



Financira
Europska unija
NextGenerationEU



Institut za hrvatski jezik
Hrvatski mrežni rječnik – Mrežnik: 2. faza

Okrugli stol
Riječ po riječ, rječnik po rječnik...

KNJIŽICA SAŽETAKA

Zagreb, 27. studenoga 2025.

Kristina Štrkalj Despot

Institut za hrvatski jezik

kdespot@ihjj.hr

Leksikografija u vremenu umjetne inteligencije: trendovi, projekti i infrastrukture

Leksikografija i dalje prolazi kroz duboku transformaciju obilježenu digitalnom tranzicijom, razvojem istraživačkih infrastrukture, integracijom velikih korpusa, sve boljim alatima za pretraživanje i analizu te primjenom velikih jezičnih modela u leksikografiji. Iz anketa provedenih u proteklih nekoliko godina, koje obuhvaćaju leksikografske prakse velikoga broja zemalja (Kosem i dr. 2019.; Tiberius i dr. 2024.) može se jasno vidjeti da, unatoč golemu tehnološkomu napretku, velik broj leksikografa i dalje ima potrebu za boljim alatima za izdvajanje i automatsku obradu podataka iz korpusa (osobito za semantičku analizu: značenja, definicije i leksičke veze) te za boljim povezivanjem sustava za pretragu korpusa (*CQS*) i sustava za pisanje rječnika (*DWS*) te se naglašava potreba za znatnim poboljšanjem (polu)automatske izrade i posteditorske leksikografije jer je trenutačno kvaliteta i pouzdanost takvih resursa upitna. Veliki izazovi koji se ističu jesu održavanje, nadogradnja i upravljanje podatcima te dakako nedostatak financiranja i nedostatan vrednovanje leksikografskoga rada, koji se očituju u ukidanju velikih nacionalnih projekata (npr. *Veliki poljski rječnik* ili njemački *Elexiko*). Pozitivna promjena koju je donijela digitalna revolucija, a koja je u anketama istaknuta, jest bolja interakcija s korisnicima (iako postoji potreba za boljim alatima za analizu korisničkoga ponašanja). Ankete su, očekivano, pokazale da je leksikografska zajednica vrlo heterogena: ono što neki vide kao napredak, drugi smatraju gubitkom. Među novim trendovima ističu se *crowdsourcing* i igrifikacija, koji se zasad većinom primjenjuju u obliku korisničkih povratnih informacija. U nekim su se projektima (poput slovenskoga tezaurusa, slovenskoga kolokacijskoga rječnika, estonskoga rječnika asocijacija ili nizozemskog *Taalradara*) te metode pokazale visokoučinkovitima. No najveći je i najaktualniji izazov integracija umjetne inteligencije i velikih jezičnih modela u leksikografiju (De Schryver 2023.; Lew 2024.), što se vidi i po popularnosti i zastupljenosti znanstvenih radova u časopisu *International Journal of Lexicography* te po temama konferencijskih izlaganja na prošloj EURALEX konferenciji te po temi sljedeće konferencije. Dosadašnji radovi pokazuju da su ljudska stručnost i evaluacija zasad nezamjenjivi u osiguravanju točnosti i kulturne relevantnosti rječničkih zapisa. Rezultati infrastrukturnoga leksikografskog projekta ELEXIS i COST akcije *EneL* pokazali su i da

postoji potreba za zajedničkim standardima i alatima za usklađivanje formata rječnika. Stoga su leksikografske mreže suradnje važnije nego ikad. EURALEX kao najvažnija takva mreža redovito održava svoj međunarodni kongres, na svojoj mrežnoj stranici nudi korisne informacije, snimke izlaganja i pristup leksikografskim radovima te putem svoje *EURALEX Talks* serije omogućuje umrežavanje i razmjenu ideja i u razdoblju između dviju konferencija. Na globalnoj razini takve platforme jesu: Globalex, DSNA, AFRILEX, ASIALEX i AustraLex. Međunarodni studij *European Master in Lexicography* (EMLex) nudi međusveučilišnu i višejezičnu obuku u leksikografiji. eLex konferencije usmjerene su inovativnim trendovima u leksikografiji. Kao i drugdje u Europi, i hrvatska leksikografija prati digitalni zaokret, ponajviše projektom *Mrežnik* (Hudeček i Mihaljević 2017.), koji se razvija u Institutu za hrvatski jezik kao prvi digitalno rođen jednojezični mrežni rječnik hrvatskoga standardnog jezika.

de Schryver, G.-M. 2023. Generative AI and lexicography: The current state of the art using ChatGPT. *International Journal of Lexicography* 36/4. 355–387.

Hudeček, L.; Mihaljević, M. 2017. The Croatian Web Dictionary Project – Mrežnik. *Electronic lexicography in the 21st century. Proceedings of eLex 2017 conference*. Ur. Kosem, Iztok i dr. Lexical Computing CZ s.r.o. Brno – Leiden. 172–192.

Kosem, I.; Lew, R.; Müller-Spitzer, C.; Ribeiro Silveira, M. i dr. 2019. The image of the monolingual dictionary across Europe. Results of the European survey of dictionary use and culture. *International Journal of Lexicography* 32/1. 92–114. <https://doi.org/10.1093/ijl/ecy022>.

Lew, R. 2024. Dictionaries and lexicography in the AI era. *Humanities and Social Sciences Communications* 11/426. <https://doi.org/10.1057/s41599-024-02889-7>.

Tiberius, C.; Kallas, J.; Koeva, S.; Langemets, M.; Kosem, I. 2024. A Lexicographic Practice Map of Europe. *International Journal of Lexicography* 37/1. 1–28. <https://doi.org/10.1093/ijl/ecad023>.

Nina Ledinek

ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša

nina.ledinek@zrc-sazu.si

Pomen projekta *eSSKJ in korpus: na poti k najsodobnejšim jezikovnim podatkom pri oblikovanju slovarja eSSKJ*

V prispevku predstavljamo vlogo projekta *eSSKJ in korpus: na poti k najsodobnejšim jezikovnim podatkom* pri oblikovanju najnovejšega enojezičnega splošnega razlagalnega slovarja za slovenščino *eSSKJ: Slovarja slovenskega knjižnega jezika, tretja izdaja*.

Infrastrukturni projekt *eSSKJ in korpus: na poti k najsodobnejšim jezikovnim podatkom* smo med marcem 2023 in oktobrom 2023 izvajali na Inštitutu za slovenski jezik Frana Ramovša ZRC SAZU. Cilja projekta sta bila:

- prenoviti format zapisa podatkov o tipičnih kolokatorjih in skladenjskih vzorcih leksemov v slovarski podatkovni bazi eSSKJ, hkrati pa kolokatorje, ki so kot iztočnice že opisani v slovarski bazi, tudi pomensko razdvoumiti;
- oblikovati nov specializirani pisni korpus slovenščine, *Korpus znanstvenih besedil sodobne slovenščine*.

Motivacija za delo na obeh gradnikih slovenske jezikovne infrastrukture je bila vsebinska in tehnološka nadgradnja (podatkovne baze) eSSKJ. Posodobitev formata zapisa podatkov o tipičnih kolokatorjih in skladenjskih vzorcih leksemov pri več kot 30.000 besedah v več kot 12.000 kolokacijskih nizih v slovarski podatkovni bazi eSSKJ je bila ključna, saj omogoča povezljivost in izmenljivost podatkov z različnimi drugimi jezikovnimi viri, zlasti leksikalnopodatkovnimi in korpusnimi, kar pomembno prispeva k učinkovitejši in hitrejši izdelavi eSSKJ ter k učinkovitejšemu posodabljanju njegovih podatkov. Za opisano tehnološko nadgradnjo eSSKJ smo se odločili, saj je strokovno verodostojna interpretacija podatkov o tipičnem besedilnem okolju leksemov časovno eden najbolj zahtevnih delov oblikovanja temeljnih enojezičnih splošnih razlagalnih slovarjev, vendar pa je ta segment opisa zaradi potreb uporabnikov, ki se vedno pogosteje in bolj demokratično vključujejo v jezikovni diskurz, relevantnejši kot kadarkoli doslej.

Za oblikovanje geslovnika in leksikalnega opisa eSSKJ je bila pomembna tudi izgradnja *Korpusa znanstvenih besedil sodobne slovenščine*. Delež stvarnih besedil v slovenskih referenčnih korpusih, ki so temeljna podlaga za oblikovanje eSSKJ, je namreč skromen, zato ti korpusni viri ne odsevajo realne pomenske razplatenosti slovenskega knjižnega jezika,

omenjeno stanje pa negativno vpliva na leksikalni opis, in sicer tako determinologizirane leksike kot splošne leksike, ki se tipično rabi v strokovnih besedilih. Korpus Gigafida 2.0, ki kot najodobnejši referenčni korpus služi kot temeljna gradivska osnova za slovar eSSKJ, vključuje le 3,8 % stvarnih besedil, pri čemer je bilo le 5 % teh besedil objavljenih po letu 2010. Za slovenščino sicer obstajajo tudi obsežnejši specializirani korpusi strokovnih besedil, npr. korpus akademske slovenščine KAS in korpus znanstvenih objav s portala Openscience.si OSS, ki referenčni korpus gradivsko dopolnjujejo, vendar pa vključujejo zlasti zaključna dela z različnih nivojev visokošolskega izobraževanja, delež besedil osrednjih strokovnih besedilnih vrst v njih pa je majhen. Težava je tudi, da se ne posodablja (redno). Nastali *Korpus znanstvenih besedil sodobne slovenščine* se od omenjenih korpusov nekoliko razlikuje, s čimer se korpusni infrastrukturni manko slovenščine zmanjšuje, z interpretacijo podatkov v korpusu pa se povečuje tudi kvaliteta leksikalnega opisa sodobne slovenščine v najnovejšem enojezičnem splošnem razlagalnem slovarju.

Anja Nikolić-Hoyt

Zavod za lingvistička istraživanja, Hrvatska akademija znanosti i umjetnosti

anhoyt@hotmail.com

Benešićev književnojezični rječnik od spomenika hrvatskoga povijesnog leksika do njegove suvremene obrade

Benešićev *Rječnik hrvatskoga književnoga jezika od preporoda do I. G. Kovačića* jedinstven je društvenopovijesni dokument i jezičnokulturni spomenik hrvatskoga jezika i književnosti od sredine XIX. do polovine XX. stoljeća, to jest od preporoda i kodifikacije ilirskoga, odnosno općehrvatskoga jezičnog standarda do kodifikacijskoga zaokreta prema artikuliranju i pobjedi jezičnoga unitarizma, s kratkim periodom reafirmacije i restauracije posebnosti hrvatskoga jezika u vrijeme NDH. Premda je riječ o mlađoj prošlosti, stvarnost prožeta previranjima i smjenom različitih paradigmi koju *Rječnik* odražava pripada prošlosti, kao što je i leksičko blago koje sadržava iz perspektive suvremenih jezičnih odnosa znatnim dijelom povijesno i prošlo.

Nastao sredinom prošloga stoljeća iz potrebe za izdavanjem rječnika koji bi nadomjestio praznine i nedostatke tada postojećih rječnika, poglavito *Broz-Ivekovićeve rječnika* iz 1901. u kojemu hrvatska jezična i književna tradicija nisu bile u dovoljnoj mjeri zastupljene, *Benešićev*

se *rječnik* temelji na citatima iz djela hrvatske književnosti kao repozitorija identitetskih obilježja hrvatskoga jezika. Za potrebe *Rječnika* ekscerpirano je više od šesto književnih djela romantizma, realizma i moderne te više od stotinu hrvatskih književnika od Antuna Mihanovića do I. G. Kovačića, koji su stvarali u burnome vremenu usustavljanja i ustaljivanja hrvatskoga standardnog jezika, vremenu u kojemu supostoje različite grafije, pravopisi i gramatike kao odrazi različitih jezičnopolitičkih praksi. Svaka je natuknica u *Benešićevu rječniku* oprimjerena nizom kronološki poredanih književnih citata koji u funkciji kontekstnih okvira pridonose opisu, odnosno razrađenijemu profiliranju značenja i proširenju znanja o denotatu.

Prvih dvanaest svezaka (*a – rzati*) objavila je Jugoslavenska akademija znanosti i umjetnosti zajedno s Nakladnim zavodom Globus od 1985. do 1990., tridesetak godina nakon Benešićeve smrti, a rad na dovršavanju *Rječnika* nastavio se tek dva desetljeća kasnije, to jest 2008. u Zavodu za lingvistička istraživanja Hrvatske akademije znanosti i umjetnosti. Otad su objavljena tri nova sveska: 13. *s – spužvast* (2013.), 14. *spužvica – švrljuga* (2017.) i 15. *ta – uljudba* (2022.), pred zaključenjem je i pretposljednji, 16. *uljuditi – vrebuške*. Oni se odlikuju opsežnom metodološkom obnovom i suvremenim pristupom obradi. Postupci i procedure metodološke obnove novih svezaka, o kojima će se više govoriti u samome izlaganju, bitno su utjecali na prerastanje *Benešićeva rječnika* u iznimno vrijednu sastavnicu hrvatskoga nacionalnog i kulturnog bića, čije dimenzije i dubine još uvijek otkrivamo.

Jozo Vela

Staroslavenski institut

jvela@stin.hr

Rječnik crkvenoslavenskoga jezika hrvatske redakcije – RCJHR

Izrada *Rječnika crkvenoslavenskoga jezika hrvatske redakcije* (RCJHR) temeljni je projekt Staroslavenskoga instituta od sredine prošloga stoljeća. Njegova povijest seže u 1958. kad je na Četvrtome međunarodnom slavističkom kongresu u Moskvi nastala ideja za stvaranje rječnika svih redakcija crkvenoslavenskoga jezika te se tada pristupilo prikupljanju građe i izradi korpusa za njegovo sastavljanje. Na RCJHR su radili svi (istaknuti) znanstvenici Staroslavenskog instituta, osobito do 2000-ih godina. Prvi je svešćić objavljen 1991. godine. Dosad su objavljena 23 svešćića (*a – izgnilstvo*), od kojih su po deset uvezani u velike sveske.

Prvi je takav svezak objavljen 2002., a drugi 2016. Dosad je RCJHR bio objavljivao isključivo tiskom. Posljednji je svešćić tiskan prije četiri godine (*i – izgnilstvo*).

Riječ je o rječniku koji se može odrediti kao rječnik razdoblja, tj. povijesni rječnik. Rječnik sadržava četiri pisma – glagoljicu, ćirilicu, latinicu i grčki alfabet, te pet jezika – hrvatski crkvenoslavenski, hrvatski, engleski, starogrčki i latinski. U rječničkome članku, u kojemu se natuknica navodi na glagoljicu i ćirilicu, donosi se niz znanstveno relevantnih podataka. Navode se istovrijednice na hrvatskome, engleskome, latinskome i starogrčkome, gramatičke odrednice, primjeri uporabe s latinskim i/ili latinskim usporednim tekstom. O riječima se donose ortografski, sintaktički, morfološki, semantički, stilistički, distribucijski, uporabni, ilustrativni i međujezični podatci. Korpus za RCJHR temelji se na 66 osnovnih glagoljičnih izvora.

U izlaganju će se predstaviti trenutačno stanje leksikografske obrade, njezine mogućnosti i daljnji planovi. Na odjelu Staroslavenskoga instituta koji je zadužen za RCJHR i korpusnu građu RCJHR trenutačno djeluju dva projekta financirana sredstvima Europske unije (NextGenerationEU), a to su RCJHR i DigiSTIN (*Razvoj modela digitalne infrastrukture Staroslavenskoga instituta*). Dok prvi projekt ima za cilj nastaviti »tradicionalnu« leksikografsku obradu hrv. crkvenoslavenskoga jezika, ali ju pretočiti u suvremenije sučelje digitalne e-leksikografije, drugom je projektu cilj otvaranje resursa RCJHR većem krugu zainteresiranih znanstvenika i korisnika objavom na mrežnim stranicama Staroslavenskoga instituta, uz istovremeno daljnje razvijanje i dopunjavanje tih resursa. Najvažnija je u tom smislu planirana retrodigitalizacija RCJHR koja će omogućiti jednostavnije služenje dosad objavljenim svescima rječnika, ali i postaviti okvir za daljnju leksikografsku obradu. Oba su spomenuta trenutačno djelatna projekta zamišljena tako da se međusobno nadopunjuju.

Željka Brlobaš

Institut za hrvatski jezik

zbrlobas@ihjj.hr

Rječnik hrvatskoga kajkavskoga književnog jezika

Povijesni *Rječnik hrvatskoga kajkavskoga književnog jezika* jedan je od leksikografskih projekata u Institutu za hrvatski jezik i leksikografska publikacija koja se objavljuje u svescima. U izlaganju se navodi izvedbeni tijek od klasičnoga tiskanog oblika *Rječnika* (odluka o izradi,

probir i odabir pisanih izvora, stvaranje korpusa, metodološki postupci leksikografske obrade, struktura rječničkoga članka) preko postupka digitalizacije (uvođenje računala u leksikografsku obradu, digitalizacija prvotno tiskanih svezaka) do mrežnoga objavljivanja, čime je *Rječnik* osim ustaljenoga tiskanog oblika zadovoljio i suvremenu formu te je leksikografski postao dostupniji korisnicima u njihovu pretraživanju jezičnih podataka, posebice kada je riječ o hrvatskoj povijesnoj leksikografiji i kajkavskome hrvatskome književnom jeziku u povijesti hrvatskoga jezika.

Anita Celinć

Institut za hrvatski jezik

acelinic@ihjj.hr

Bednjanski rječnik

Bednjanski rječnik projekt je u okviru kojega se izrađuje dijalektni rječnik mjesnoga govora Bednje. Bednja je naselje i općina na sjeveru Hrvatskoga zagorja, u široj javnosti prepoznata, među ostalim, po mjesnome govoru koji se smatra teže razumljivim. Taj je govor, sa specifičnom kombinacijom konzervativnih i inovativnih elemenata u fonologiji, i u dijalektologiji razmjerno zahtjevniji za analizu.

Sa zahtjevnošću bednjanskoga govora proporcionalan je interes i lokalne zajednice i struke za objavom njegova rječnika. Lokalna zajednica intenzivno potiče izradu bednjanskoga rječnika, svjesna da taj zadatak nadilazi mogućnosti lokalnih entuzijastičnih amatera. Bednjanski se rječnik zato izrađuje kao znanstveni projekt IHJ-a. Za razliku od većine ostalih hrvatskih dijalektnih rječnika, autor mu neće biti izvorni govornik bednjanskoga govora, nego dijalektolog.

Projekt je utemeljen na opsežnome i kontinuiranome terenskom radu istraživača s izvornim govornicima bednjanskoga govora. Na taj se način prikuplja građa, koju istraživač zatim dijalektološki interpretira, unosi u računalo navodeći prozodijske i gramatičke podatke te leksikografski obrađuje. Rječničkom se tekstu pridodaje i građa iz studije Josipa Jedvaja iz 1956. godine, a izrađen će biti i standardnojezično-dijalektni indeks.

Na okruglom stolu bit će prikazani osnovni elementi rječničke natuknice i ukratko opisana načela i teškoće rada na navedenome tipu rječnika.

Ivana Lovrić Jović

Institut za hrvatski jezik

ilovric@ihjj.hr

Rječnik dubrovačkoga govora

Rječnik dubrovačkoga govora projekt je koji sufinancira *Next generation* (EU), a provodi se u Odjelu za povijest hrvatskoga jezika Instituta za hrvatski jezik. Osnovni mu je cilj izraditi dijakronijski mrežni rječnik dubrovačkoga govora temeljen na korpusu. Projekt obaseže razdoblje od 1. siječnja 2024. do 31. prosinca 2027. godine te će početkom 2028. godine biti javnosti ponuđen u besplatnoj portalnoj inačici.

Voditeljica je projekta dr. sc. Ivana Lovrić Jović, a suradnici su: dr. sc. Lana Hudeček, dr. sc. Marijana Horvat, dr. sc. Sanja Perić Gavrančić, dr. sc. Martina Kramarić, dr. sc. Domagoj Vidović i dr. sc. Josip Mihaljević. Osim potonjega, koji dolazi iz Staroslavenskoga instituta, ostali su suradnici iz Instituta za hrvatski jezik.

U tekućoj, drugoj projektnoj godini rad na rječniku odvija se prema planu, što znači da je korpus priređen i da se rječničke natuknice obrađuju. Do kraja godine bit će obrađeno 340 natuknica.

Veći je dio korpusa u čitljivoj elektroničkoj inačici, dok je manji njegov dio u faksimilima (rukopisna građa) ili u obliku zvučne snimke, što je od iznimne vrijednosti za izradu naglasaka. Korpus je izrađivan prema određenome konceptu rječnika, što znači da obaseže razdoblje od 16. do 21. st. i obuhvaća različite žanrove i funkcionalne stilove: književni, pučke čestitarske pjesme, pučki idiom stiliziran u dramskome tekstu, poslaničku prozu, administrativno-pravne tekstove, kronike, tekstove iz satiričnih novina itd.

Uz uobičajene izazove rječničarska posla, rad na ovome rječniku ima i mnoge druge. Za početak, ne postoji rječnik koji bi mu bio uzorom u konceptualnome smislu: objediniti leksik iz dubrovačke književnosti koja je označila temelj standardnoga jezika i onaj iz organskoga govora tijekom nekoliko stoljeća iznimno je zahtjevan zadatak. Jednako je tako teška bila odluka da se taj koncept odredi, ali su potrebe korisnika, a to su znanstvenici, kazalištarci i govornici, odredile nužnost izrade upravo takva rječnika. Nepostojanje pouzdana rječnika dubrovačkoga govora temeljena na korpusu već sad ima bitne posljedice, pa se i u znanstvenim radovima sve češće i bez ograda rabe podatci iz glosara koje su izradili građani u najboljoj namjeri, ali bez ikakvih suradnika iz rječničarskoga ili sličnoga područja.

Okrugli stol dobra je prigoda da izazove podijelimo s kolegama u svrhu najboljih rješenja.

Dubravka Ivšić Majić

Institut za hrvatski jezik

divsic@ihjj.hr

Etimološki rječnik – 3. faza

U izlaganju će se predstaviti projekt *Etimološki rječnik – 3. faza* i prikazati napredak aktivnosti na njemu. Projekt predstavlja logičan slijed izrade *Etimološkoga rječnika hrvatskoga jezika* urednika akademika Ranka Matasovića, a ujedno predstavlja i važan korak u modernizaciji hrvatske etimologije.

Najprije ćemo se ukratko podsjetiti na prve dvije faze, tj. na objavu prvoga i drugoga sveska Rječnika. Težište prezentacije bit će na ključnim aktivnostima 3. faze. Detaljno će se opisati izrada Indeksa riječi – govorit će se o tome zašto je indeks za približno 125 000 „sekundarnih” riječi potreban, kako se metodološki izrađuje i dokud se stiglo s tim zadatkom. Predstavit će se koncept mrežne inačice Rječnika. Pokazat će se kako će se iskoristiti strukturirana baza podataka iz programa Tshwanelex za stvaranje modernoga digitalnog resursa koji objedinjuje prednosti tiskanoga i digitalnoga izdanja.

Na kraju ćemo se dotaknuti budućnosti Rječnika. Posebno će se istaknuti funkcionalnost kontinuiranoga održavanja (ispravljanja i dopunjavanja natuknica) na novoj mrežnoj platformi, čime će Etimološki rječnik postati uvijek ažuran resurs.

Domagoj Vidović

Institut za hrvatski jezik

dvidovic@ihjj.hr

Rječnik najčešćih hrvatskih prezimena

Antroponomastička istraživanja hrvatskoga jezika jedna su od temeljnih zadaća Odjela za onomastiku i etimologiju. Izdavanjem *Rječnika suvremenih hrvatskih osobnih imena* hrvatska je onomastika dobila prvi hrvatski opći rječnik osobnih imena izrađen prema zahtjevima onomastičke struke. Izradba *Rječnika najčešćih suvremenih hrvatskih prezimena* bila bi prvim korakom u izradbi onomastički utemeljenoga rječnika hrvatskih prezimena koji bi se nadopunjavao s *Leksikom prezimena SR Hrvatske* (ur. Valentin Putanec i Petar Šimunović) iz 1976. i trosvezačnim *Hrvatskim prezimenikom* (ur. Franjo Maletić i Petar Šimunović).

Polazište bi za izradbu rječnika bila baza prezimena s podacima o broju nositelja prezimena, a izvor bi mu bili podatci dobiveni od Državnoga zavoda za statistiku. Povijesni bi se podatci crpili iz historiografskih i antroponomastičkih djela, pri čemu bi osim područja Republike Hrvatske bila obuhvaćena vrela iz Bosne i Hercegovine, Vojvodine u Srbiji te Boke kotorske i Svebarja u Crnoj Gori, područja u kojima su Hrvati autohtono stanovništvo. Time bi se dobili podatci relevantni za rekonstrukciju migracijskih kretanja važni i za upotpunjivanje povijesne hrvatske dijalektne slike. Svako bi prezime u rječniku bilo obrađeno prema modernim leksikografskim načelima, a etimologije koje se donose prethodno kritički prosuđene i raspravljene. Zbog iznimnoga zanimanja javnosti za jezično postanje prezimena i podrijetlo nositelja pojedinih prezimena riječ je o projektu kojim će se u velikoj mjeri popularizirati znanost. Iako postoje mnogi portali i publikacije u kojima se metodološki vrlo raznoliko obrađuju hrvatska prezimena, zasad ne postoji antroponomastički utemeljen rječnik s nužnim znanstvenim aparatom, pa bi izradba ovoga rječnika bila prvim korakom u tome smjeru.

Perina Vukša Nahod

Institut za hrvatski jezik

pvuksa@ihjj.hr

Mrežni dijalektni rječnici

U izlaganju će se predstaviti rad na trima projektima *Od fertuna do peruna – dijalektološka i etimološka analiza leksema povezanih sa semantičkim poljem kuhinje* (NextGenerationEU), *Rijekom neretvanskih riječi – od kuće do škole* (Ministarstvo kulture i medija Republike Hrvatske) i *Mrežni frazeološki rječnik zažapskih govora* (Općina Zažablje) koji se provode u Institutu za hrvatski jezik s ciljem izrade mrežnih dijalektnih rječnika.

Projektima se dijalektna građa želi prikazati na suvremen i standardiziran način, a svaki od njih donosi različite izazove u obradi građe prikupljene iz dijalektnih rječnika i radova ili recentnim terenskim istraživanjem.

Jedan od ciljeva projekta *Od fertuna do peruna – dijalektološka i etimološka analiza leksema povezanih sa semantičkim poljem kuhinje* tematski je rječnik koji obuhvaća građu svih triju narječja: čakavskoga, kajkavskoga i štokavskoga. Novina je to da je standardnojezični oblik temeljna natuknica za prikaz dijalektne građe. Time se želi omogućiti umrežavanje s ostalim rječnicima koji se izrađuju u Institutu za hrvatski jezik, ali i pretragu svih dijalektnih inačica za pojedini leksem. Natuknica je zamišljena u nekoliko blokova: standardološki, terminološki, dijalektološki, etimološki i frazeološki.

U sklopu projekta *Rijekom neretvanskih riječi – od kuće do škole* izrađuje se mrežni dijalektološki rječnik neretvanskih govora koji pripadaju dvama dijalektima: istočnohercegovačkomu i novoštokavskom ikavskom. Prikazat će se različiti moduli prilagođeni učenicima uključenima u projekt (od 1. razreda osnovne škole do 4. razreda srednje škole).

Mrežni frazeološki rječnik zažapskih govora bit će prvi mrežni dijalektološki frazeološki rječnik skupine novoštokavskih ijekavskih govora. Rječnik se izrađuje u programu za izradu rječnika TLex i sadržava rječničku natuknicu koja se sastoji od nadnatuknice (kanonski oblik riječi), natuknice (potvrđeni frazemi s nadnatuknicom u sastavu), značenja i rečeničnoga primjera. Svi će primjeri biti naglašeni i označeni hrvatskom dijalektološkom transkripcijom. Za pojedine će natuknice biti dostupan zvučni zapis, frazeološki članak, fotografija, ogled govora i poveznice na literaturu.

Lana Hudeček

Maja Matijević

Institut za hrvatski jezik

lhudecek@ihjj.hr, mmatijevic@ihjj.hr

Hrvatski mrežni rječnik – Mrežnik: 2. faza

Rad na izradi *Hrvatskoga mrežnoga rječnika – Mrežnika* odvijao se od 2017. do 2021. godine u sklopu istoimenoga projekta HRZZ-a (IP-2016-06-2141). Nastavljen je od 2021. do 2023. na (također istoimenome) internom projektu Instituta za hrvatski jezik, a od 2024. godine rad na izradi toga rječnika odvija se u sklopu projekta *Hrvatski mrežni rječnik – Mrežnik: 2. faza*. Projekt financira Europska unija u sklopu programa NextGenerationEU. Na projektu je u svim njegovim fazama sudjelovao velik broj suradnika (<http://ihjj.hr/mreznik/page/suradnici/3/>). Osnovni su ciljevi projekta, postavljeni i nepromijenjeni od njegova početka: 1. izrada hrvatskoga jednojezičnoga općega normativno-deskriptivnoga, na korpusu utemeljenoga i javno dostupnoga tromodulnoga rječnika utemeljena na spoznajama suvremene e-leksikografije; 2. doprinos teorijskomu području e-leksikografije.

U prvoj projektnoj fazi postavljeni su temelji na kojima se projekt zasniva i danas: uspostavljena je njegova makrostruktura (rječnik ima tri modula, odnosno tri zasebna rječnika: modul za odrasle govornike hrvatskoga jezika, modul za učenike nižih razreda osnovne škole te modul za nasljedne i neizvorne govornike) i mikrostruktura (struktura rječničkoga članka u svakome od modula), sastavljene upute za obrađivače, održane inicijalne radionice (povezne s radom u programu TLexu, uspostavom ujednačenoga modela obrade, dopunama skica riječi za hrvatski jezik), uspostavljene četiri popratne mrežne stranice koje sadržavaju frazemske etimologije, jezične savjete, popise naglašenih etnika i ktetika te etimologije riječi dostupne na portalu *Hrvatski u školi* (hrvatski.hr, rubrike: Frazemi, Savjeti, Etnici i ktetici i Etimologije). U prvoj projektnoj godini, 2017., sredstvima projekta financirana je izrada skica riječi za hrvatski jezik te je započeo i rad na *Pojmovniku e-leksikografije* (<http://ihjj.hr/mreznik/page/pojmovnik/6/>). Projekt HRZZ-a ocijenjen je 2021. najvišim ocjenama. Na njemu je na temelju natječaja u okviru programa *Projekt razvoja karijera mladih istraživača – izobrazba novih doktora znanosti* zaposleno dvoje doktoranada, Josip Mihaljević i Maja Matijević, koji su doktorirali u predviđenome roku. 2021. godine na mreži je u demoinačici objavljena građa A – F. 2019. održan je međunarodni znanstveni skup *E-rječnici i e-leksikografija* (sudionici iz 14 zemalja, radovi nastali na temelju izlaganja objavljeni su u Raspravama Instituta za hrvatski

jezik 46/2), a 2020. *Okrugli stol Rječnik i stereotipi* (radovi objavljeni u Raspravama 47/1). Do kraja 2024. godine na projektu su objavljena 92 znanstvena i znanstveno-popularna rada.

Na projektu *Hrvatski mrežni rječnik – Mrežnik: 2. faza*, koji financira Europska unija u sklopu programa NextGeneration EU, nastavljen je rad na rječniku, u prvome redu na uređivanju rječničke građe. 2024. na mreži je, u skladu s projektnim planom, objavljena građa do slova K, a 2025. do slova NJ. Objava preostale rječničke građe predviđena je za 2026. (do slova Š) te 2027. godinu (do slova Ž). Tijekom 2024. i 2025. godine, također u skladu s projektnim planom, dopunjena je i građa na portalu *Hrvatski u školi* (rubrike Frazemi, Jezični savjeti, Etnici i ktetici te Etimologije). Dopunjen je i *Pojmovnik e-leksikografije*. Provedene su brojne diseminacijske aktivnosti i aktivnosti usavršavanja, osobito mlađih suradnika, te je od 2024. objavljeno 16 znanstvenih i 8 znanstveno-popularnih radova. Ukupno je dosad u svih fazama izrade rječnika u časopisima i zbornicima objavljeno 116 znanstvenih i znanstveno-popularnih radova.

2024. objavljena je monografija *Anatomija rječnika – Hrvatski mrežni rječnik – Mrežnik*, čiji su autori suradnici na projektu i koja uključuje 42 poglavlja okupljena u 5 cjelina, a 2025. *Rječnik za nasljedne i neizvorne govornike*, autorica Lane Hudeček, Milice Mihaljević i Dinke Pasini, koji je proistekao iz *Mrežnikova* modula za osobe koje uče hrvatski kao drugi i strani jezik te je bitno dopunjen, između ostaloga i engleskim i španjolskim istovrijednicama. Monografija je dostupna i na <http://ihjj.hr/mreznik/page/monografija/34/>.

Tijekom rada na projektu uspostavljene su brojne veze s drugim institutskim projektima te su u Mrežnik uključene poveznice na mrežne stranice na kojima se nalaze podatci nastali u njihovom okviru (Kolokacijska baza hrvatskoga jezika, e-Glava, Struna, baza morfoloških dubleta – s kojom se Mrežnik povezao 2025.) i drugim znanstvenim ustanovama, npr. s Odsjekom za informacijske i komunikacijske znanosti Filozofskoga fakulteta Sveučilišta u Zagrebu, Croaticumom – Centrom za hrvatski kao drugi i strani jezik, Katedrom za hrvatski jezik i književnost Filozofskoga fakulteta u Pečuhu i dr.



ORGANIZACIJSKI ODBOR

Lana Hudeček, predsjednica Organizacijskoga odbora, voditeljica projekta *Hrvatski mrežni rječnik – Mrežnik*, Institut za hrvatski jezik

Željko Jozić, ravnatelj Instituta za hrvatski jezik

Milica Mihaljević, Institut za hrvatski jezik

Kristian Lewis, Institut za hrvatski jezik

Maja Matijević, Institut za hrvatski jezik

Dorotea Cakić, Institut za hrvatski jezik



**Financira
Europska unija**
NextGenerationEU