrich Auersperg

### Literatura:

1. Stjepan Ivšić: Podsjedanje i osvojenje Sigeta u glagoljskom prijepisu hrvatskoga opisa iz g. 1566. ili 1567., *Starine XXXVI*. 1918. str. 390–429.
2. Mala početnica kurzivne glagoljice, [www.croatianhistory.net/etf/kurziv.html](http://www.croatianhistory.net/etf/kurziv.html) (rujna 2020)



Ivan Mance, Hrvatska pošta d.d., Velika Gorica, [ivan.mance@gmail.com](mailto:ivan.mance@gmail.com)

Darko Žubrinić, Faculty of Electrical Engineering and Computing, Zagreb, [darko.zubrinic@gmail.com](mailto:darko.zubrinic@gmail.com)

## Črnko's record about the defense of Siget from the Turks and the death of Nikola Zrinski in 1566, kept in the State Archive in Vienna

A famous record of Ferenc (Franjo) Črnko, transcribed by an anonymous Croatian Glagolitic priest in the region of Ozalj in Glagolitic quickscript in 1566 or 1567, contains precisely 40 pages. It is kept in the Austrian State Archive in Vienna as part of a private collection belonging to the family of Heinrich Auersperg. Until now, only two pages of this document have been published, in a well-known study by Stjepan Ivšić from 1918. The text was written as a mixture of three Croatian dialects, Chakavian, Kajkavian and Štokavian, with occasional expressions in Hungarian, German, Turkish and Italian. During our lecture, all forty pages of this exceptionally important Croatian language and historical document will be shown for the first time, given their purchase in 2019 by Ivan Mance and Darko Žubrinić. Here, we feature a manuscript of high international value (written in 1566 or 1567), which is of interest for understanding not only the history of Europe (in particular of Croatia, Hungary and Austria), but also the history of Turkey. An extensive study by the same two authors will appear by the end of 2021 in *Dometi* (Achievements, Rijeka). We also draw the reader's attention to the web page [www.croatianhistory.net/etf/crnkogl.html](http://www.croatianhistory.net/etf/crnkogl.html), which contains additional details.

**Keywords:** Franjo Črnko, Nikola Zrinski, Siget, Croatian Glagolitic script, Heinrich Auersperg

### References:

1. Stjepan Ivšić: Podsjedanje i osvojenje Sigeta u glagoljskom prijepisu hrvatskoga opisa iz g. 1566. ili 1567., *Starine XXXVI*. 1918. str. 390–429.
2. Mala početnica kurzivne glagoljice, [www.croatianhistory.net/etf/kurziv.html](http://www.croatianhistory.net/etf/kurziv.html) (September 2020)

Jelena Vignjević, Učiteljski fakultet Sveučilišta u Zagrebu, [jelena.vignjevic@ufzg.hr](mailto:jelena.vignjevic@ufzg.hr)

## Odgojitelji i učitelji pred izazovima hrvatske dijalektne raznolikosti

Hrvatska narječja – čakavsko, kajkavsko i štokavsko – živo su prisutna u svakodnevnoj komunikaciji. Govornici hrvatskoga primarno se koriste svojim dijalektnim govorom, a djeca ga usvajaju kao prvi idiom.

Iskorakom u okolinski svijet, posebno u svijet institucijskoga odgoja i obrazovanja, svaki se govornik susreće i sa standardnim hrvatskim jezikom te njime počinje ovladavati. Usvaja ga već od vrtića, a sustavno ga započinje učiti od prvoga razreda osnovne škole. Tako se dva jezična koda, dijalektni i standardni, kod svakoga govornika na poseban način isprepleću, miješaju i utječu jedan na drugi. Drugim riječima, većina govornika hrvatskoga jezika jezikom ovladava u kontekstu okomite dvojezičnosti. U toj kompleksnoj komunikacijskoj okolini pred odgojiteljima i učiteljima stoji zadatak uspostavljanja kvalitetnih odnosa s djecom i s njihovim roditeljima, kao i poticanje djece u ovladavanju standardnim hrvatskim jezikom, ali i u njegovanju njihova zavičajnog idioma, dijalektnoga govora. Taj zadatak, posebno u onim jezičnim sredinama koje su odgojitelju ili učitelju daleke i slabo poznate, može biti itekako zahtjevan i izazovan.

U ovome se radu stoga promišlja o izazovima poučavanja i odgajanja djece u kontekstu opisane jezične kompleksnosti, o pripremljenosti odgojitelja i učitelja za to, kao i o njihovoj osviještenosti za dijalektnu raznolikost hrvatskoga jezika i konkretnu govornu zbilju u različitim hrvatskim krajevima – mogućim lokacijama njihova učiteljskoga ili odgojiteljskoga rada. U tom su kontekstu anketirani studenti Učiteljskoga fakulteta u Zagrebu na trima studijskim programima: odgojiteljski studij, učiteljski studij te učiteljski studij s engleskim/njemačkim jezikom, a kako bi se stekao uvid u njihovo poznavanje hrvatske dijalektne raznolikosti i govora u različitim hrvatskim krajevima. Rezultati ankete pokazuju da je ukupno gledano znanje studenata o govoru u pojedinim hrvatskim krajevima, posebno u čakavskima, ispod očekivanoga. Za obrazovanje učitelja i odgojitelja u Hrvatskoj ti rezultati ukazuju na potrebu jačanja svjesnosti o hrvatskoj dijalektnoj raznolikosti.

**Ključne riječi:** hrvatska dijalektna raznolikost, odgojitelj, učitelj



Jelena Vignjević, University of Zagreb, Faculty of Teacher Education,  
[jelena.vignjevic@ufzg.hr](mailto:jelena.vignjevic@ufzg.hr)

## Teachers and pre-school teachers facing the challenge of the Croatian dialectal diversity

The three Croatian dialects – Chakavian, Shtokavian and Kajkavian – are very much present in everyday communication in Croatia. Croatian speakers predominantly use their dialect of origin in their daily lives, while children acquire it as their first idiom.

As they step out of their family environment, and especially as they start their pre-school and primary education, all Croatian language speakers encounter and begin to master the standard Croatian language. They acquire it already in kindergarten and systematically start to learn it in the first grade of primary school. This way the two linguistic codes, the dialect and the standard, intertwine in every speaker's linguistic system; they mix and affect one another. In other words, most Croatian language speakers master the Croatian language in the context of vertical bilingualism. In this complex communication environment, teachers and pre-school teachers need to establish good relationships with the children and their parents to stimulate the children to master the standard Croatian language, as well as to cherish their dialect of origin. This can be a difficult and challenging task, especially in language communities that a teacher or a pre-school teacher is not well acquainted with.

Thus, this paper contemplates some challenges teachers and pre-school teachers face when working within linguistically complex contexts; their readiness to face such tasks; their awareness of the dialectal diversity of the Croatian language; and the actual linguistic reality in various parts of the country – all potential places where teachers and pre-school teachers might find employment. To tackle these issues, a questionnaire was administered to students of three study programmes at the Faculty of Teacher Education of the University of Zagreb: pre-school teacher education study, primary teacher education study and primary teacher education study with English/German. The goal was to investigate the students' awareness of the Croatian dialectal diversity present in various parts of the country. The results show that their overall knowledge of the idioms used in different parts of Croatia, and Chakavian in particular, is below expectations. For teacher education in Croatia, these results signify a need to improve awareness of Croatian dialect diversity.

**Keywords:** Croatian dialectal diversity, pre-school teacher, teacher

Ivana Kurtović Budja, Institut za hrvatski jezik i jezikoslovlje, Zagreb, [ikurtov@ihji.hr](mailto:ikurtov@ihji.hr)

## Upitnici za Hrvatski jezični atlas i Zlatna formula hrvatskoga jezika ča-kaj-što na primjeru govora Makarskoga primorja

Hrvatski su mjesni govori istraživani za višejezične atlase: Općeslavenski lingvistički atlas (28 punktova), Europski lingvistički atlas – Atlas linguarum Europae i za srednjojužnoslavenski dijalektološki atlas (pod nazivom Hrvatsko-srpski dijalektološki atlas, u organizaciji Međuakademijskoga odbora za dijalektologiju, s predstavnicima JAZU-a, sada HAZU-a) (236 punktova). Godine 1996. započet je u Institutu za hrvatski jezik i jezikoslovlje projekt Hrvatski jezični atlas (HJA) za koji je predviđena mreža od 399 punktova hrvatskih govora (101 čakavski, 110 kajkavskih, 188 štokavskih). Nastao je izdvajanjem od Hrvatsko-srpskoga dijalektološkog atlasa. Mjesni se govori sve ubrzanije mijenjaju, smrću posljednjih govornika oni i nestaju te je krajnji čas da se istraži prema jedinstvenome upitniku ukupno hrvatsko jezično blago. Istraživanja za Hrvatski jezični atlas provode se na temelju Upitnika koji je za potrebe projekta HRZZ-a, Govori Makarskoga primorja – dijakronija i sinkronija, digitaliziran i građa u njemu prikupljena postaje pretraživa prema fonološkim, akcenatskim, morfološkim, leksičkim i semantičkim kriterijima. Članak „Upitnici za Hrvatski jezični atlas i Zlatna formula hrvatskoga jezika ča-kaj-što na primjeru govora Makarskoga primorja“ govori o tome koliko se u govorima Makarskoga primorja potvrđuju temeljne odrednice Zlatne formule hrvatskoga jezika ča-kaj-što. One se najbolje vide na leksičkoj razini. Rad dokazuje da hrvatski mjesni govori čuvaju neprekinutu nit razvoja hrvatskoga jezika od najranijih njegovih potvrda.

**Ključne riječi:** upitnici za dijalektološko istraživanje, Zlatna formula, govori Makarskoga primorja

### Literatura

1. BROZOVIĆ, DALIBOR (1970). O Makarskom primorju kao jednom od središta jezično-historijske i dijalekatske konvergencije. Makarski zbornik 1: 381-405.
2. ČILAŠ ŠIMPRAGA, ANKICA (2002). Fonologija ikavskih štokavskih govora između rijeke Krke i Neretve, u: Celinić, Anita; Kurtović Budja, Ivana; Čilaš Šimpraga, Ankica; Jozić, Željko, Prinosi hrvatskoj dijalektnoj fonologiji. Split – Zagreb: Književni krug Split i Institut za hrvatski jezik i jezikoslovlje.
3. KURTOVIĆ BUDJA, IVANA (2009). Govori Makarskoga primorja - fonologija i morfologija, doktorska disertacija, Filozofski fakultet, Zagreb, u rukopisu.
4. ŠUNDE, SMILJANA (2001). Sutikla – crkva, legende i običaji u Podgori. vlastita naklada.

Ivana Kurtović Budja, Institute of Croatian Language and Linguistics, Zagreb, [ikurtov@ihji.hr](mailto:ikurtov@ihji.hr)

## Questionnaires for the Croatian Language Atlas and the Golden Formula of the Croatian language ča-kaj-što – Dialect of Makarska coast

Local speeches are increasingly changing. They disappear with the deaths of the last speakers, meaning it is the "last minute" to investigate the total Croatian linguistic treasure according to a unique questionnaire. Accordingly, this paper reports research on local Croatian speeches for multilingual atlases: the General Slavic Linguistic Atlas (28 reference points), the European Linguistic Atlas – Atlas linguarum Europae and the Central-South Slavic Dialectological Atlas (also called the Croatian-Serbian Dialectological Atlas, organized by the InterAcademy Committee on Dialectology with representatives of JAZU, now HAZU; 236 reference points). In 1996 the Institute of Croatian Language and Linguistics first started the project Croatian Linguistic Atlas (HJA), for which a network of 399 Croatian speech reference points were planned (101 Čakavian, 110 Kajkavian, 188 Štokavian). It was created when separated from the Croatian-Serbian Dialectological Atlas. This research was carried out through a questionnaire, which was later digitized for the HRZZ project "Makarska littoral speeches – Diachrony and synchrony". The material collected from the questionnaire became searchable according to phonological, accentual, morphological, lexical and semantic criteria. This current paper now shows how the fundamental determinants of the Golden Formula are confirmed in the dialect of the Makarska Coast, best seen at the lexical level. It also proves that Croatian local speeches maintain the continuous thread of the Croatian language's development from its earliest affirmation.

**Keywords:** questionnaires for dialectological research, Golden Formula, dialect of the Makarska coast

### References:

1. BROZOVIĆ, DALIBOR (1970). O Makarskom primorju kao jednom od središta jezično-historijske i dijalekatske konvergencije. Makarski zbornik 1: 381-405.
2. ČILAŠ ŠIMPRAGA, ANKICA (2002). Fonologija ikavskih štokavskih govora između rijeke Krke i Neretve, u: Celinić, Anita; Kurtović Budja, Ivana; Čilaš Šimpraga, Ankica; Jozić, Željko, Prinosi hrvatskoj dijalektnoj fonologiji. Split – Zagreb: Književni krug Split i Institut za hrvatski jezik i jezikoslovlje.
3. KURTOVIĆ BUDJA, IVANA (2009). Govori Makarskoga primorja - fonologija i morfologija. Filozofski fakultet, Zagreb, manuscript.
4. ŠUNDE, SMILJANA (2001). Sutikla – crkva, legende i običaji u Podgori.

Antonio Sammartino, zaklada "Agostina Piccoli", [anton.sammartino@gmail.com](mailto:anton.sammartino@gmail.com)

## Moliškohrvatski idiom: kvantitativna analiza dijafaznih varijacija u dvojezičnom kontekstu

U radu se prikazuju dosadašnji rezultati kvantitativnih analiza dijafaznih varijacija u moliškohrvatskom dvojezičnom kontekstu. Korpus se sastoji od rječnika, gramatike, etnografskih istraživanja, lingvističkih studija i govornih zapisa snimljenih od 1999. godine.

Moliškohrvatski govor pripada štokavsko-ikavskom narječju, čuva se u trima selima: Mundimitar (tal. Montemitro), Filič (tal. San Felice del Molise) i Kruč (tal. Acquaviva Collecroce) u pokrajini Molise (Italija) i broji oko 1500 govornika.

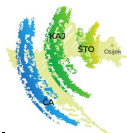
Petstoljetna izolacija od pradomovine i kontakt s talijanskim jezičnim okruženjem rezultirali su inovacijama u moliškohrvatskom sustavu. Na leksičkoj se razini istovremeno bilježi spori gubitak hrvatskog materijala te obogaćenje talijanskim posuđenicama koje se fonološki i morfološki prilagođavaju jeziku primatelju.

Nerijetko su talijanske posuđenice sinonimi već postojećih leksema, a moliškohrvatski govornici pokazuju tendenciju prema jezičnoj ekonomiji tako da u neformalnim situacijama češće biraju talijanske posuđenice, odnosno sinonime hrvatskih leksema. Stoga se događa da talijanski leksički materijal zamijeni hrvatski, kao što je to sa sinonimskim parovima: pisat (s hrvatskog) često se zamjenjuje sinonimom skrivit (s talijanskog) ili razumit s akapit, nedilja s domenika, počinit s ripozat, itd.

Uvijek čuvajući hrvatske temelje, moliški dijalekt „kroatizirao“ je sve što je ušlo u njegovu sferu, ponekad se čak i obogatio u leksiku. S morfološkog se stajališta moliškohrvatski može zasigurno smatrati manje propusnim, doista dobro ukorijenjenim i spremnim za primanje bilo kakve jezične novosti.

Analiza s jedne strane pokazuje ranjivost moliškohrvatskog idioma, izloženog neizbježnoj evoluciji. S druge strane, zaključuje se da moliškohrvatski sustav ni nakon pola tisućljeća nije zasićen.

**Ključne riječi:** Molise, moliškohrvatski idiom, dijafazija, štokavsko-ikavski govor



#### Literatura:

1. Sammartino A., Piccoli A. – Dizionario dell'idioma croato-molisano di Montemitro/Rječnik moliškohrvatskoga govora Mundimitra. Montemitro/Zagreb 2000
2. Breu W., Piccoli G. – Dizionario croato molisano di Acquaviva Collecroce – Dizionario plurilingue della lingua slava della minoranza di provenienza dalmata di Acquaviva Collecroce in Provincia di Campobasso. Campobasso 2000
3. Menac-Mihalić M., Sammartino A. – Adaptacija talijanizama u moliškohrvatskom govoru Mundimitra (u: Hrvatski dijalektološki zbornik. Knjiga 12 – HAZU). Zagreb 2003
4. Sammartino A. – Grammatica della lingua croato-molisana/Gramatika moliškohrvatskoga jezika. Montemitro-Zagreb 2004
5. Razni autori – S našimi riči – Raccolta di componimenti in croato molisano/Zbirka literarnih ostvarenja na moliškohrvatskome. Montemitro 2004, 2007, 2010, 2013, 2016, 2019

Silvija Ćurak, Filozofski fakultet Sveučilišta J. J. Strossmayera u Osijeku, [scurak@ffos.hr](mailto:scurak@ffos.hr)

Vera Blažević Krezić, Filozofski fakultet Sveučilišta J. J. Strossmayera u Osijeku, [vblazevic1@ffos.hr](mailto:vblazevic1@ffos.hr)

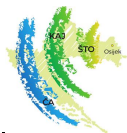
## U rič sakrita Šokadija, med Đakova i Vinkovci položita

Ovomu je izlaganju zadatak interpretirati šokačke priče i *pismice* kojima je autorica Milica Lukić, osječka sveučilišna profesorica, jezikoslovka i kroatistica, Đakovštine zavičajnica, bogatila stranice „Revije Đakovačkih vezova“, etnografsko-popularnoga časopisa koji od 1970. godine raznorodnim tekstovima narodne tematike prati istoimenu međunarodnu smotru izvornoga folklora, ilinštakov iliti srpanjski ures položen u samo srce Šokadije, Strossmayerovim koordinatama omeđene Slavonije.

Obzoru najavljene jezično-stilske analize priveli smo pet pripovjedaka, od čega jednu standardnojezičnu, koja se može okarakterizirati kao fragmentarna zavičajna autobiografija znakovita naziva „Ravničarski bluz“ ili „Jedan zakašnjeli oproštaj s tugom“, odnosno četiri nestandardnojezične, u potpunosti ispisane materinskim šokačkim govorima: mrzovićkim ikavsko-jekavskim i starim ikanovačkim ikavsko-ekavskim govorom, koje suvremena hrvatska dijalektologija imenuje govorima tzv. posavskoga poddijalekta staroštokavskoga slavonskog dijalekta, crpeći iz njihove snage na zalazu vrijednu dijalektološku građu, pravu kulturnu dragocjenost čija su razlikovna obilježja aktivirana u izgradnji hrvatskoga nacionalnog i kulturnog identiteta, a najzad uvrštena i u Zlatnu formulu hrvatskoga jezika ča-kaj-što. Četirima šokačkim pričama – „Srce dvaput pripuknito“, „Od cviča moji uspomena“, „Miris svibanjski ružica i zumbula na zalazu“, „Od zavjeta na svetoj Vodici Gojspinoj primitog“ – pridružila se i jedna *pismica*, intermedijalna šokačka varijacija, zavičajem začeta, filološkim obrazovanjem autorice (ot)hranjena „Kvintesencija panonska“.

Tako zaokružen korpus podvrći će se zakonitostima jezično-stilske analize, s naglaskom na fonološke, morfološke, sintaktičke i leksičke značajke slavonskoga dijalekta te osebujne šokačke *riči* koja u književnim tekstovima iz 2010-ih zapravo skladišti ostavštinu naših predaka – Šokaca starinaca, poglavito Šokica – porodičnih stanovnica od starine, u Đakovštini i Đakovštinom naseljenih, zbog čega se tekstovi Milice Lukić promatraju kao svojevrsan muzej riječi podignut slavonskim narodnim govorima i običajima u čast; kao dragocjen prilog suvremenoj dijalektnoj književnosti; angažirano žensko i etnografsko pismo; ukratko – uspio znanstveno-umjetnički eksperiment kojemu je zadaća, kako to autorica opetovano ističe, uzajamno prožimati jezik znanosti, umjetnosti i duhovnosti, svih civilizacijskih vrednota koje nas obvezuju, ali na koje – kao sudionici globalne zajednice, uostalom i one istomislećih i istogovorećih – sveudilj polažemo pravo.

**Ključne riječi:** slavonski dijalekt, suvremena zavičajna književnost, šokačke priče i *pismice*, jezično-stilska analiza



Silvija Ćurak, Faculty of Humanities and Social Sciences University of Osijek, [scurak@ffos.hr](mailto:scurak@ffos.hr)

Vera Blažević Krezić, Faculty of Humanities and Social Sciences, Univ. of Osijek, [vblazevic1@ffos.hr](mailto:vblazevic1@ffos.hr)

## Between Đakovo and Vinkovci Placed Šokadija – Hidden in the Word

This presentation aims to interpret select short stories and poems by Milica Lukić, a Croatian linguist, philologist, paleoslavist and professor at the University of Osijek. The stories, written partly in the Slavonian dialect, were first published in *The Đakovo Cuts Journal*, a popular ethnographic magazine that follows the international folklore placed in the very heart of Strossmayer's Slavonia.

This linguistic-stylistic analysis encompasses five stories, one written in standard Croatian and the others in the Slavonian dialect, or, more precisely, the Ikavian-Jekavian idiom spoken in Mrzović and the Ikavian-Ekavian idiom spoken in Stari Mikanovci. In modern Croatian dialectology, both are recognized as speeches of the Posavina subdialect of the Old Štokavian Slavonian dialect. The stories are accompanied by a dialect poem entitled "Pannonian Quintessence" (Kvintesencija panonska).

The current analysis focuses on the phonological, morphological, syntactic and lexical features of the Slavonian dialect apparent in Lukić's texts, which are viewed not only as guardians of tradition, but also as valuable contributions to contemporary dialectal literature and engaged women's and ethnographic writing. In short, the stories can be labelled as a successful scientific-artistic experiment whose task, as Lukić herself repeatedly accentuates, is to intertwine the language of science, art and spirituality, all civilizational values which oblige us, but to which we also claim our right.

**Keywords:** Slavonian dialect, contemporary Croatian dialectal literature, short stories and poems, linguistic-stylistic analysis

Biserka Goleš Glasnović, Osnovna škola Tituša Brezovačkog, Zagreb,  
[biserkagg@yahoo.com](mailto:biserkagg@yahoo.com)

## Smutica za zdrav(l)je hrvatskog jezika

Tekst tumači dijakronijsku i sinkronijsku povezanost triju hrvatskih narječja i zagovara Zlatnu formulu hrvatskog jezika ča-kaj-što kao teorijsku i djelatnu aktivnost u stvaralaštvu njezina autora, hrvatskog književnika i diplomata Drage Štambuka. U dijakronijskom slijedu poziva se na djela Ozaljskoga književnog kruga, djela Tituša Brezovačkog i opsežnije analizira jezik Josipa Šipuća u djelu „Temelji žitne trgovine“<sup>1</sup>. U sinkronijskom pregledu rad se osvrće na Štambukovu zbirku pjesama „Kad su miši balali molfrinu“, svojevrsni pjesnički manifest Zlatne formule te interpretira dvije Štambukove metafore koje Formulu prikazuju kao molekulu vode i mitsko stablo s tri grane.

U sudbonosno vrijeme kada i jezik treba svjedočiti o pravu naroda na samostalnost, upravo je Drago Štambuk osvijestio i osvjetlio pravu narav hrvatskoga jezika – vrlo djelatno. Godine 1991. osnovao je pjesničku manifestaciju Croatia rediviva: ča-kaj-što u Selcima na Braču, na kojoj već 29 godina sudjeluju pjesnici štokavci, kajkavci i čakavci.

Vjera u riječ Štambuka pjesnika, theurga i jezikofila jednako je snažna u njegovu poetskom govoru kao i u traženju identiteta i definiranju naravi hrvatskog jezika. Zlatna je formula poput zlatnoga reza kanon koji se ne mijenja, a njezina trodjelnost nedjeljiva poput sakralnoga trojstva. Kozmogonijski pak slijedi kršćanska vjerovanja budući da na početku bijaše Riječ, ali i mnoga koja im prethode pa će je predočiti i slikom mitskog stabla s tri grane ča-kaj-što, koje nosi svijet, narod, jezik i pojedinca. Na stilističkoj razini ta je slika svojim alegorijskim i hiperboličnim prikazom barokna, gotovo preuzetna, no štambukovski šarmantna jer istodobno s njezinim uzvišenim i svečanim tonom, korespondira s uzet(n)om sličicom skromnoga bračkog napitka iz svakodnevnoga tegobnog života, smuticom od mlijeka i crnoga vina. Spoj bliskog i dalekog, malog i velikog, stvarnog i mitološkog – ona je začudnost koju prepoznajemo i u Štambukovoj poetici – i u jeziku kao sredstvu kojim se obraćamo vidljivom i nevidljivom svijetu.

**Ključne riječi:** Croatia rediviva, smutica, trojezičnost



Biserka Goleš Glasnović, Osnovna škola Tituša Brezovačkog, Zagreb,  
[biserkagg@yahoo.com](mailto:biserkagg@yahoo.com)

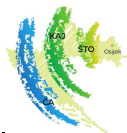
## Smoothie for the health of the Croatian language

This text explores the diachronic and synchronic connections between the three Croatian dialects and justifies the "Golden Formula" of the Croatian language, or *ča-kaj-što* (Engl. what), advocated as the definitive theoretical and practical activity in the work of its creator, Croatian writer and diplomat Drago Štambuk. This paper's diachronic sequence makes reference to the works of the Ozalj Literary Circle, those of Tituš Brezovački and, in a more detailed analysis, the language used by Josip Šipuš in his work *The foundation of the grain trade*.<sup>1</sup> In its synchronic review, the paper also refers to Štambuk's collection of poems *When mice danced the Molfrina*, effectively a poetry manifesto in favour of the Golden Formula utilising Štambuk's two metaphors, which depict the Formula as a water molecule or a mythical tree with three branches.

In challenging times, when the Croatian language should have also testified to the right of the people to independence, it was Štambuk who actively established and highlighted the language's true nature. In particular, in 1991 he founded an annual poetry manifesto festival called Croatia rediviva: *ča-kaj-što* in Selce on the Island of Brač, in which many poets from Štokavian, Kajkavian and Čakavian regions still participate.

Štambuk as poet, theurgist and linguaphile manifests a belief in the word that is just as strong in his poetic speech as in his search for the identity and defining nature of the Croatian language. The Golden Formula, like the Golden Ratio, is a canon that does not change, and its trinity is indivisible, just like the holy trinity. Still, Štambuk's cosmogony, while following Christian beliefs, since "in the beginning there was the Word", nonetheless follows many other beliefs that preceded Christianity. Accordingly, he represents this concept with the image of a mythical tree with three branches – *ča-kaj-što* – that carry the world, the people, the language and the individual. On a stylistic level, this word-painting is baroque in its allegorical and hyperbolic imagery, almost presumptuously so, but it also has a distinctively Štambukian charm: at the same time as it presents a sublime and solemn tone, it corresponds with a picture capturing a modest Brač drink composed of daily burdensome life, a smoothie made of milk and red wine. The combination of the near and the far, the small and the large, the real and the mythological – this is exactly the estrangement that we recognise in Štambuk's poetics, and in language as a means by which we address both the visible and invisible worlds.

**Keywords:** Croatia rediviva, smoothie, trilingualism



## Literatura

1. Šipuš, Josip, Temelj žitne trgovine (Temely xitne tergovine – polag – narave y dogacsajev razborito – po – Joseff Sipus – Horvatsanu Karlovacskome. – vu Zagrebu, 1796.). Pretisak Matica Hrvatska ogranak Karlovac, 1993.
2. Goleš Glasnović, Biserka (2015) Spomenek na Tituša Brezovačkog. U: Monografija Osnovne škole Tituša Brezovačkog. Str. 157. – 159.
3. Goleš Glasnović, Biserka (2019.) Duhovni versovnik. U: Forum. Zagreb. Hrvatska akademija znanosti i umjetnosti. Br. 10-12. Str. 1344. – 1350.
4. Štambuk, Drago, Theurgia, Školska knjiga, 2017.
5. Štambuk, Drago, Kad su miši balali molfrinu, Naklada Đuretić, 2017.

## References

1. Šipuš, Josip, Temelj žitne trgovine / The foundation of the grain trade (Temelj žitne trgovine – polag – narave y dogacsajev razborito – po – Joseff Sipus – Horvatsanu Karlovacscome. – vu Zagrebu, 1796.). Reprint by Matica Hrvatska, Karlovac branch, 1993.
2. Goleš Glasnović, Biserka (2015) Spomenek na Tituša Brezovačkog / A memorial to Tituš Brezovački. In Monograph of the Elementary School of Tituš Brezovački, p. 157 – 159.
3. Goleš Glasnović, Biserka (2019) Duhovni versovnik / Spiritual Hymnal. In Forum. Zagreb. Hrvatska akademija znanosti i umjetnosti. No. 10-12, p. 1344 – 1350.
4. Štambuk, Drago, Theurgia, Školska knjiga, 2017.
5. Štambuk, Drago, Kad su miši balali molfrinu / When Mice Danced the Molfrina, Naklada Đuretić, 2017.

Lobel Filipić, Institut za hrvatski jezik i jezikoslovlje, Zagreb, [lobel.filipic@gmail.com](mailto:lobel.filipic@gmail.com)

Marija Čubelić, 7. gimnazija, Zagreb, [marija.cubelic55@gmail.com](mailto:marija.cubelic55@gmail.com)

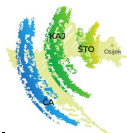
## Zlatna formula hrvatskoga jezika u Hrvatskome zvučnom atlasu

Zlatna formula hrvatskoga jezika ča-kaj-što predstavlja raznolikost hrvatske dijalektalne stvarnosti, odnosno strukturu hrvatskoga jezika koji se iz narječja dijeli na dijalekte, dijalekti na skupine govora, a skupine govora na mjesne govore. U općoj komunikaciji potiče se uvjerenje o različitosti mjesnih govora, o idiomima koji su značajno udaljeni pa se poneki primjeri te različitosti anegdotalno redovito ponavljaju. Za takvo stajalište uvelike ima opravdanja i te su tvrdnje istinite, a i potiču da se hrvatski jezik promatra kao skup svih idioma, koliko god oni različiti bili, ako im je zajedničko obilježje to što pripadaju hrvatskoj jezičnoj strukturi. Međutim, upravo zbog tog naglašavanja različitosti dijelom ostaju nezabilježene sličnosti koje i jesu opravdanje pripadnosti mjesnih govora istome višem idiomu, odnosno hrvatskomu jeziku. Ovaj se rad bavi proučavanjem leksičkih sličnosti u mjesnim govorima različitih narječja. Rad se bavi građom Hrvatskoga zvučnog atlasa u kojoj se na referentim točkama skupljala jezična baština. Ona je snimljena te predstavljena na nekoliko jezikoslovno utemeljenih razina koje su važne za proučavanje dijakronijskih i sinkronijskih stanja hrvatskoga jezika. Zvučni zapisi podijeljeni su po određenim dijalektološkim obilježjima; odraz pojedinih praslavenskih glasova, sudbina određenih glasova ovisno o mjestu u riječi, naglasni inventar, morfološka obilježja te riječi različite od njihovih istoznačnica u hrvatskome standardnom jeziku. Ovo potonje temelj je analize ovoga rada. Svrha je istraživanja dokazivanje leksičke sličnosti hrvatskih govora. Proučit će se koliko je sličnost česta, u kojim se govorima pojavljuje, jesu li to govori dodirnih područja ili prostorno udaljenih idioma. Na temelju tih rezultata moći će se istražiti moguće povijesne promjene koje su dovele do leksičke sličnosti govora, odnosno utjecaj migracija, dominantnosti određenoga narječja, povezanosti postojanja pisane literature koja je doprinijela širenju određenih leksema. Analiza će se temeljiti na bilježenju pojavnosti određenoga leksema u određenome govoru, odnosno za svaki će se leksem mjeriti u koliko je različitih mjesnih govora zabilježen dostupnih u Hrvatskome jezičnom atlasu.

**Ključne riječi:** zlatna formula hrvatskoga jezika, Hrvatski zvučni atlas, mjesni govor, leksik

### Literatura

1. Brozović, D. O tronarječnoj dimenziji hrvatske književnosti. *Croatiaca: časopis za hrvatski jezik, književnost i kulturu*. Vol. 7 No. 7-8, 1976.
2. Katičić, R. Zlatna formula hrvatskoga jezika: ča-kaj-što. *Kolo: časopis Matice hrvatske za književnost, umjetnost i kulturu*. Zagreb. 2014.
3. Težak, S. Ozaljski govor. *Hrvatski dijalektološki zbornik* 5. Zagreb 1981.



Lobel Filipić, Institute of Croatian Language and Linguistics, Zagreb, [lobel.filipic@gmail.com](mailto:lobel.filipic@gmail.com)

Marija Čubelić, 7<sup>th</sup> Gymnasium, [marija.cubelic55@gmail.com](mailto:marija.cubelic55@gmail.com)

## The Golden formula of the Croatian language in the Croatian Sound Atlas

This paper represents the diversity of the Croatian dialect reality, that is, the structure of the Croatian language as divided into dialects and local speeches. It widely encourages belief in the diversity of local speeches, in idioms that are significantly distant, via examples and differences that repeatedly appear in Croatian literature and general communication. There are many valid justifications for this view that promote the view of the Croatian language as a set of all of its idioms, however different they may be, if their common feature is that they belong to the standard Croatian linguistic structure. However, it is because of the emphasis on the dialects' differences that their similarities remain in part unrecorded. These similarities justify the belonging of local speeches to the same higher idiom, or standard Croatian. This paper thus addresses the lexical similarities in local speeches of different Croatian dialects. It also studies the structure of the Croatian Sound Atlas, which assembles Croatian language heritage from reference points across the country. The material is recorded and presented at several language-based levels important to the study of the diachronic and synchronic state of the Croatian language. The audio recordings are divided according to certain dialectological characteristics: the reflection of certain pre-Slavic voices, the fate of certain voices depending on their position in a word, an accent inventory, morphological characteristics and words different from their equivalents in standard Croatian. The latter is the basis for this paper's analysis, whose purpose is to prove the lexical similarity of various Croatian speeches recorded in the Croatian Sound Atlas. The paper examines how common similarities are, in which speeches they appear and whether these are the speeches of contact areas or spatially distant idioms. The results then allowed an investigation of the historical changes that led to the lexical similarity of certain speeches, including the influence of migrations, the dominance of a certain dialect and the existence of written literature that contributed to the spread of certain lexemes.

**Keywords:** the Golden Formula of the Croatian language, Croatian Sound Atlas, local speech, vocabulary

### References:

1. Brozović, D. O tronarječnoj dimenziji hrvatske književnosti. *Croatiaca: časopis za hrvatski jezik, književnost i kulturu*. Vol. 7 No. 7-8, 1976.
2. Katičić, R. Zlatna formula hrvatskoga jezika: ča-kaj-što. *Kolo: časopis Matice hrvatske za književnost, umjetnost i kulturu*. Zagreb. 2014.
3. Težak, S. Ozaljski govor. *Hrvatski dijalektološki zbornik* 5. Zagreb 1981.



Marija Čubelić, 7. gimnazija, Zagreb, [marija.cubelic55@gmail.com](mailto:marija.cubelic55@gmail.com)

Lobel Filipić, Institut za hrvatski jezik i jezikoslovlje, Zagreb, [lobel.filipic@gmail.com](mailto:lobel.filipic@gmail.com)

## Mjesni govori kao baština u jezičnoj politici i odgojno-obrazovnom procesu

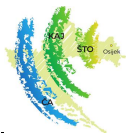
U vremenu koje obilježavaju intenzivne migracije i globalizacija, a smanjuje se broj govornika malih govora, u okviru državnih politika razmatraju se mogućnosti očuvanja jezično-kulturnoga identiteta. Škole kao mjesta odgoja i obrazovanja novih generacija imaju važnu ulogu, posebice nastava jezika kao dio jezične politike. Smatrali jezičnu politiku dijelom jezičnoga planiranja ili nadređenim mu pojmom, ona obuhvaća ideje, zakone, prakse na koje utječu određeni autoriteti kako bi se mijenjao jezik u društvu. Nastava Hrvatskoga jezika izvodi se na temelju službenih dokumenta (planova i programa te kurikulā) Ministarstva znanosti i obrazovanja Republike Hrvatske u kojima se očituje mjesto standardnih i nestandardnih idioma, odnosno cjelokupne jezične baštine u obrazovnoj jezičnoj politici. Baštinu se pritom u literaturi ne povezuje samo s prošlošću pojedine društvene zajednice nego se ističe i njezin utjecaj na sadašnjost i budućnost zajednice. Za mnoge mjesne govore u Ministarstvu kulture Republike Hrvatske utvrđeno je da imaju svojstvo nematerijalnoga kulturnog dobra (termin koji nalazimo u hrvatskim dokumentima), a Zlatnom formulom ča-kaj-što, kulturnim dobrom kojim se ističe međupovezanost hrvatskih narječja, nastojat će se zaštititi i drugi neupisani, ugroženi govori. Mjere zaštite istaknute u tom zakonskom okviru ne zaobilaze odgojno-obrazovni proces pa se odnose i na

„mogućnost prenošenja tradicije nasljednicima putem formalnog i neformalnog obrazovanja“ (Rješnje 2019.). U ovom se radu na temelju dokumenata prikazuju glavni razlozi potrebe za zaštitom mjesnih govora u danom trenutku, kao i tko se zalaže za njihovo čuvanje. Analizira se diskurs propisanih mjera zaštite s posebnim osvrtom na one koje se odnose na odgojno-obrazovni sustav te se istražuje i položaj jezične baštine u nacionalnim odgojno-obrazovnim dokumentima. Navode se prijedlozi za provođenje mjera u školama po uzoru na dosad provedene aktivnosti.

**Ključne riječi:** mjesni govori, baština, jezična politika, nastava Hrvatskoga jezika

Literatura:

1. Jelinčić, Daniela Angelina. 2010. Kultura u izlogu: kratki vodič za upravljanje kulturnim dobrima. Meandar. Zagreb.
2. Kaplan, Robert B.; Baldauf, Richard B. 1997. Language Planning: From Practice to Theory. Multilingual Matters. Clevedon.
3. Rješnje Ministarstva kulture, 29. 11. 2019.
4. Spolsky, Bernard. 2009. Language Management. Cambridge University Press. New York.



Marija Čubelić, 7<sup>th</sup> Gymnasium, [marija.cubelic55@gmail.com](mailto:marija.cubelic55@gmail.com)

Lobel Filipić, Institute of Croatian Language and Linguistics, Zagreb, [lobel.filipic@gmail.com](mailto:lobel.filipic@gmail.com)

## Local speeches as a heritage in language policy and educational process

In times marked by intense migration, globalization and decreasing numbers of small language speakers, the preservation of linguistic and cultural identity is often considered within the framework of state policies. Schools as places of education for new generations play an important role in this regard, especially the teaching of languages as part of language policy. Whether language policy is considered within language planning or above it, it encompasses ideas, laws and practices influenced by certain authorities to change language in society. Teaching of the Croatian language in particular is carried out on the basis of official documents (plans, programmes and curricula) from the Croatian Ministry of Science and Education, which describe the place of standard and non-standard idioms, that is, the entire language heritage, in educational language policy. In the literature, heritage is not only associated with the past of a particular social community, but also its impact on the present and the future of the community. Many local speeches in the Croatian Ministry of Culture have also been established as intangible cultural property, and the Golden Formula *ča-kaj-što*, a cultural property that emphasizes the interconnectedness of Croatian dialects, seeks to protect unregistered and endangered speeches in particular. Such protection measures prescribed in the legal framework do not bypass the education process, and rather refer to the "possibility of transmitting tradition to inheritors through formal and informal education" (Resolution 2019). This paper presents the main reasons for the need to protect local speeches, as well as who advocates their preservation. It analyses the discourse of prescribed protection measures with special emphasis on those related to the Croatian education system and the position of language heritage in national educational documents. The paper's final suggested implementation of measures in schools is modelled on the activities undertaken so far.

**Keywords:** local speeches, heritage, language policy, Croatian language education

References:

1. Jelinčić, Daniela Angelina. 2010. *Kultura u izlogu: kratki vodič za upravljanje kulturnim dobrima*. Meandar. Zagreb.
2. Kaplan, Robert B.; Baldauf, Richard B. 1997. *Language Planning: From Practice to Theory*. Multilingual Matters. Clevedon.
3. Rješenje Ministarstva kulture, 29. 11. 2019.
4. Spolsky, Bernard. 2009. *Language Management*. Cambridge University Press. New York.

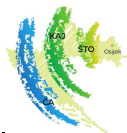


Vlatka Štimac Ljubas, Leksikografski zavod Miroslav Krleža, Zagreb, [vlatka.stimac@lzmk.hr](mailto:vlatka.stimac@lzmk.hr)

## Mjesto višerječnih farmaceutskih naziva u općejezičnim rječnicima hrvatskoga jezika

U novijim općejezičnim rječnicima hrvatskoga jezika farmaceutsko-farmakološki nazivi nisu sustavno i kriterijski dosljedno popisani, a ni leksikografski obrađeni. Nazivi s područja farmaceutske struke dijelom su kemijskoga i medicinskoga nazivlja, ali i naziva drugih prirodoslovnih disciplina. Stoga leksikografski popis i opis farmaceutskih naziva u općejezičnome rječniku hrvatskoga treba provesti u suradnji s predmetnim stručnjacima jer kada on ovisi iznimno o uredniku/leksikografu koji obrađuje određenu struku, može biti subjektivan. Stoga će se u članku ponajprije istražiti koje se natuknice u općejezičnim rječnicima hrvatskoga određuju strukovnom odrednicom (*farm*), odnosno prema kojim kriterijima i koliko sustavno ti rječnici donose farmaceutske nazive. Istraživanje će se provesti prema korpusu novijih hrvatskih jednojezičnih rječnika („Rječnik hrvatskoga jezika“ Leksikografskoga zavoda i Školske knjige, 2000.; „Veliki rječnik hrvatskoga jezika“ Vladimira Anića, 2006. i „Hrvatski enciklopedijski rječnik“ Novog Libera, 2003. te „Veliki rječnik hrvatskoga standardnog jezika“ Školske knjige, 2015.), a strukovna će se relevantnost određivati s obzirom na farmakološko nazivlje pretraživo u sustavu terminološke baze Strune. Budući da su u hrvatskim jednojezičnim rječnicima natuknice uglavnom jednorječne, a većina pak farmakološkoga nazivlja u Struni višerječna, postavlja se i pitanje odnosa: analizira se leksikografska obradba višerječnih naziva sa strukovnom odrednicom *farm* u rječnicima (burova voda, farmaceutska tehnologija, generički lijek, gotovi lijekovi), a koji možda nisu u Struni te se na kraju donosi odgovor na pitanje na kojoj leksikografskoj razini višerječne nazive iz Strune uvrstiti u buduće općejezične rječnike hrvatskoga.

**Ključne riječi:** farmaceutsko-farmakološko nazivlje, općejezični rječnici hrvatskoga jezika, natuknice, leksikografska obradba, Struna



Vlatka Štimac Ljubas, The Miroslav Krleža Institute of Lexicography, [vlatka.stimac@lzmk.hr](mailto:vlatka.stimac@lzmk.hr)

## The lexicographical level of multi-word pharmaceutical terms in general-language Croatian dictionaries

In more recent standard Croatian dictionaries, pharmaceutical-pharmacological terms are not listed systematically according to consistent criteria, nor are they treated lexicographically. The nomenclature in the pharmaceutical profession forms part of that of chemistry and medicine, but also of other natural science disciplines. Therefore, there is a pressing need for a lexicographical list and description of pharmaceutical terms in a standard Croatian language dictionary, created in collaboration with experts in related fields to avoid the subjectivity that often emerges in a list created by one editor/lexicographer. As such, this article first explores which entries in standard Croatian dictionaries are given a specialized vocabulary label (i.e. *farm*) based on what criteria and how consistently they present pharmaceutical terms. The research is based on the corpus of recent Croatian-Croatian dictionaries (Rječnik hrvatskoga jezika, Leksikografski zavoda and Školska knjiga, 2000; Veliki rječnik hrvatskoga jezika by Vladimir Anić, 2006; Hrvatski enciklopedijski rječnik, Novi Liber, 2003; and Veliki rječnik hrvatskoga standardnog jezika, Školska knjiga, 2015), while the individual terms' relevance within the professional field is determined in reference to the pharmacological nomenclature searchable in the Croatian professional terminology database Struna. Since single-language Croatian dictionaries mostly offer single-word entries, and the majority of pharmacological terms in Struna are multi-word, this also raises the question of their relationship. Accordingly, this article analyses the lexicographical treatment of multi-word dictionary entries bearing the label *farm* (*burova voda*, *farmaceutska tehnologija*, *generički lijek*, *gotovi lijekovi*) that might not appear in Struna. Lastly, it suggests on which lexicographical level multi-word terms from Struna should be included in future standard Croatian dictionaries.

**Keywords:** pharmaceutical-pharmacological nomenclature, standard Croatian dictionaries, dictionary entries, lexicographical treatment, Struna

Milan Kranjčević, Katedra Čakavskog sabora pokrajine Gacke, Otočac, [mcranjcevic@hgk.hr](mailto:mcranjcevic@hgk.hr)

## Katolički kalendar u zrcalu Gacke čakavštine

Gacki čakavski govori slabije su istraživani govori, premda su zaštićeno nematerijalno kulturno dobro Republike Hrvatske. Ovim radom nastoji se uputiti na životnost tih govora, inkorporiranih u katolički kalendar, odnosno zadržavanje u kalendaru brojnih fonoloških značajki govora koje se počinju gubiti. Problem je relevantan zato što se malotko time bavi, a kako su govori podložni nestajanju, taj dio jezične baštine u najvećoj je opasnosti.

Metoda istraživanja je terensko istraživanje gackih čakavskih govora s područja Otočca, prije čega je konzultirana sva objavljena dijalektološka građa koja se odnosi na to područje (npr. „Ričnik gacke čakavštine“).

Tipično je za gacku čakavštinu da blagdani muških svetaca uglavnom nose ženska imena (Josipova, Jurjeva, Ilinja...), zapravo je riječ o pridjevima koji su postali imenice e-deklinacije (Josipova ili Kiževa *maša*...). Npersonalizirani blagdani u srednjem rodu imaju ista svojstva (*Spasova, Telova*). Blagdani svetica ne mijenjaju se, u nekim slučajevima gube početni vokal (sveta Lizabeta). Kod nekih se javlja metateza (*Fajbanova* – Fabijanovo; *Lozarija* – Rozalija, Majka Božja od Ružarija), inače rijetka u tim govorima. Jednako tako očuvan je izostanak fonema /h/ (*Dova* – Duhovi, *Mijolja* – Miholje), a kod nekih se javlja i pokrata (*Dova* – Du(h)ova). U nazivlju blagdana mogu se pronaći izgubljeni elementi, poput brojnih pridjeva (*na Blagovisti trojega mesa*). Očituje se i dosljedna primjena Jakubinski-Meyerova pravila (*Svičnica, Telova*), dok je u nekim nazivima sačuvano odstupanje ostvarenja jata po pravilu, koje se u suvremenosti više ne koristi (*Blagovisti – Pripovisti*). Neki od blagdana imaju dvojni naziv (*Svičnica/Kandelora*...), odnosno mogu biti u deminutivu (sveti Lovrenc).

Imena mjeseci ne odstupaju od standardnoga nazivlja (siječanj,...), s tim da se rabi i *ožuljak* (ožujak), dok lipanj ima još naziv – *ivanjski mīsec*. Pravo je obilje pučkih izreka vezanih za blagdane (*sveta Sofija rakiju popila; ognjeni dani*...).

Veliki značaj puk je pridavao i kvatrama (grgurevske, dovske, mijoljske i božićne kvatre) te raznim vjеровanjima i običajima uz njih. Navedene osobine tih govora izdvajaju ih iz (štokavskoga) jezičnog okruženja.

**Ključne riječi:** gacka čakavština, kalendar, crkveni blagdani

Milan Kranjčević, Katedra Čakavskog sabora pokrajine Gacke, Otočac, [mkranjcevic@hgk.hr](mailto:mkranjcevic@hgk.hr)

## Der katholische kalender im spiegel des cakavischen dialekts der region Gacka

Die čakavischen Dialekte der Region Gacka gehören zu den weniger erforschten Dialekten, obwohl sie geschütztes immaterielles Kulturgut der Republik Kroatien darstellen. In diesem Beitrag soll daher versucht werden auf die Lebendigkeit dieser Dialekte hinzuweisen, die im katholischen Kalender enthalten sind, bzw. aufzuzeigen, dass zahlreiche phonologische Merkmale dieser Dialekte, die langsam verloren gehen, darin zu finden sind. Dieses Problem ist insofern relevant, als dass sich selten jemand damit beschäftigt, und da Dialekte dazu tendieren zu verschwinden, befindet sich dieser Teil des Sprachguts in größter Gefahr.

Die Untersuchungsmethode beruht auf einer Feldforschung, die zum Thema der čakavischen Dialekte auf dem Gebiet von Otočac durchgeführt wurde. Vor Beginn wurde die gesamte veröffentlichte dialektologische Fachliteratur, die sich auf dieses Gebiet bezieht (z.B. Ričnik gacke čakavščine) zu Rate gezogen.

Das Typische an dem Čakavischen der Region Gacka ist, dass die Feiertage männlicher Heiliger meist einen weiblichen Namen tragen (Josipova, Jurjeva, Ilinja...), wobei es sich eigentlich um Adjektive handelt, die zu Substantiven mit e-Deklination wurden (Josipova oder Kiževa maša...). Unpersönliche Feiertage im Neutrum weisen dasselbe Merkmal auf (Spasova, Telova). Die Feiertage weiblicher Heiliger verändern sich nicht, verlieren aber in einigen Fällen den Anfangsvokal (sv. Lizabeta). Bei einigen kommt es zur Metathese (Fajbanova – Fabijanovo; Lozarija – Rozalija, MB od Ružarija), die sich sonst selten in diesen Dialekten finden lässt. Ebenso ist in manchen Fällen das Weglassen des Phonems /h/ (Dova – Duhovi, Mijolja - Miholje) erhalten, während es bei einigen Bezeichnungen zur Verkürzung kommt (Dova – Du(h)ova). Bei der Bezeichnung der Feiertage lassen sich verlorene Elemente nachweisen, wie zum Beispiel zahlreiche Adjektive (na Blagovisti trojega mesa). Evident ist auch eine konsequente Anwendung der Jakubinski-Meyer Regel (Svičnica, Telova), während in einigen Bezeichnungen eine Abweichung von der Regel zum Jat-Reflex erhalten bleibt, die heutzutage so nicht mehr angewendet wird (Blagovisti – Pripovisti). Einige der Feiertage tragen eine Doppelbezeichnung (Svičnica / Kandelora...) oder können im Diminutiv (sv. Lovrenc) erscheinen.

Die Monatsbezeichnungen unterscheiden sich nicht von den Standardbezeichnungen (sječanj, ...), wobei darüber hinaus die Bezeichnung ožuljak (ožujak) verwendet wird, während der Juni (lipanj) die zusätzliche Bezeichnung ivanjski misec trägt. In Verbindung mit den Feiertagen kann man überdies eine Fülle von Sprichwörtern finden (sv. Sofija rakiju popila; ognjeni dani ...).

Für das gemeine Volk waren die Quatembren sowie zahlreiche religiöse Überzeugungen und Bräuche, die damit in Verbindung standen, ebenfalls von großer Bedeutung (grgurevske, dovske, mijoljske und božićne kvatre). Durch die angeführten Merkmale grenzen sich diese Dialekte klar von der (štokavischen) Sprachumgebung ab.

Schlüsselwörter: čakavischer Dialekt der Region Gacka, Kalender, kirchliche Feiertage

Jasna Horvat, Udruga Ivana Perkovca, Opća pučka škola, Šenkovec,  
[jasnahorv@yahoo.com](mailto:jasnahorv@yahoo.com)

Sanja Rajter, Udruga Ivana Perkovca, Opća pučka škola, Šenkovec, [sanja.rajter@gmail.com](mailto:sanja.rajter@gmail.com)

## Kajkavski donjosutlanski ikavski dijalekt u projektu Mrežni okvir "Ča-Kaj-Što" sastavnice hrvatskoga identiteta

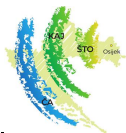
Potporom Zaklade Adris 2020. godine Odjelu za matematiku Sveučilišta u Osijeku, Udruga Ivana Perkovca sudjeluje kao supartner u projektu pod nazivom „Mrežni okvir ča-kaj-što sastavnice hrvatskoga identiteta“. U proteklom vremenu karantene i socijalne izolacije članovi i članice Udruge Ivana Perkovca posvetili su se radu na projektu „Mrežni okvir ča-kaj-što sastavnice hrvatskoga identiteta“ u kojem su supartneri s Odjelom za matematiku Sveučilišta u Osijeku. Projekt financira Zaklada Adris (program: Stvaralaštvo, ekologija, baština i dobrot), a nastavlja se na realizirani projekt iz 2018. godine pod nazivom „Hrvatski jezik u mrežnom oblaku svjetskih jezika“.

Prvim projektom otvorio se novi put modernoga proširenja identiteta hrvatskoga jezika i hrvatske kulturne baštine te je stvoren i provjeren SSF rječnik/tezarij standardnoga jezika s oko milijun pojava te njegova uključenost u lingvistički oblak (LLOD *cloud*) svjetskih jezika. U sklopu projekta „Mrežni okvir ča-kaj-što sastavnice hrvatskoga identiteta“ te se iste riječi proširuju čakavskim, kajkavskim i štokavskim inačicama koje se prilagođavaju u SSF-u. Kajkavcima ikavcima bit će pritom od velike pomoći objavljeni Mrežni rječnik <http://ikavci.ihj.hr/> te Zvučni atlas hrvatskih govora (V. Piškorec; I. Kurtović Budja) <http://hrvatski-zvucni-atlas.com/> u kojem rečenice na ikavskoj kajkavštini kazuju izvorni govornici ikavske kajkavštine.

„Mrežni okvir ča-kaj-što sastavnice hrvatskoga identiteta“ – omogućit će i dopunjavanje korpusa riječi kajkavskoga donjosutlanskog ikavskog dijalekta te će se uz postojećih 11.500 objavljenih riječi dodati, dosad novoprikupljenih osam stotina riječi. Vrijednost je projekta i u dokumentiranju digitaliziranih dokumenata ikavske kajkavštine (izdane knjige, brošure, dokumenti) koji će se također nalaziti u SSF tezariju, a za koje se razvio program smještanja izvora u prostor i vrijeme, poput zvučnoga atlasa.

Taj projekt još jednom je pokazao kako je povezivanje znanosti i nositelja kulturnog dobra – udruge koja djeluje na području svoje lokalne zajednice, presudna u očuvanju i dokumentiranju hrvatske kulturne baštine i osobnoga jezičnog nasljeđa.

**Ključne riječi:** kajkavski donjosutlanski ikavski dijalekt, rječnik, SSF / tezarij standardnog jezika, Udruga Ivana Perkovca, identitet



Jasna Horvat, Udruga Ivana Perkovca, Opća pučka škola, Šenkovec,  
[jasnahorv@yahoo.com](mailto:jasnahorv@yahoo.com)

Sanja Rajter, Udruga Ivana Perkovca, Opća pučka škola, Šenkovec, [sanja.rajter@gmail.com](mailto:sanja.rajter@gmail.com)

## Kajkavian Lower Sutla Ikavian dialect included in the project Network framework "Ča-Kaj-Što" components of Croatian identity

With the support of the Adris Foundation and the University of Osijek's Department of Mathematics, the Ivan Perkovac Association participated in the 2020 project Network framework "Ča-Kaj-Što" components of Croatian identity. The project is funded by the Adris Foundation (Program: Creativity, ecology, heritage and humanity), and is a continuation of an already realized project from 2018 entitled Croatian language in the network cloud of world languages. The first project opened a new path for modern expansion in Croatian language identity and cultural heritage, and resulted in the creation of an SSF dictionary/thesaurus of standard Croatian, containing about 1 million words. The SSF dictionary was sufficiently tested and included in the linguistic cloud (LLOD cloud) of world languages. With the current project Network framework "Ča-Kaj-Što" components of Croatian identity, the SSF dictionary has expanded to include words of the same meaning in the Chakavian, Kajkavian and Shtokavian dialects, which have also been adjusted to SSF standards. The Kajkavian Ikavci benefits from an online dictionary (<http://ikavci.ihj.hr/>) and the Sound Atlas of Croatian Speeches (V. Piškorec; I. Kurtović Budja; <http://hrvatski-zvucni-atlas.com/>) in which sentences in Ikavian Kajkavian are spoken by native speakers. The current project also enables the expansion of the word corpus of the Kajkavian Lower Sutla Ikavian dialect, adding 800 new words to the 11,500 already published words. The project's value lies in the digitalization of Ikavian Kajkavian documents (published books, brochures and various other texts) which will be included in the updated SSF dictionary/thesaurus. This project has shown once again that the connection of science and local community to safeguard cultural heritage (as represented by the Association) is crucial to preserving the Croatian language identity for the future.

**Keywords:** Kajkavian Lower Sutlan Ikavian dialect, Dictionary, SSF, Ivan Perkovac Association, Cultural identity

Matej Vinarić, OŠ „Ivan Goran Kovačić“, Đakovo, [matejblack.vinari@gmail.com](mailto:matejblack.vinari@gmail.com)

## „Bećarac – deseterac s porukom“

Bećarac je oblik napjeva s područja Šokadije (geografska kulturna regija na kojoj žive Hrvati Šokci) koji se sastoji od dviju sastavnica teksta (deseterački dvostisi/trostisi) i pripadajućega melodijskog obrasca. Obje sastavnice promjenjive su i ovise o više elemenata. Melodija ovisi o pjevaču i geografskom području. Melodije su uglavnom slične, imaju jednaku glavnu dionicu, ali imaju određene varijacije u pojedinim elementima. Međutim, naglasak je ovoga izlaganja na tekstu i kontekstu. Tekst bećarca odraz je vremena u kojemu nastaje te se smatra komunikacijskim kanalom između različitih društvenih skupina i/ili pojedinaca. Budući da je bećarac od 2011. godine uvršten na UNESCO-vu Listu nematerijalne kulturne baštine čovječanstva, izvođači bećarca, etnolozi, folkloristi, etnomuzikolozi i dr., dobili su zadatak proučavati i pratiti načine očuvanja bećarca u današnjoj svakodnevici. Dosadašnja istraživanja te glazbeno-folklorne forme temeljila su se na zapisivanju tekstova i djelomičnoj analizi teksta i konteksta. Ovaj rad poseban naglasak stavlja na analizu konteksta bećarca te donosi primjere iz prošlosti i iz novijih vremena iz kojih se može iščitavati načelo aktualizacije. Promišlja se o mogućnosti korištenja bećarca kao povijesnog izvora na trostruki način: analiza vremena u kojemu je nastao, analiza teksta (arhaizmi, novotvorenice i tuđice u navedenoj formi), analiza autora/izvođača (dob, obrazovanje). Navedeno izlaganje doprinosi povijesnoj, lingvističkoj i etnološkoj znanosti.

**Ključne riječi:** bećarac, folklor, tekst, kontekst, povijest

Literatura:

1. Barišić, Ana Tereza. „Književni prikaz Prvog svjetskog rata u kontekstu bećaraca.“ Pannoniana 1, br. 1 (2016)
2. Ceribašić, Naila. „Norma bi individuacija u deseteračkim napjevima s područja Slavonije, 1. dio“ Narodna umjetnost 31, br. 1 (1994): 145-280
3. Dundes, Alen. „Tekstura, tekst i kontekst“ u: Folkloristička čitanka, ur. Marijana Hmerške, Suzana Marjanić, Zagreb: AGM, 2010.
4. Grgurovac, Martin, ur. Slavonske pismice, Vinkovci: SN „Privlačica“, 2005.
5. Toldi, Zvonimir. Sve je cure Mile vezertiro! (pjesmice, bećarci, natpjevanja). Slavonski Brod: Folklorni ansambla Broda, 2013.

Matej Vinarić, OŠ „Ivan Goran Kovačić“, Đakovo, [matejblack.vinari@gmail.com](mailto:matejblack.vinari@gmail.com)

### „Bećarac – ten-part with message“

Bećarac is a music form originating in the area of Šokadija (the geographical cultural region where the Šokci people live) and consists of two components: the text (ten-part) and the corresponding melodic pattern. Both components are variable and depend on multiple elements, with the melody specific to the singer and the geographical area. These melodies are mostly similar, with the same main section, but they have certain variations. A bećarac text is a reflection of the time when it was created, and is considered a communication channel between different social groups and individuals. With bećarac included on the UNESCO list of intangible cultural heritage of mankind since 2011, bećarac performers, ethnologists, folklorists and ethnomusicologists have been given the task of studying and monitoring ways of preserving bećarac in everyday life. Previous research of this musical/folklore form is based on written texts and the partial analysis of text and context. In turn, this paper places special emphasis on the analysis of the bećarac text and context, and utilizes examples from the recent and distant past to read the principle of actualization. The paper also considers the possibility of using bećarac as a historical source in three ways: to analyse the time in which it was created, to analyse the text (archaisms, newcomers and strangers in this form) and to analyse the author/performer (age, education). Accordingly, this presentation contributes to historical, linguistic and ethnological science.

**Keywords:** bećarac, folklore, text, context, history

#### References:

1. Barišić, Ana Tereza. „Književni prikaz Prvog svjetskog rata u kontekstu bećaraca.“ Pannoniana 1, br. 1 (2016)
2. Ceribašić, Naila. „Norma bi individualizacija u deseteračkim napjevima s područja Slavonije, 1. dio“ Narodna umjetnost 31, br. 1 (1994): 145-280
3. Dundes, Alen. „Tekstura, tekst i kontekst“ u: Folkloristička čitanka, ur. Marijana Hmerške, Suzana Marjanić, Zagreb: AGM, 2010.
4. Grgurovac, Martin, ur. Slavonske pismice, Vinkovci: SN „Privlačica“, 2005.
5. Toldi, Zvonimir. Sve je cure Mile vezertiro! (pjesmice, bećarci, natpjevanja). Slavonski Brod: Folklorni ansambla Broda, 2013.



Marko Orešković, Nacionalna i sveučilišna knjižnica u Zagrebu, [moreskovic@nsk.hr](mailto:moreskovic@nsk.hr)

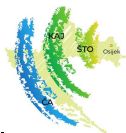
Iva Pribisalić, studentica Odjela za matematiku Sveučilišta u Osijeku, [ipribisa@mathos.hr](mailto:ipribisa@mathos.hr)

Mario Essert, suradnik Odjela za matematiku Sveučilišta u Osijeku, [messert@inet.hr](mailto:messert@inet.hr)

## Tri hrvatska e-leksikona u mrežnome okviru

Riječ, izgovorena ili napisana, proizvod je ljudskoga uma kao obilježje nekoga pojma. Štoviše, i samo je obilježje riječ, a kako jedan pojam obično traži više obilježja, može se reći da riječ opisuju riječi. U izjavi ili rečenici takva struktura dolazi u interakciju s drugim riječima, odnosno s njihovim obilježjima. Pritom sama riječ ima i svoju podstrukturu pa ako gradimo rječnik kao skup riječi, moramo mu dati tri razine informacija: subatomske, atomske i molekularne. Na tom su temelju u mrežnom sintaktičko-semantičkom okviru (SSF) izgrađena tri međusobno čvrsto povezana mrežna rječnika: e-leksikon LEX – koji sadrži oko 950.000 riječi (svih pojavnica), MSY – e-rječnik subatomske dijelova riječi (slogova, morfema i silabomorfema) s oko 10.000 sastavnica te MWE – e-rječnik višerječnica (kolokacija, frazema, termina i sl.) s oko 170.000 primjeraka. Zato ne čudi što računalna baza tog modela zauzima oko 5 GB memorijskoga prostora. Model LEX uključuje polje s naglaskom riječi, može sadržavati sliku koju riječ predstavlja i izgovor (zvuk) riječi s ispisanim naglaskom. Posebnost je tog e-leksikona i njegova vidljiva povezanost sa subatomskom strukturom MSY e-rječnika i s višerječničkom strukturom MWE e-rječnika unutar iste natuknice te s mrežnim znanjem drugih domaćih (Leksikografski zavod „Miroslav Krleža“ – LZMK, Hrvatski jezični portal – HJP, Croatian WordNet – CroWN) i stranih (BabelNet, WordNet) mrežnih repozitorija. Uz to, SSF je kroz CroLLOD u DataHub-u uključen u lingvistički oblak povezanih podataka (LLOD), što znači da je ostvarena mrežna interoperabilnost hrvatskoga jezika s ostalim svjetskim jezicima i narječjima. Te i takve povezanosti imaju izravan učinak na jezikoslovne analize koje SSF nudi, a koje se do sada nisu mogle provoditi u hrvatskom jezikoslovlju.

**Ključne riječi:** mrežni leksikon, tri vrste rječnika, poveznice s resursima, lingvistički mrežni oblak



Marko Orešković, National and University Library in Zagreb, [moreskovic@nsk.hr](mailto:moreskovic@nsk.hr)

Iva Pribisalić, student of the Department of Mathematics, University of Osijek, [ipribisa@mathos.hr](mailto:ipribisa@mathos.hr)

Mario Essert, associate of the Department of Mathematics, University of Osijek, [messert@inet.hr](mailto:messert@inet.hr)

## The Three Croatian E-lexicons in the Network Framework

A word, whether spoken or written, is the product of the human mind, representing a characteristic of a concept. Moreover, the characteristic is also a word and, since a concept usually requires more features, it can be said that words describe a word. When added together in a sentence, words interact with other words, or their features. Each word also has its own substructure. In science, this is similar to the ways in which molecules are made up of atoms, which are, themselves, made up of subatomic particles that come together to create something entirely new. Therefore, if we build a dictionary to contain a set of words, we need to give it three levels of information: subatomic, atomic and molecular. Based on this, the Syntax-Semantic Framework (SSF) network consists of three mutually-linked network dictionaries: LEX, a lexicon which contains over 950,000 words (in all forms); MSY, a lexicon of subatomic word parts (syllables, morphs and syllablemorphs) with over 10,000 parts; and MWE, a lexicon of multiword expressions (collocations, phrases, terms, etc.) with over 170,000 entries. The database containing these lexicons uses 5 GB of storage space. In addition to words, the LEX model also contains each word's accent, image, or sound (spoken variant of the word). The specialty of this lexicon is also its link to other local network resources (e.g. the Miroslav Krleža Institute of Lexicography – LZMK, Croatian Language Portal – HJP, Croatian WordNet – CroWN), as well as foreign ones (e.g. BabelNet and WordNet). Furthermore, the SSF is included in the global Linguistic Linked Open Data (LLOD) cloud, which provides interoperability with all other world languages and dialects. These links produce a direct impact on language analysis, enabling the SSF to conduct what was, until now, not possible to conduct.

**Keywords:** network lexicon, three kind of dictionary, links to resources, linguistic network cloud

Suzana Molčanov, umir. XI gimnazije u Zagrebu, [suzana.molcanov@gmail.com](mailto:suzana.molcanov@gmail.com)

Marko Landeka, umir. Državni arhiv Vukovar/Vinkovci, [marko.landeka@gmail.com](mailto:marko.landeka@gmail.com)

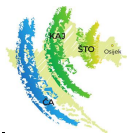
Blaženka Budimčić, Ekonomska i upravna škola u Osijeku, [blbudimcic@gmail.com](mailto:blbudimcic@gmail.com)

## Uređivanje mrežnih e-rječnika u sintaktičko-semantičkom okviru

Prirodni je jezik dinamičan i živ sustav u kojem svakodnevno nastaju nove riječi. Dok u jezik ulaze nove strane riječi, neke se uopće više ne koriste, a neke poprimaju i nova značenja. Zato je uređivanje mrežnih rječnika prijeko potreban i zahtjevan zadatak. Budući da se radi o desetcima i stotinama tisuća riječi jednog jezika, nužno je uređivanje povjeriti grupi stručnjaka. Pri izradi rječnika svih pojava nameću se sljedeća pitanja: hoće li uredništvo ručno unositi svaku od pojava i koliko je to vremenski zahtjevan posao? Osim problematike vremena, potrebno je riješiti i problematiku različitih stavova struke te problem kontrole i sprječavanja ljudskih pogrešaka.

U sintaktičko-semantičkom okviru (SSF) koji sadrži tri vrste rječnika načinjena su tri modula koja maksimalno objektiviziraju uređivački posao. Modul unosa dohvaća pojedinačne riječi koje nakon raščlambe (parsiranja) dokumenta nekog korpusa nisu pronađene među dotadašnjim riječima u rječniku. Ako se radi o promjenjivoj vrsti riječi, urednik je pretvara u kanonski oblik (lemu imenice, pridjeva ili glagola) te poziva morfološki generator (J. Markučića) koji po zadanim morfonološkim zakonima generira odgovarajuće sintagme. Urednikova je zadaća samo odabrati pravu riječ i odabrati naredbu 'Spremi'. Sve pojavnice novoodabrane riječi dobit će sva pripadna gramatička obilježja (WOS) po svim kategorijama. Za nepromjenjivu riječ dovoljno je odabrati jednu od ponuđenih opcija. Modul za promjenu riječi nudi vizualni prikaz trenutnog stanja raščlambe riječi i njezinih gramatičko-semantičkih obilježja. Odabirom pojedinog prikaza omogućuje se promjena (npr. naglasak na vokalu; razdjelnika među slovima, npr. morfema ili slogova; umetanje ništičnog morfema i sl.). Urednik uvijek ima samo mogućnost izbora, što smanjuje mogućnosti pogrešaka i povećava sustavnost, ujednačenost i dosljednost. Neispravne riječi, koje su se pojavile kao nepostojeće u jeziku, urednik izbacuje odabirom naredbe 'Odbaci', kao što ispravnu riječ potvrđuje odabirom naredbe 'Unesi'. Sličan pristup postoji i u uređivanju rječnika višerječnica.

**Ključne riječi:** uređivanje mrežnih leksikona, vizualizacija, naglasci i raščlambe, morfološki generator



Suzana Molčanov, XI Gymnazije u Zagrebu, retired, [suzana.molcanov@gmail.com](mailto:suzana.molcanov@gmail.com)

Marko Landeka, ret. State archives in Vukovar/Vinkovci, [marko.landeka@gmail.com](mailto:marko.landeka@gmail.com)

Blaženka Budimčić, School of Economics and Management in Osijek,  
[blbudimcic@gmail.com](mailto:blbudimcic@gmail.com)

## Editing Online E-dictionaries in the Syntactic-Semantic Framework

Natural language is a dynamic system in which new words emerge on a daily basis. Some become meaningful (e.g. the word "net" once referring to fisheries and nowadays to computations), while others are not used at all due to the influence of foreign languages and other factors. This is why editing online dictionaries is a necessary though not always simple task. Since we handle thousands of a language's words in this process, it is necessary to trust a group of editing professionals, and not just one person. Further, when editing a lexicon of all word forms, must the editing group manually enter each one into the edited dictionary? Along with time, the problem of human differences emerges, so how can we avoid this? In the Syntactic and Semantic Framework (SSF), which contains three types of lexicons, three different modules maximize editing objectivity. The insertion module in particular extracts all words which the parser did not find in the lexicon, while for open word types, the editor changes them to the canonical form (lemma) by using a morphological generator (by J. Markučić). The editor then only needs to choose the correct form and click "Save". All word forms will automatically be saved in the lexicon with all related grammatical tags. For unchangeable words, it is enough to select one. The SSF's word editing module also offers a visual user interface for the word breakdown, accent editing or insertion of null morphemes; the editor only needs to choose the necessary action to minimize errors. The editor can then delete words from the lexicon that are not valid by clicking "Discard"; similarly, correct words can be accepted by clicking "Accept". The same approach is also available in the multiword expression editing module.

**Keywords:** editing network lexicons, visualizations, accents and disassembly, morphological generator

Domagoj Ševerdija, Odjel za matematiku Sveučilišta u Osijeku, [dseverdi@mathos.hr](mailto:dseverdi@mathos.hr)

Ivan Židov, student Odjela za matematiku Sveučilišta u Osijeku, [izidov@mathos.hr](mailto:izidov@mathos.hr)

## Obrada lingvističkih podataka koristeći API SSF-a

Mrežni okvir za obradu lingvističkih podataka imenom Sintaktičko-semantički okvir (SSF) koji, osim posjedovanja triju rječnika morfema, riječi i višerječnica, nudi programerski modul u kojem korisnik može programski obrađivati vlastite korpuse koristeći informacije dostupne u SSF rječnicima i sintaktičko-semantičke informacije kroz T-strukture i O-strukture.

U ovoj radionici pokazat će se korištenje SSF-a u kontekstu programerskog sučelja za obradu lingvističkih podataka kroz JupyterHub sučelje. Demonstrirat će se kako dohvatiti i pripremiti tekstove za sintaktičku i semantičku analizu koristeći NLTK module za oblikovanje korpusa. Na praktičnim primjerima demonstrirat će se trenutačno implementirane API mogućnosti SSF-a kao što je lematizacija riječi, određivanje vrste riječi i imenskih fraza, pronalaženje metonimije i metafore itd. Pored implementiranih mogućnosti, SSF uključuje popularne moderne biblioteke kao što je Spacy za obradu prirodnog jezika, posebno izvježban za hrvatski jezik. Te biblioteke mogu se koristiti za nadopunu i predikciju određenih značajki unutar leksikona. U konačnici, predstaviti će se Word2Vec neuronski model jezika koji projicira entitete iz leksikona u vektorski prostor. Ideja je tog modela za jedinstvenu riječ pronaći reprezentativni vektor koji opisuje interakciju te riječi s ostalim riječima u nekom kontekstu (princip distribuirane semantike). Model izgrađuje vektore riječi koje su izvježbane na korisnički-definiranom korpusu. Vektori riječi mogu modelirati semantička svojstva kao što su analogije i sličnosti, što će se i pokazati na vlastito-definiranom korpusu.

Za polaznike radionice poželjno je osnovno poznavanje Pythona, ali nije obvezujuće.

**Ključne riječi:** sintaktička i semantička analiza, SSF, API, paket NLTK



Domagoj Ševerdija, Department of Mathematics, University of Osijek, [dseverdi@mathos.hr](mailto:dseverdi@mathos.hr)

Ivan Židov, student of the Department of Mathematics, University of Osijek,  
[izidov@mathos.hr](mailto:izidov@mathos.hr)

## Processing of linguistic data using SSF API

The linguistic data processing Syntactic-Semantic framework (SSF), in addition to having three lexicons of morphemes, words and multiword expressions, offers a programming module through which users can process their own corpora with data from SSF lexicons and syntactic-semantic information via T-structures and O-structures, respectively.

The goal of this workshop is to present an SSF application programming interface (API) for linguistic data processing within the JupyterHub service. It demonstrates the retrieval and preparation of texts for syntactic and semantic analysis, as well as shows the current state of API functionality through practical examples, such as word lemmatization, word tagging, named entity recognition, and metonymy and metaphor recognition. Moreover, the SSF incorporates state-of-the art libraries, such as Spacy, for natural language processing specifically trained for the Croatian language. Such libraries can be used to add or predict features derived from the SSF. Finally, this workshop presents a Word2Vec neural language model which projects lexicon items into a vector space. This model's main principle is to find a representative vector of a word which captures its interaction with other words within a specific context (i.e. the distributed semantics principle). The model also builds word vectors that are trained on a user-defined corpus. Lastly, this workshop demonstrates how word vectors can model semantic properties like analogies and similarities on an extant user-defined corpus.

**Keywords:** syntactic and semantic analysis, SSF API, NLTK package

Josip Užarević, Filozofski fakultet Sveučilišta u Zagrebu, [juzarevic@ffzg.hr](mailto:juzarevic@ffzg.hr)

Suzana Molčanov, umir. XI gimnazije u Zagrebu, [suzana.molcanov@gmail.com](mailto:suzana.molcanov@gmail.com)

Antonio Jovanović, student Odjela za matematiku Sveučilišta u Osijeku,  
[ajovanov@mathos.hr](mailto:ajovanov@mathos.hr)

## Hrvatski mrežni subčestični (strukturni) rječnik

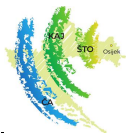
Riječi bi se mogle promatrati kao atomi koji tvore jezik, a njihovi dijelovi (morfemi i slogovi) kao subčestice koje tvore riječi.

Morfemi su bitni u oblikoslovnoj tvorbi riječi, bez obzira na to radi li se o sklonidbi (fleksiji) ili izvođenju (derivaciji), a slogovi su važni u glasovnom smislu. Uz to postoje i razlozi za izdvajanje posebne međurazine koja bi spajala morfeme i slogove u tzv. silabomorfeme – radi boljšega razumijevanja naglasnih zakonitosti u hrvatskome jeziku. Mrežni sintaktičko-semantički okvir (SSF) podržava sve subčestice, ali ne pokriva potpuno i njihovo automatsko formiranje. Naime, automatsko rastavljanje riječi u potpunosti je dovršeno samo za slogove, a tek dijelom za morfove. Međutim, zahvaljujući dobro razvijenoj morfonologiji hrvatskoga jezika (Marković, 2013.), algoritmi se ubrzano razrađuju i za tu vrstu raščlambe. Tako se uz e-leksikon (pojedinačnih riječi – LEX), stvara i mrežni subčestični rječnik (MSY).

Pretraživanje MSY-ja slično je pretraživanju LEX-a u SS okviru, samo su umjesto abecednog poretka riječi, u ovom slučaju abecedno poredani slogovi/morfovi. Na vrhu MSY-jeva zaslona za postavljanje uvjeta pretrage napisan je trenutačni broj morfema (2.153) i slogova (9.783) koji su spremljeni u bazi. Riječi i njihove subčestice prikazane su u odnosu na poziciju subčestice u pripadajućoj riječi (prema broju utora). Ako je korisnik izabrao prvu poziciju (utor broj 1) na stranici postavki, onda će se kao rezultat za odabrano slovo prikazati svi morfovi/slogovi koji počinju tim slovom, a nalaze se u prvom utoru riječi, zajedno s riječima koje imaju taj slog na tom mjestu. Na stranici postavki moguće je odabrati broj od 1 do 11 za apsolutnu poziciju utora jer najduža hrvatska riječ ima 11 slogova. Također, postoji izbor relativne pozicije: 'bilo koji' i 'zadnji' slog/morf, što je neovisno o broju subčestica u riječi.

Takva metoda pretraživanja subčestica i njihove okolice donosi vrijednost poput one koju je konkordancija donijela sintaktičkoj analizi.

**Ključne riječi:** slogovi i morfovi, silabomorfemi, strukturni rječnik, utori



Josip Užarević, Faculty of Humanities and Social Sciences University of Zagreb,  
[juzarevic@ffzg.hr](mailto:juzarevic@ffzg.hr)

Suzana Molčanov, XI Gymnazije u Zagrebu, retired, [suzana.molcanov@gmail.com](mailto:suzana.molcanov@gmail.com)

Antonio Jovanović, student, Department of Mathematics, University of Osijek,  
[ajovanov@mathos.hr](mailto:ajovanov@mathos.hr)

## Croatian Network Subparticle (Structural) Dictionary

If words are like atoms which form a certain language, then parts of a word (morphs and syllables) are like the sub-particles that form those atoms. Morphemes are relevant in morphological formation, regardless of flexion or derivation, and syllables are relevant in a phonetic sense. There are serious arguments for monitoring the relationships between morphs and syllables, through so-called syllabomorphemes. The Syntactic and Semantic Framework (SSF) supports all sub-particles by implementing algorithms to automatically decompose words into syllables and, to a lesser extent, morphs. Due to the well-documented morphology of the Croatian language, further research to develop algorithms that degrade words into both morphs and syllables is expected. The process of searching the SSF's sub-particle dictionary (MSY) is similar to that used to search the classical lexicon (LEX). The only difference is that, rather than an alphabetical list of words, the SSF displays a list of syllables and/or morphs. As with the classical lexicon, this sub-particle dictionary includes a setup screen which can be used to configure how the output appears. The total number of morphs (2,153) and syllables (9,783) currently stored in the database is displayed at the top of the screen, while the sub-particles and words which use them is displayed in relation to the position of the sub-particle within the word (related to the slot's number). For instance, if position one is selected in the setup screen, the output will be listed in ascending order, starting with the first sub-particle. Since the longest word in the SSF lexicon consists of 11 syllables, that is the highest position setting in the setup screen. The value of this search technique is similar to the value that Concordance provided to syntactical research.

**Keywords:** syllables, morphemes, syllabomorphemes, structural dictionary, slots



Mario Essert, suradnik Odjela za matematiku Sveučilišta u Osijeku, [messert@inet.hr](mailto:messert@inet.hr)

Lucija Tatarević, studentica Odjela za matematiku Sveučilišta u Osijeku,  
[lucija.tatarevic@gmail.com](mailto:lucija.tatarevic@gmail.com)

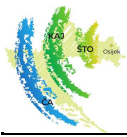
Marko Orešković, Nacionalna i sveučilišna knjižnica u Zagrebu, [moreskovic@nsk.hr](mailto:moreskovic@nsk.hr)

## Hrvatski mrežni e-rječnik višerječnica

Kao što su molekule tvorene od atoma, tako su višerječnice tvorene od riječi. One obično sadrže različit tip informacije: terminološki (npr. ratni brod), sociološki (npr. labava carinska unija), metaforički (npr. sunce moje drago), idiomatski/frazeološki (npr. brz na jeziku), poslovični (npr. griješeći se uči) i sl. Za svaki je jezik rječnik višerječnica (MWE – eng. *multiword expression*) iznimno važan jer daje paradigmatsku i semantičku komponentu koju rječnik pojedinačnih riječi nema. Mrežni sintaktičko-semantički okvir (SSF), uz mrežni leksikon i rječnik subčestica (morfemi i slogovi), sadrži i MWE rječnik koji je s njima čvrsto (u oba smjera) povezan.

Trenutačno u SSF-u postoji 170.000 višerječnica koje su programom skupljene iz vanjskih izvora (poštujući autorsko pravo) ili mrežnih dokumenata (uz statističku provjeru tipa, tj. veze riječi u višerječnici). Uporaba i korisničko sučelje MWE rječnika slično je načinu rada ostalih rječnika. Poredak višerječnica temelji se na abecednom poretku riječi koje su povezane i mjestom, tj. pozicijom odabrane riječi u višerječnici. Mogu se odabrati apsolutna mjesta utora odabrane riječi (od 1 do 11) ili relativna mjesta (npr. korisnik može odabrati da tražena riječ bude na kraju višerječnice ili na bilo kojoj poziciji). Uz to postoji filtracija višerječnica pa se mogu prikazivati samo izabrane kategorije, npr. kolokacije, idiomi/frazemi, čvrste sveze riječi, termini i sl. Uz svaku višerječnicu prikazana je i potpuna struktura svake pojedinačne riječi u njoj – njezina morfemska, slogovna, gramatička i semantička obilježja te poveznice s drugim izvorima. Posebna pozornost posvećena je statističkim alatima za obradu višerječnica (npr. mjerama asocijativne povezanosti). Ti alati (Studentov t-test, Pearsonov hi-kvadrat test, Log-likelihood i Fisherov egzaktni test) služe za otkrivanje te provjeru kolokacija i koligacija. Dodatni naponi uloženi su u aglomerativno grupiranje kako bi se provela kolostrukcijska (Stefanowitsch, Gries) istraživanja i u hrvatskome jeziku.

**Ključne riječi:** mrežni okvir, rječnik višerječnica, provjera kolokacija i koligacija, kolostrukcijska istraživanja



Mario Essert, associate of the Department of Mathematics, University of Osijek,  
[messert@inet.hr](mailto:messert@inet.hr)

Lucija Tatarević, student, Dept. of Mathematics, University of Osijek,  
[lucija.tatarevic@gmail.com](mailto:lucija.tatarevic@gmail.com)

Marko Orešković, National and University Library in Zagreb, [moreskovic@nsk.hr](mailto:moreskovic@nsk.hr)

## Croatian network e-dictionary of multiword expressions

Similar to how molecules are formed from atoms, multiword expressions are formed from words and usually hold different information types (e.g. terminological: "war ship"; sociological: "loose customs union"; metaphorical: "my sweet sunshine"; phrasematic; proverbial; etc.). Having an MWE lexicon is important for a language, because it offers paradigmatic and semantic components that a simple dictionary does not contain. The Syntactic and Semantic Framework (SSF) includes one such "molecular lexicon", which is similar to already described lexicons of words and syllables/morphs. Currently there are over 170,000 multiword expressions automatically collected from external resources (respecting copyrights) and network documents (with statistical type verification, i.e. type of bonding within the MWE). The SSF's display settings are also similar to other lexicons. By default it organizes multiwords in ascending order based on the position of the words within the multiword, but this can be further configured in the setup screen (e.g. to show only collocations, phrasemes, etc.). Near each multiword is a list of words (atoms) from which the multiword is formed. Special attention is given to statistical tools for MWE processing (e.g. measures for collocational and collostructional strength). These tools (Student's t-test, Pearson's chi-square test, loglikelihood and the Fisher exact test) are used for collocation detection and verification. Additional efforts are made for agglomerative clustering to conduct collostruction research (Stefanowitsch, Gries) in Croatian.

**Keywords:** network framework, multiword lexicon, collocation detection, collostruction research



Autori	Stranica
Blažeka, Đuro	<a href="#">3</a>
Blažević Krezić, Vera	<a href="#">19</a>
Budimčić, Blaženka	<a href="#">43</a>
Ćubelić, Marija	<a href="#">25, 29</a>
Ćurak, Silvija	<a href="#">19</a>
Essert, Mario	<a href="#">41, 49</a>
Filipić, Lobel	<a href="#">25, 29</a>
Goleš Glasnović, Biserka	<a href="#">21</a>
Guberac, Vlado	<a href="#">VIII</a>
Horvat, Jasna	<a href="#">37</a>
Jovanović, Antonio	<a href="#">47</a>
Jozić, Željko	<a href="#">7</a>
Kranjčević, Milan	<a href="#">35</a>
Kurtović Budja, Ivana	<a href="#">15</a>
Landeka, Marko	<a href="#">43</a>
Mance, Ivan	<a href="#">11</a>
Molčanov, Suzana	<a href="#">43, 47</a>
Orešković, Marko	<a href="#">41, 49</a>
Pribisalić, Iva	<a href="#">41</a>
Rajter, Sanja	<a href="#">37</a>
Sammartino, Antonio	<a href="#">17</a>
Stolac, Diana	<a href="#">9</a>
Ševerdija, Domagoj	<a href="#">45</a>
Štambuk, Drago	<a href="#">IX</a>
Štimac Ljubas, Vlatka	<a href="#">33</a>
Tatarević, Lucija	<a href="#">49</a>
Užarević, Josip	<a href="#">47</a>
Vignjević, Jelena	<a href="#">13</a>
Vinarić, Matej	<a href="#">39</a>
Zubčić, Sanja	<a href="#">1</a>
Židov, Ivan	<a href="#">45</a>
Žubrinić, Darko	<a href="#">11</a>