

UDK 811.163.42'322

811.163.42'36(091)

Pregledni rad

Rukopis primljen 27. II. 2024.

Prihvaćen za tisak 15. XI. 2024.

doi.org/10.31724/rihjj.50.2.5

Marijana Horvat

Institut za hrvatski jezik

Ulica Republike Austrije 16, HR-10000 Zagreb

orcid.org/0000-0003-1798-1098

mhorvat@ihjj.hr

Josip Mihaljević

Staroslavenski institut

Demetrova ulica 11, HR-10 000 Zagreb

orcid.org/0000-0002-7482-7663

jmihaljevic@stin.hr

PORTAL HRVATSKIH GRAMATIKA DO ILIRIZMA – MODEL IZRADE I PRIMJENA VIŠERAZINSKOGA OZNAČIVANJA

U članku¹ se analiziraju rezultati rada na projektu *Retrodigitalizacija i interpretacija hrvatskih gramatika do ilirizma – RETROGRAM*, koji su financirali Hrvatska zaklada za znanost (IP-2018-01-3585) i Institut za hrvatski jezik. Na projektu se, među ostalim, radilo na označivanju sklonidbenih paradigma i paradigma spreznja u odabranim dopreporodnim gramatikama i nazivlja u gramatikama pisanim hrvatskim jezikom. Članak se sastoji od dvaju glavnih dijelova. U prvome se dijelu u kratkim crtama opisuje proces označivanja i izrade digitalnih izdanja gramatika, a u drugome se objašnjava krajnji rezultat – uspostava *Portala hrvatskih gramatika do ilirizma*.

¹ Ovaj je rad nastao u okviru projekta Razvoj i primjena modela za normalizaciju grafije starih latiničnih tiskanih tekstova (MONOGRAF) i projekta *Razvoj modela digitalne infrastrukture Staroslavenskoga instituta – DigiSTIN*, koje financira Europska unija – NextGenerationEU. Za iznesene stavove i mišljenja odgovorni su samo autori te ti stavovi ne odražavaju nužno službena stajališta Europske unije ili Europske komisije. Ni Europska unija ni Europska komisija ne mogu se smatrati odgovornima za njih.

1. Uvod

Projekt *Retrodigitalizacija i interpretacija hrvatskih gramatika do ilirizma – RETROGRAM* (HRZZ IP-2018-01-3585 i IHJ, voditeljica dr. sc. Marijana Horvat; <https://retrogram.jezik.hr/>)² obuhvaća, među ostalim, višerazinsko označivanje transkribiranoga ili prevedenoga teksta odabranih hrvatskih dopreporodnih gramatika i njihovo povezivanje sa slikama izvornoga teksta (faksimilima). Cilj projekta bio je izraditi portal hrvatskih gramatika dopreporodnoga razdoblja sa slikama izvornika, metapodacima i tekstovima transkripcija ili prijevoda odabranih gramatika, omogućiti pretraživost gramatika na morfološkoj razini po zadanim elementima pretrage, istražiti i opisati gramatike te izraditi popis hrvatskoga dopreporodnog gramatičkog nazivlja s kraticom gramatike u kojoj se pojedinačni nazivi nalaze. Također je već prvotna zamisao bila da portal bude otvorenoga pristupa s mogućnošću trajnoga razvoja nakon dovršetka projekta, a pridonio bi očuvanju zastupljene građe i baštine te omogućio daljnja istraživanja. Opravdanost i nužnost pokretanja takvoga projekta, kao i kasniji prikaz određenih problema i trenutačnih rezultata rada na projektu, obrazloženi su u nekim ranijim radovima (v. Horvat 2020, Horvat i Kramarić 2021, Bago 2022, Barbarić 2023, Horvat i Perić Gavrančić 2023, Mihaljević, Perić Gavrančić i Tvrtković 2023). Stoga će se u ovome radu naglasak staviti na opis i objašnjenje digitalnih izdanja i označivanja gramatičkih kategorija te na računalne aspekte izrade *Portala hrvatskih gramatika do ilirizma*. Budući da su rezultati rada na spomenutom projektu inovativni ne samo u hrvatskoj nego i u europskoj filologiji jer je stvoren model za buduća slična istraživanja, smatramo potrebnim jezikoslovnu struku upoznati s tim rezultatima i osnovama postupka označivanja.

² Projekt se provodio u Institutu za hrvatski jezik od 25. studenoga 2019. do 24. studenoga 2023. Na projektu su surađivali: Petra Bago (Filozofski fakultet, Zagreb), Martina Kramarić (Institut za hrvatski jezik, Zagreb), Ivana Lovrić Jović (Institut za hrvatski jezik, Zagreb), Ana Mihaljević (Staroslavenski institut, Zagreb), Sanja Perić Gavrančić (Institut za hrvatski jezik, Zagreb), Ivo Pranjković (Filozofski fakultet, Zagreb), Diana Stolac (Filozofski fakultet, Rijeka), Ljiljana Šarić (Humanistički fakultet, Oslo, Kraljevina Norveška), Barbara Štebih Golub (Institut za hrvatski jezik, Zagreb), Tamara Tvrtković (Fakultet hrvatskih studija, Zagreb), Vuk-Tadija Barbarić (Institut za hrvatski jezik, Zagreb, izvanprojektni suradnik), Toma Tasovac (Centar za digitalne humanističke nauke, Beograd, Republika Srbija, savjetnik na projektu).

2. O gramatikama i označivanju

2.1. Digitalna izdanja

Rad na projektu *RETROGRAM* od samih je početaka bio obilježen dvama glavnim izazovima: 1. različite gramatike svesti na zajedničku strukturu potrebnu za jednoznačno označivanje; 2. osmisliti model označivanja gramatika jer ne postoji repozitorij koji bi sadržavao pretražive digitalizirane gramatike ni model koji bi se na njih mogao primijeniti.

Naime, nisu sve gramatike pisane hrvatskim jezikom niti su sve gramatike hrvatskoga jezika.

Hrvatske gramatike sa stranim metajezikom sadržavaju obilnu hrvatsku građu u sklonidbenim paradigama, paradigama sprezanja i hrvatskim primjerima. Gramatike stranih jezika (talijanskoga, latinskoga, njemačkoga) zbroju hrvatskih gramatika pridružuju odlomci i primjeri na hrvatskom te hrvatski kao metajezik (Horvat 2020: 635). Da bi se tako različite gramatike mogle svesti na zajedničku strukturu, što je bio preduvjet za izradu digitalnih izdanja i označivanje, morale su biti transkribirane suvremenim slovopisom (ako su pisane hrvatskim jezikom i starom grafijom) ili prevedene na hrvatski jezik (ako su pisane stranim jezikom). Zato govorimo o označivanju transkribiranoga ili prevedenoga teksta odabranih hrvatskih dopreporodnih gramatika (usp. Horvat i Perić Gavrančić 2023: 137). Korpus za izradu digitalnih izdanja čine ove gramatike objavljene u nizu *Pretisci* Instituta za hrvatski jezik:

Tablica 1: Izvori s obilježjima metajezika, jezika primjera u paradigmama i jezika predmeta opisa

gramatika	metajezik	primjeri u paradigmama	jezik predmet opisa	izvori za digitalno izdanje i označivanje
Bartol Kašić: <i>Institutionum linguae Illyricae libri duo</i> . Rim. 1604.	latinski	hrvatski i latinski	hrvatski	Bartol Kašić: <i>Institutionum linguae Illyricae libri duo / Osnove ilirskoga jezika u dvije knjige</i> . Pretisak prvoga izdanja (1604). Prijevod: Sanja Perić Gavrančić. Pogovor: Darija Gabrić-Bagarić. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2002.
Jakov Mikalja: <i>Gramatika talijanska ukratko ili kratak nauk za naučiti latinski jezik</i> . Loreto. 1649.; Loreto – Ancona, 1649. /1651.	hrvatski	hrvatski i talijanski	talijanski	Jakov Mikalja: <i>Gramatika talijanska ukratko ili kratak nauk za naučiti latinski jezik</i> . Pretisak prvoga izdanja (1649.). Popratna studija: Darija Gabrić-Bagarić i Marijana Horvat. Transkripcija: Marijana Horvat. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2008
Ardelio Della Bella: <i>Istruzioni grammaticali della lingua illirica</i> . Venecija. 1728.	talijanski	hrvatski i talijanski	hrvatski	Ardelio Della Bella: 2006. <i>Istruzioni grammaticali della lingua illirica / Gramatičke pouke o ilirskome jeziku</i> . Pretisak prvoga izdanja (1728.). Prijevod: Nives Sironić-Bonefačić. Pogovor: Darija Gabrić-Bagarić. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2006.

Blaž Tadijanović: <i>Svašta po malo iliti kratko složenje imena, riči u ilirski i njemački jezik.</i> Magdeburg. 1761.	hrvatski	hrvatski i njemački	njemački i hrvatski	Marijana Horvat i Ermina Ramadanović: <i>Jezikoslovni priručnik Blaža Tadijanovića Svašta po malo iliti kratko složenje imena, riči u ilirski i njemački jezik (1761.).</i> Pretisak prvoga izdanja (1761.). Transkripcija: Marijana Horvat. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2012.
Marijan Lanosović: <i>Uvod u latinsko riči slaganje s nikima nimačkog jezika biličkama za korist slovinskih mladića složen.</i> Osijek. 1776.	hrvatski	hrvatski, latinski i njemački	latinski	Sanja Perić Gavrančić: <i>Latinska gramatika i hrvatski jezik Marijana Lanosovića. Jezičnopovijesna studija s pretiskom i transkripcijom izvornika</i> Uvod u latinsko ricsih slaganje (1776.). Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2021.
Ignacije Szentmártony: <i>Einleintung zur kroatischen Sprachlehre für Deutsche.</i> Varaždin. 1783.	njemački	hrvatski i njemački	hrvatski/kajkavski	Barbara Štebih Golub: <i>Ignacije Szentmártony i</i> Uvod u nauk o horvatskome jeziku. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2014.

Josip Voltić: <i>Grammatica illirica</i> . Beč. 1803.	talijanski i njemački	hrvatski, talijanski i njemački	hrvatski	Josip Voltić: <i>Grammatica Illirica / Ilirska gramatika (1803.)</i> . Priredile i popratnu studiju napisale: Marijana Horvat i Sanja Perić Gavrančić. Transkripcija: Marijana Horvat i Sanja Perić Gavrančić. Prijevod: Sanja Perić Gavrančić. Institut za hrvatski jezik i jezikoslovlje. Zagreb. 2016.
Francesco Maria Appendini: <i>Grammatica della lingua Illirica</i> . Dubrovnik. 1808.	talijanski	hrvatski i talijanski	hrvatski	Ivana Lovrić Jović: <i>Appendinijeva Gramatika ilirskoga jezika – jezičnopovijesna studija s prijevodom i transkripcijom uz pretisak</i> . IHJJ – NSK. Zagreb. 2022.

Višejezičnost je dakle temeljno obilježje hrvatskih gramatika 17. i 18. stoljeća te se hrvatski jezik u tim gramatikama opisivao kontrastivnom metodom.³ Bartol Kašić prvu je hrvatsku gramatiku *Institutiones linguae Illyricae* (1604.) sastavio na latinskom jeziku i prema modelu latinske gramatike.⁴ „Posljedica je preuzimanja gramatičkoga modela drugoga jezičnog sustava nedostatan ili nedosljedan opis narodnoga (u ovome slučaju hrvatskoga) jezika: to je vidljivo u neprepoznavanju nekih jezičnih kategorija karakterističnih za jezik koji se opisuje ili u unošenju nekih novih kategorija koje opisani jezik ne poznaje.” (Mihaljević, Perić Gavrančić i Tvrtković 2023: 308). Nemogućnost potpunoga uklapanja hrvatskih oblika u latinski gramatički nacrt izrazit je, primjerice, u obradbi padeža. Dok latinski jezik ima šest padeža u obama brojevima, prema Kašiću hrvatski jezik nema isti broj padeža u jednini i množini: u jednini ima sedam padeža (nominativ, genitiv, dativ, akuzativ, vokativ, ablativ, sedmi padež (= instrumental)), a u množini osam (nominativ, genitiv, dativ, akuzativ, vokativ, ablativ, sedmi padež

³ *Nova ričoslovica ilirička* (1812.) Šime Starčevića, pisana ikavskom novoštokavštinom, prva je gramatika koja opisuje i normira hrvatski jezik bez usporedbe s drugim jezicima.

⁴ O uzorima v. Katičić 1981: 9; Mihaljević, Perić Gavrančić, Tvrtković 2023: 308–309.

(= instrumental), osmi padež (= lokativ)).⁵ Kašić prema latinskome sustavu u hrvatski padežni sustav (nepotrebno) uvodi ablativ, koji donosi s prijedlogom *od* i oblično identičnim s genitivom. U jednini izostavlja lokativ, tumačeći to njegovom podudarnošću s dativom, a u množini ga dodaje, što upućuje na zaključak da je Kašić uočio posebnosti hrvatskoga padežnog sustava u odnosu na latinski. Također je važno spomenuti da se u prilagodbi padežnoga sustava dopreporodni gramatičari nisu uvijek koristili istim nazivima za iste morfološke kategorije hrvatskoga jezika, što je prouzročilo nejasnoće u kategorizaciji padeža, pa tako nazivom *ablativ* u Appendinijevoj gramatici nije označen isti padež kao u Kašićevoj (Mihaljević, Perić Gavrančić i Tvrković 2023: 304), te su i takve situacije bile izazovne u prilagođivanju gramatičkoga teksta digitalnomu izdanju namijenjenom za označavanje i kasnije pretraživanje morfoloških paradigma i nazivlja. U označavanju su se primjenjivali standardnojezični nazivi. Stoga je kod označavanja, u skladu sa suvremenim nazivljem, Kašićev sedmi padež označen kao instrumental, a osmi kao lokativ. U gramatikama čiji padežni sustav sadržava ablativ koji je oblikom isti kao genitiv i dolazi s prijedlogom *od*, ablativ je i označen kao ablativ da bi se kod pretraživanja razlikovao od genitiva. Prijedlog *od* nije sastavni dio označenoga ablativa. Appendinijevi su ablativi označeni prema suvremenom nazivlju, npr. ablativ (npr. *vjetrimi*) kao instrumental, a ablativ 2. (npr. *vjetrijeh*, *vjetrih*) kao lokativ. Appendinijev padežni sustav u jednini čine *nominativ*, *genitiv*, *dativ*, *akuzativ*, *vokativ* i *ablativ* (= instrumental), a u množini *nominativ*, *genitiv*, *dativ*, *akuzativ*, *vokativ*, *ablativ* (= instrumental) i *ablativ drugi* (= lokativ). Naime, Appendini kritizira svoje prethodnike koji u padežnom sustavu upotrebljavaju ablativ kao padež koji oblično odgovara genitivu s prijedlogom *od*, smatrajući da sam genitiv može preuzeti ulogu toga ablativa (v. Appendini u Lovrić Jović 2022: 200–202). S obzirom na to da su se u označavanju primjenjivali standardnojezični nazivi, suvremenim su nazivima označeni i padeži u gramatikama koje padeže ne imenuju (npr. u Tadijanovićevoj gramatici) ili ih označuju brojevima (npr. u Voltićevoj gramatici).

⁵ Prvi je paradigmu od sedam padeža u jednini i množini u suvremenom smislu uveo Ignjat Alojzije Brlić u svojoj gramatici *Grammatik der illyrischen Sprache* (1833.), s tim što je lokativ sedmi, a instrumental šesti padež.

Sustav od sedam padeža u jednini i množini ima i Matija Antun Relković (*Nova slavonska i nimačka gramatika / Neue Slavonische und Deutsche Grammatik*, 1767.), ali on još zadržava ablativ i ne registrira lokativ, tj. padežni sustav čine *nominativus*, *genitivus*, *dativus*, *accusativus*, *vocativus*, *ablativus*, *sedmi casus instrumentalis seu societatis* (Relković 1767: 58–59).

Navedenim je primjerom s ablativom ilustriran jedan od mnogih važnih problema pri kojima se na projektu susretalo tijekom dubinske raščlambe svake gramatike i njezine pripreme za označivanje, kao i odluka koja je u vezi s tim donesena. Ovdje je također važno spomenuti da je u zaglavlje digitalnoga izdanja svake gramatike uneseno objašnjenje odluka za koje se procijenilo da ih je potrebno obrazložiti korisnicima, npr.: Ablativ je morfološki jednak genitivu i u sklonidbenim paradigmama dolazi s prijedlogom *od.*; Digitalno izdanje, u odnosu na tiskano, ne sadržava znanstvenu studiju ni ostale prateće tekstove povezane sa znanstvenom studijom.; Budući da je uz lične/osobne zamjenice uobičajeno donošenje kategorije lica, primjenjujemo to u označivanju.

Označivale su se dvije skupine: primjeri paradigma u svim gramatikama i nazivlje u gramatikama pisanim hrvatskim jezikom. Oznake su izrađene prema smjernicama TEI-ja (Text Encoding Initiative)⁶ za humanističke znanosti, unificirane su za sve gramatike i svedene na zajedničku strukturu.⁷ Elementi digitalnih izdanja i način označivanja osmišljeni su i izrađeni u sklopu projekta *RETROGRAM* jer, kao što je rečeno, nije postojala TEI shema za označivanje gramatika. Označivanje su ručno obavljali suradnici na projektu.⁸ Digitalno izdanje svake gramatike sadržava zaglavlje i glavni tekst. Zaglavlje na hrvatskome i engleskome jeziku donosi podatke o izdanju, predlošku/izvoru, financiranju, naslovu gramatike u prijevodu/transkripciji i u izvornome zapisu, podatke o au-

⁶ TEI (Text Encoding Initiative) konzorcij je koji razvija i održava smjernice namijenjene svima koji se bave pripremom i/ili obradom tekstnih resursa u digitalnome obliku. Smjernice su preporuke konzorcija kako na odgovarajući način prikazati svojstva onih tekstnih resursa koji se radi lakše računalne obrade moraju izričito označiti. TEI smjernice temelje se na jeziku XML te su time neovisne o računalnim programima i operativnim sustavima. Svaka XML oznaka unutar TEI-ja ima detaljan opis povezan s digitalnim prikazom teksta. Tim oznakama obilježava se struktura teksta i moguća druga svojstva koja se radi kasnije obrade žele posebno označiti. Kodiranje prema unaprijed definiranoj TEI shemi omogućava razmjenu podataka bez gubitaka informacija i bez mogućnosti da neka informacija bude pogrešno interpretirana (Bago 2015: 81).

⁷ Oznake je izradila suradnica Petra Bago.

⁸ Na označivanju gramatika radile su: Mikaljina gramatika – Marijana Horvat, Della Bellina gramatika – Martina Kramarić, Appendinijeva gramatika – Ivana Lovrić Jović, Voltičeva gramatika – Ana Mihaljević, Kašićeva gramatika i Lanosovićeva gramatika – Sanja Perić Gavrančić i Tamara Tvrtković, Tadijanovićeva gramatika – Ljiljana Šarić, Szentmártonyjeva gramatika – Barbara Štebih Golub i Diana Stolac. Nakon što je završeno označivanje svih gramatika, provedeno je testiranje, koje se odvijalo u nekoliko faza. Prvo je svatko pregledao gramatiku za koju je bio zadužen te ispravio uočene pogreške i propuste. Potom je suradnica Sanja Perić Gavrančić pregledala sve gramatike, ispravila pogreške i ujednačila označivanja. Nakon nje sve je gramatike pregledala voditeljica projekta Marijana Horvat, koja je također ispravila uočene pogreške i propuste. Završno testiranje označenih gramatika napravila je suradnica Petra Bago. Konačne inačice gramatika priređene za programera, odnosno za izradu *Portala hrvatskih gramatika do ilirizma* potvrdila je voditeljica projekta.

toru gramatike i samoj gramatici, podatke o projektu, voditelju projekta i suradnicima te o osobitostima transkripcije/prijevoda i samoga označivanja.

```

<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE TEI SYSTEM "retrogram-dtd-v03.dtd">
<TEI xmlns="http://www.tei-c.org/ns/1.0">

  <teiHeader>
    <fileDesc>
      <titleStmt>
        <title xml:lang="hr">Gramatika talijanska ukratko: digitalno izdanje</title>
        <title xml:lang="en">Short Italian Grammar: a digital edition</title>
        <author xml:lang="hr">
          <persName xml:id="jmikalja">
            <forename>Jakov</forename>
            <surname>Mikalja</surname>
          </persName>
        </author>
      </titleStmt>
    </fileDesc>
    <biblFull>
      <titleStmt>
        <title xml:lang="hr" type="main">
          <choice>
            <orig>Grammatika talianska u kratho illi kratak nauk za naucitti latinski jezik</orig>
            <reg>Grammatika talianska u kratho illi kratak nauk za naucitti latinski jezik</reg>
          </choice>
        </title>
        <title xml:lang="hr" type="sub">
          <choice>
            <orig>Koga zlovinski upiša Otac Jacov Mikaglja Druzgbe Ifufove</orig>
            <reg>Koga zlovinski upissa Otac Jacov Mikaglja Druzgbe Isusove</reg>
          </choice>
        </title>
        <author>
          <persName xml:lang="hr" corresp="#jmikalja">
            <forename>Jacov</forename>
            <surname>Mikaglja</surname>
          </persName>
        </author>
      </titleStmt>
      <publicationStmt>
        <publisher> Po Paulu, i Ivanu Batiŕti Serafinu</publisher>
        <pubPlace xml:id="loreto">U Loretu</pubPlace>
        <date when="1649">MDCXLIX.</date>
        <authority>
          <choice>
            <orig>Superiorum Permifsu</orig>
            <reg>Superiorum Permissu</reg>
          </choice>
        </authority>
      </publicationStmt>
    </biblFull>
  </teiHeader>

```

Slika 1 i 2: Jakov Mikalja, *Gramatika talijanska ukratko ili kratak nauk za naučiti latinski jezik* – iz zaglavlja digitalnoga izdanja

<note xml:lang="hr" type="bio">Blaž Tadijanović rođen je u Rastušju 1728. godine. Kao franjevački kapelan zarobljen je 1757. godine kod Breslave (danas Wrocław, Republika Poljska). U logoru Küstrin (danas Kostrzyn) ostao je do kraja Austrijsko-pruskoga rata 1763. godine, dragovoljno jer nije htio ostaviti svoje sunarodnjake bez dušobrižnika. Od 1777. do 1780. godine bio je franjevački provincijal u Baji. Umro je u Cerniku 1797. godine. Boraveći u zarobljeništvu, napisao je i tiskao dvije knjige. Jezični priručnik Svašta po malo iliti kratko složenje imena i riči u ilirski i njemački jezik objavljen je u Magdeburgu 1761. (drugo izdanje 1766. u Tropavi, danas Opava, Češka Republika). Njime je Tadijanović nastojao pomoći hrvatskim vojnicima pri komuniciranju u njemačkoj sredini, ali se djelo moglo koristiti i za usvajanje gramatike hrvatskoga jezika, odnosno pri učenju hrvatskoga jezika. Molitvenik Sertum ex diversis floribus tiskan je također u Magdeburgu 1761. godine, ali djelo nije sačuvano.</note>

<note xml:lang="en" type="bio">Blaž Tadijanović was born in Rastušje in 1728. He was captured as a Franciscan chaplain in 1757 near Breslau (now Wrocław, Poland). He voluntarily remained in the camp at Küstrin (now Kostrzyn) until the end of the Austro-Prussian War in 1763 because he did not want to leave his compatriots without a pastor. From 1777 to 1780, he was the Franciscan Provincial of Baja. He died in Cernik in 1797. In captivity, he wrote and published two books. The language manual Svašta po malo ili kratko složenje imena i riči u ilirski i njemački jezik (Miscellany, or a Short Illyrian and German Grammar) was published in Magdeburg in 1761 (second edition 1766 in Tropava, now Opava, Czech Republic). With it, Tadijanović tried to help Croatian soldiers to communicate in a German environment, but the book could also be used for learning Croatian and its grammar. His prayer book Sertum ex diversis floribus (A Garland of Diverse Flowers) was also published in Magdeburg in 1761, but it has not been preserved.</note>

Slika 3: Blaž Tadijanović, *Svašta po malo iliti kratko složenje imena i riči u ilirski i njemački jezik* (1761.) – iz zaglavlja digitalnoga izdanja

2.2. Označivanje paradigma i naziva

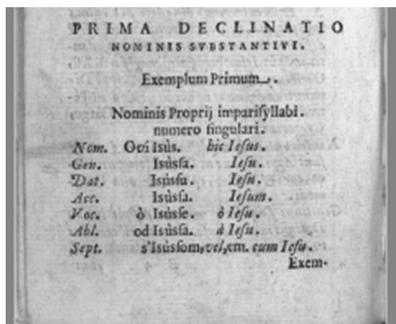
Glavni tekst svakoga digitalnog izdanja čini tekst transkripcije ili prijevoda gramatike u formatu XML s oznakama paradigma i nazivlja. Kod označivanja paradigma označivane su sljedeće vrste riječi: imenice, zamjenice, pridjevi, brojevi i glagoli. Kako TEI smjernice nemaju poseban modul za označivanje gramatika, odlučeno je prilagoditi postojeći modul za označivanje rječnika. Budući da se često morfološke paradigme prikazuju u tabličnome obliku, za potrebe projekta odlučeno je da se zanemari prezentacijski oblik paradigme te da se označe informacije sadržane u tablicama. Za označivanje jedne leksičke jedinice u paradigmi upotrijebljen je element <form>. Dogovorene su i utvrđene obvezatne i neobvezatne informacije, koje su označivane za svaku vrstu riječi te je razvijena prilagođena shema prema TEI-ju, tj. baza za razvoj modula za gramatike.⁹

⁹ Cijeli ulomak prema Bago 2022: 237.

Primjer obvezatnih i neobvezatnih informacija za imenice i glagole:

a) Kod imenica označujemo:

- obvezatne informacije: vrsta riječi, vrsta imenice, rod, broj, padež, sklonidba
- neobvezatne informacije: život.



```

<p>PRVA DEKLINACIJA IMENICA</p>
<p>Prvi primjer</p>
<p>Vlastita nejednakosložna imenica</p>
<p>Jednina</p>
<p>Nom. ovi Iesus                hic Iesus</p>
<p>Gen. Isusæ                    Iesuc</p>
<p>Dat. Isusu                    Iesuc</p>
<p>Ak. Isusæ                    Iesum</p>
<p>Vok. ò Isuse                  ò Iesu</p>
<p>Abl. od Isusæ                à Iesuc</p>
<p>Sedmi s Isusom ili em      cum Iesuc</p>
<p>Nom. ovi
<form type="inflectedForm" xml:lang="hr">
<orth>Isus</orth>
<gramGrp>
<gram type="pos" corresp="#imenica"/>
<gram type="nounType" corresp="#I_vlastita"/>
<gram type="gender" corresp="#muski"/>
<gram type="number" corresp="#jednina"/>
<gram type="case" corresp="#nominativ"/>
<gram type="animacy" corresp="#I_zivo"/>
<gram type="inflectionType" corresp="#I_a_sklonidba"/>
</gramGrp>
</form>
hic Iesus</p>

```

Slika 4, 5 i 6: Bartol Kašić, *Institutionum linguae Illyricae libri duo* (1604.) – primjer slike izvornika, transkripcije u formatu XML te označene imenice u NJD.

```

<p>Kad se pita čigovo je, rečemo
<form type="inflectedForm" xml:lang="hr">
<orth>vojnika</orth>
<gramGrp>
<gram type="pos" corresp="#imenica"/>
<gram type="nounType" corresp="#I_opca"/>
<gram type="gender" corresp="#muski"/>
<gram type="number" corresp="#jednina"/>
<gram type="case" corresp="#genitiv"/>
<gram type="inflectionType" corresp="#I_a_sklonidba"/>
<gram type="animacy" corresp="#I_zivo"/>
</gramGrp>
</form>
del soldato</p>
<p>Kad hoćemo da mu se što da, rećemo  daj
<form type="inflectedForm" xml:lang="hr">
<orth>vojniku</orth>
<gramGrp>
<gram type="pos" corresp="#imenica"/>
<gram type="nounType" corresp="#I_opca"/>
<gram type="gender" corresp="#muski"/>
<gram type="number" corresp="#jednina"/>
<gram type="case" corresp="#dativ"/>
<gram type="inflectionType" corresp="#I_a_sklonidba"/>
<gram type="animacy" corresp="#I_zivo"/>
</gramGrp>
</form>
al soldato</p>

```

Slika 7: Jakov Mikalja, *Gramatika talijanska ukratko* – primjer označivanja dvaju padeža imenice *vojnika*

b) Kod glagola označujemo:

- obvezatne informacije: vrsta riječi, vrsta glagola
- ako se može dodijeliti informacija za glagolsko vrijeme
 - obvezatne informacije: glagolsko vrijeme, klasifikacija glagolskih vremena, lice, broj, glagolsko stanje
 - ako je dodijeljena informacija „pluskvamperfekt” ili „perfekt”:
 - obvezatne informacije: rod
 - ako je dodijeljena informacija „futur II”:
 - neobvezatne informacije: rod
- ako se može dodijeliti informacija za glagolski način:
 - obvezatne informacije: glagolski način, lice, broj, glagolsko stanje
 - ako je dodijeljena informacija „kondicional I” ili „kondicional II”:
 - obvezatne informacije: rod
- ako se može dodijeliti kategorija glagolskoga pridjeva:
 - obvezatne informacije: glagolski pridjev, rod, broj
 - neobvezatne informacije: padež
- ako se može dodijeliti kategorija glagolskoga priloga:

- obvezatne informacije: glagolski prilog
- ako se može dodijeliti kategorija infinitiva:¹⁰
 - obvezatne informacije: infinitiv
 - neobvezatne informacije: glagolsko stanje (kod infinitiva).

```
<form type="inflectedForm" xml:lang="hr">
  <orth>imam</orth>
  <gramGrp>
    <gram type="pos" corresp="#glagol"/>
    <gram type="verbType" corresp="#G_glavni"/>
    <gram type="tense" corresp="#sadasnie"/>
    <gram type="tenseType" corresp="#G_prezent"/>
    <gram type="person" corresp="#prvo"/>
    <gram type="number" corresp="#jednina"/>
    <gram type="voice" corresp="#aktiv"/>
  </gramGrp>
</form>
```

Slika 8: Jakov Mikalja, *Gramatika talijanska ukratko* – primjer označivanja glagola

Kad se u paradigmi nalazi samo nastavak, a ne cijeli oblik riječi (tipa *velik*, *-ka*, *-ko*; *imajući*, *-ća*, *-će*), u element „form” dodan je atribut „expand”, čija je vrijednost cijeli oblik riječi.¹¹

¹⁰ Informacije za označivanje supina:

```
<form type="inflectedForm" xml:lang="hr">
  <orth>neki supin</orth>
  <gramGrp>
    <gram type="pos" corresp="#glagol"/>
    <gram type="verbType" corresp="#G_glavni"/>
    <gram type="infinitive" corresp="#G_supin"/>
  </gramGrp> </form>
```

Primjer označivanja supina (Ignacije Szentmártony, *Einleitung zur kroatischen Sprachlehre für Deutsche*):

```
<p>
  <form type="inflectedForm" xml:lang="hr">
    <orth>malat</orth>
    <gramGrp>
      <gram type="pos" corresp="#glagol"/>
      <gram type="verbType" corresp="#G_glavni"/>
      <gram type="infinitive" corresp="#G_supin"/>
    </gramGrp> </form>
```

¹¹ Primjer: Marijan Lanosović, *Uvod u latinsko riči slaganje*:

```
<form type="inflectedForm" expand="velike" xml:lang="hr">
  <orth>-ke</orth>
  <gramGrp>
    <gram type="pos" corresp="#pridjev"/>
    <gram type="adjectiveType" corresp="#opisni"/>
    <gram type="gender" corresp="#zenski"/>
    <gram type="number" corresp="#mnोजना"/>
    <gram type="case" corresp="#vokativ"/>
  </gramGrp>
</form>
```

U gramatikama pisanim hrvatskim jezikom (*Gramatika talijanska ukratko* Jakova Mikalje, *Svašta po malo* Blaža Tadijanovića i *Uvod u latinsko riči slaganje* Marijana Lanosovića) označeno je morfološko nazivlje.

```
<pb n="5"/>
<p>II. <term corresp="#imenica">Samostavna</term> mogu od dvi vrsti broj imati, to
jest: <term corresp="#jednina">jedini ili jednostavni</term> (singularis), <term
corresp="#mnozina">veći ili višjestavni</term> (pluralis). Prvi zlamenuje jednu samo
stvar, nap: ovaj gospodin. Drugi zlamenuje više stvari, kano: ova gospoda.</p>
<p>III. <term corresp="#imenica">Samostavna</term> trpe niko <term
corresp="#I_sklonidba">izvođenje ili pregibanje</term> (declinatio). </p>
<pb n="6"/>
<p><term corresp="#jednina">Jednostavno</term></p>
<p><term corresp="#padez">Spadanje</term></p>
<pb n="7"/>
<p><term corresp="#mnozina">Većestavno</term></p>
<pb n="37"/>
<p><term corresp="#vrijeme">Vrimena</term> (tempora), koja
<term corresp="#glagol">rič</term> prikazuje, brojimo pet:
<term corresp="#G_prezent">sadanje</term> (tempus praesens),
<term corresp="#G_imperfekt">polprošasto</term> </p>
```

Slika 9: Marijan Lanosović, *Uvod u latinsko riči slaganje* – primjer označavanja naziva

Izrađen je i popis dopreporodnoga gramatičkog nazivlja (<https://retrogram.jezik.hr/nazivlje/>). Nazivi su okupljeni pod standardnojezičnim hiperonimima s kriticom gramatike u kojoj se nalaze (npr. **množina**: broj većine Mik., broj mlogostruki Tad., broj veći ili višjestavni, većestavno Lan.). Osim iz Mikaljine, Tadijanovićeve i Lanosovićeve gramatike nazivlje je ispisano i iz četiriju priručnika: Josip Matijević, *Pomum granatum* (1771.), Tomo Mikloušić, *Syllabus vocabulorum Grammaticae Emmanuelis Alvari, in Croaticam linguam conversorum* (1796.), Antun Rožić, *Kratko naputjenje vu kruto hasnoviteh, i zevsema potrebnih temeljih dijačkoga jezika* (1820.), Antun Rožić, *Prvi temelji dijačkoga jezika za početnike vu domorodnom jeziku* (1821.) te iz gramatike Josipa Matijevića *Horvacka gramatika oder kroatische Sprachlehre* (1810.), koja nije bila obuhvaćena projektnim planom.

Završetkom opisanih aktivnosti dobivena je građa za izradu *Portala hrvatskih gramatika do ilirizma* kao temeljnoga rezultata rada na projektu *RETROGRAM*. Portal sadržava slike izvornika dvanaest gramatika s metapodacima

o gramatikama,¹² od kojih je tekstnim digitalnim izdanjem predstavljeno osam gramatika, odnosno transkripcijama ili prijevodima na kojima je provedeno označavanje (v. u ovome radu Tablicu 1). Gramatike se mogu listati i čitati. Pritiskom na svaku gramatiku dobivaju se metapodatci o gramatikama te osnovni podatci o projektu *RETROGRAM*. Stranice transkripcija ili prijevoda gramatika povezane su s pripadajućim slikama izvornoga teksta (faksimilima). Pretraživanje označenih morfoloških paradigma i nazivlja dobiva se pritiskom na oznaku *Tražilica*.

Zamišljeno je da će se *Portal* dalje razvijati nakon završetka projekta *RETROGRAM*, što se od početka 2024. čini u sklopu projekta *Razvoj i primjena modela za normalizaciju grafije starih latiničnih tiskanih tekstova (MONOGRAM)*¹³ digitalizacijom i označavanjem *Nove ričoslovice iliričke* (1812.) Šime Starčevića, kasnije i nekih drugih gramatika.

Portal hrvatskih gramatika do ilirizma nalazi se na stranici projekta *RETROGRAM* (<https://retrogram.jezik.hr/portal-hrvatskih-gramatika-do-ilirizma/>).¹⁴ Uz njega se na projektnim stranicama nalazi i cjelovit pregled rada na projektu *RETROGRAM*, zbornik radova s Međunarodnoga znanstvenog skupa *Hrvatske dopreporodne gramatike u europskome kontekstu te popis dopreporodnoga gramatičkoga nazivlja*.

Izrada stranice projekta *RETROGRAM* i *Portala* objašnjava se u nastavku ovoga rada.

¹² To su: Bartol Kašić, *Institutionum linguae Illyricae libri duo* (1604.), Jakov Mikalja, *Gramatika talijanska ukratko ili kratak nauk za naučiti latinski jezik* (1649.), Ardelio Della Bella, *Istruzioni grammaticali della lingua illirica* (1728.), Blaž Tadijanović, *Svašta po malo iliti kratko složenje imena i riči u ilirski i njemački jezik* (1761.), Marijan Lanosović, *Uvod u latinsko riči slaganje s nikima nimačkog jezika biličkama za korist slovinskih mladića složen* (1776.), Ignacije Szentmártony, *Einleitung zur kroatischen Sprachlehre für Deutsche* (1783.), Josip Voltić, *Grammatica illirica* (1803.), Francesco M. Appendini, *Grammatica della lingua Illirica* (1808.); Josip Matijević, *Pomum granatum* (1771.), Tomo Mikloušić, *Syllabus vocabulorum Grammaticae Emmanuelis Alvares, in Croaticam linguam conversorum* (1796.), Antun Rožić, *Kratko naputjenje vu kruto hasnoviteh i zevsema potrebnih temeljih dijačkoga jezika* (1820.), Antun Rožić, *Prvi temelji dijačkoga jezika za početnike vu domorodnom jeziku* (1821.).

Slike izvornika dobivene su suradnjom s Nacionalnom i sveučilišnom knjižnicom u Zagrebu.

¹³ Voditelj je toga projekta Vuk-Tadija Barbarić, a financira se iz fonda NextGenerationEU. Riječ je o projektu koji se nastavlja na raniji projekt *Razvoj modela za normalizaciju grafije starih latiničnih tiskanih tekstova* istoga voditelja. Uspostavljena je dobra suradnja između projekta *RETROGRAM* i projekta *Razvoj modela* tijekom trajanja obaju projekata (v. članak Barbarić 2023.) te se nastavak rada u sklopu novoga projekta nametnuo kao logičan izbor.

¹⁴ Tehnički i programski izradio ga je Josip Mihaljević.

3. O stranici projekta *RETROGRAM*

Stranica za projekt *RETROGRAM* izrađena je u sustavu za upravljanje sadržajem¹⁵ WordPress. Postoje besplatni sustavi otvorenoga koda slični sustavu WordPress poput Joomla i Drupala, ali je WordPress već dulje vremena najkorišteniji sustav za izradu mrežnih stranica jer trenutačno po statistici W3Techs – World Wide Web Technology Surveys (2023.) sačinjava oko 43 % mrežnih mjesta na internetu te je zabilježeno da među mrežnim mjestima na internetu koja su izrađena s pomoću sustava ili programima WordPress zauzima 63 %.¹⁶ Razlog za to je njegovo lako korištenje, mogućnosti koje nudi s dodatcima te aktivna zajednica koja radi na njegovu održavanju. Nakon provedene analize WordPress bio je očit izbor za izradu stranice projekta *RETROGRAM* te unutar nje *Portala hrvatskih gramatika do ilirizma*. Odabran je zato što:

- omogućuje da se lako izradi dizajn stranice s odabirom jednoga od ponuđenih grafičkih predložaka koji se dodatno uređuju. Predložci sustava grafički su responzivni te se time prilagođuju različitim veličinama zaslona, pa se sadržaj stranica može pregledavati na mobilnim uređajima i računalima (vidi Sliku 10).
- sustav sadržava različite funkcionalnosti za izradu različitih tipova sadržaja te mnoge dodatke koji se mogu instalirati radi proširenja funkcionalnosti stranice.
- omogućuje da više urednika koji pristupaju sustavu s pomoću korisničkih računa mogu uređivati i stvarati sadržaje stranice unutar grafičkoga sučelja. To je važno i pri pisanju objava na stranici povezanih s aktivnostima na projektu te pri doradi tekstova.

¹⁵ Sustav za upravljanje mrežnim sadržajem (engl. CMS) sučelje je koje se može primijeniti za izradu, objavu i posuvremenjivanje sadržaja mrežnih stranica. Značajke su toga sustava pokretanje na mrežnome pregledniku, pohrana podataka u bazi, usporedna mogućnost rada više korisnika na mrežnim resursima stranice te mogućnost dopuna sustava različitim programskim dodatcima (Cvitković 2017: 15).

¹⁶ <https://w3techs.com/technologies/details/cm-wordpress> (pristupljeno 14. studenoga 2023.)



Slika 10: Prikaz dizajna i responzivnosti mrežne stranice projekta *RETROGRAM*

Naslovnica za *Portal hrvatskih gramatika do ilirizma* napravljena je unutar sustava u kojemu se može odabrati određena gramatika koja se želi pregledati. Gramatike su složene u blokovima po tri stupca te su prikazane naslovnicama. Gramatika se bira pritiskom na naslovnicu ili na naziv ispod nje. Prikaz slika i tekst gramatike nije izrađen u WordPressu, nego u sustavu za listanje Tify, ali su sadržaji obaju sustava povezani s pomoću mrežne poveznice te se nalaze unutar istoga poslužioaca (servera).



Slika 11: Prikaz prvih triju gramatika u izborniku *Portal hrvatskih gramatika do ilirizma*

3.1. Programsko rješenje za pregled gramatika

Portal sadržava slike izvornika (faksimile) dvanaest gramatika, od kojih je, kao što je prethodno rečeno, njih osam označeno TEI oznakama. Označivalo se na tekstovima transkripcija ili prijevoda te je svaku stranicu (oznaka <pb>) zapisa-

nu u TEI dokumentu trebalo povezati s pripadajućom stranicom slike izvornika (faksimila). Za svaku je gramatiku također trebalo izvaditi metapodatke na hrvatskome i engleskome jeziku (naslov, životopis, urednici itd.). Utvrđeno je da sustav za pregled gramatika treba imati:

- mogućnost uzastopnoga pregleda slike i teksta koji će funkcionirati na različitim rezolucijama
- mogućnost listanja i prebacivanja između stranica knjige
- prikaz impresuma i dodatnih metapodataka povezanih s gramatikom koji moraju biti dvojezični.

Prva demoinačica takvoga sustava za pregled gramatika funkcionirala je tako da se na jednome dijelu zaslona nalazi PDF stranica dokumenta, a na drugome dijelu tekst te se s pomoću tipka mijenjaju stranice i tekst. Unatoč tomu što je rješenje bilo funkcionalno, bilo je nepregledno (posebno na uređajima s manjim zaslonom). Usto je i dizajn izgledao prejednostavno. Na kraju je odlučeno iskoristiti IIF¹⁷ preglednik dokumenata Tify¹⁸, koji je razvijen na sveučilištu Göttingen. Isti program upotrijebljen je za pregled stranica mrežne inačice knjige *Ujevićeva Hrvatska enciklopedija (1941. – 1945.)* Leksikografskoga zavoda Miroslav Krleža¹⁹ te na temelju implementacije programa na toj stranici, pa smo ga odlučili isprobati. Program Tify dostupan je za besplatno korištenje pod licencijom GNU Affero General Public License 3.0.²⁰ Tify omogućuje da se unutar JSON²¹ formata povežu slike, tekst te po potrebi ostali podatci. Osim toga, ima lijepo dizajnirano i responzivno grafičko sučelje koje omogućuje lak i ugodan pregled dokumenata (vidi 12.).

¹⁷ The International Image Interoperability Framework (IIF) podrazumijeva skup programskih rješenja koja korisniku omogućuju da standardiziranim metodama prikažu sliku ili skup povezanih slika na mreži koje sadržavaju očuvane strukturirane metapodatke (IIF 2023).

¹⁸ <https://tify.rocks/> (pristupljeno 25. kolovoza 2023.)

¹⁹ <https://hemu.lzmk.hr/> (pristupljeno 14. studenoga 2023.)

²⁰ GNU Affero General Public License 3.0. omogućuje slobodno korištenje, izmjenu i distribuciju programskih kodova pod uvjetom da korišteni kodovi budu zaštićeni istom licencijom kako se ne bi ograničila uporaba otvorenoga koda i slobodnih programskih rješenja (GNU, 2008).

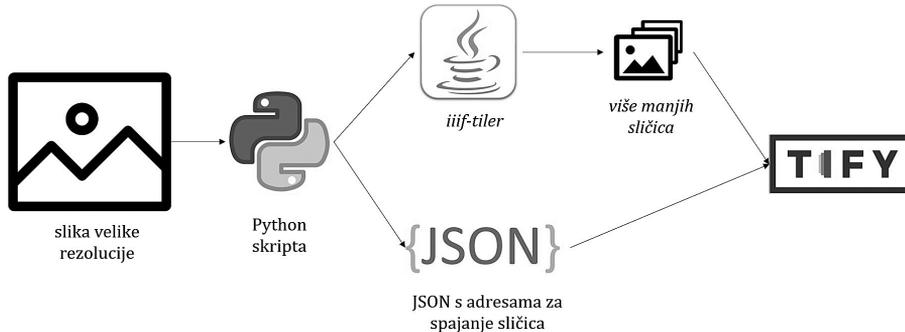
²¹ JSON (JavaScript Object Notation) tekstno je otvoreni standard dizajniran za čitljivu razmjenu podataka koji se temelji na sintaksi JavaScripta. Koristi se u mrežnome razvoju jer omogućuje interoperabilnu, brzu i neometanu razmjenu informacija između različitih sustava i programa (Techopedia 2011).



Slika 12: Izgled sučelja i funkcionalnost u programu Tify

Kodovi za korištenje programom Tify mogu se preuzeti sa službene mrežne stranice, ali je za korištenje i prilagodbu programa potrebno pročitati priručnik o korištenju. Tekstovi i slike kojima se želi koristiti unutar programa Tify dodatno se trebaju obraditi. Najveći izazov bio je prilagoditi slike gramatika jer program prikazuje slike na više sličica različitih rezolucija. Ovisno o povećanju slike i rezoluciji ekrana slike knjiga prikazuju se na više međusobno povezanih manjih sličica koje se učitavaju na određenim pozicijama (npr. kad povećamo određeni dio slike, program učita samo sličice u većoj rezoluciji koje čine taj povećani dio slike). Razlog je za takav način prikaza slika da se održi dobra kvaliteta slike tijekom pregleda na različitim rezolucijama uređaja te da se optimizira vrijeme učitavanja visokorezolucijskih slika jer kod manjega povećanja slike nije potrebno da se učita slika u najvećoj kvaliteti. Ipak, kako bi slike bile kompatibilne za Tify, trebale su biti u velikim rezolucijama prije podjele na sličice (npr. 15 000 x 12000 px) te je trebalo naći programsko rješenje koje stvara sličice koje su kompatibilne s IIIF standardom kako bi radile u programu Tify. Za izradu sličica koristio se program `iiif-tiler` napravljen u Javi (Robson 2023). Budući da se za svaku sličicu morala izravno napisati i adresa unutar JSON datoteke, potrebno je bilo unaprijed odrediti domene na kojima će se nalaziti sličice te ih integrirati u kod. To je riješeno prilagođenim korištenjem programa `iiif-tiler` unutar Python-skripte. Python-skripta visokorezolucijsku sliku podijeli na sličice s pomoću programa `iiif-tiler` te unutar svih JSON-a upiše točnu adresu. S pomoću druge

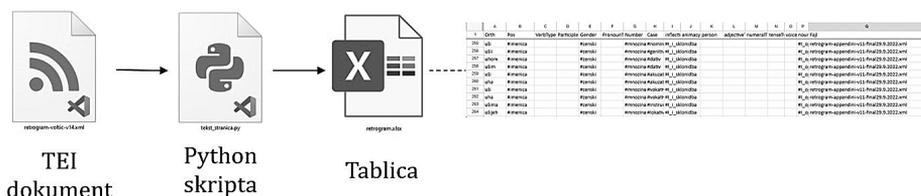
Python-skripte također se dohvati iz TEI dokumenata tekstove za stranice slika koristeći se skriptom Beautiful Soup, koja se onda povezala u datoteke JSON slika. Svaka gramatika također sadržava metapodatke koji su prikazani u dvama jezicima (hrvatski i engleski).



Slika 13: Proces izrade sličica za prikaz u programu Tify

3.2. Programsko rješenje za pretraživanje označenih morfoloških paradigma i naziva

Osim programa za pregled gramatika trebalo je omogućiti korisnicima da s pomoću tražilice mogu pretraživati vrste riječi i nazive unutar označenih gramatika. Budući da se pretražuje prema unaprijed određenim oznakama i svojstvima, nije trebalo napraviti traku za pretraživanje, nego se pretražuju vrste riječi i nazivi na temelju odabira vrijednosti u poljima povezanim s njihovim svojstvima. Polja u kojima se biraju vrijednosti padajući su izbornici koji čitaju vrijednosti koje se nalaze u TEI dokumentima. Ovisno o izboru vrste riječi (imenica, zamjenica, pridjev, broj, glagol) otvaraju se dodatna polja za pretraživanje (npr. za imenice se može dodatno pretraživati vrsta, rod, broj, sklonidba, padež, živost, a za glagole se mogu pretraživati glagolska vremena, lice, broj, glagolski način, glagolski pridjev, glagolski prilog, glagolsko stanje itd.). Baza tražilice izrađena je na temelju oznaka i vrijednosti koje se nalaze u TEI dokumentima gramatika. TEI dokumenti automatski su obrađeni te su polja i njihove vrijednosti pretvoreni u tablicu unutar baze s pomoću Python-skripte.



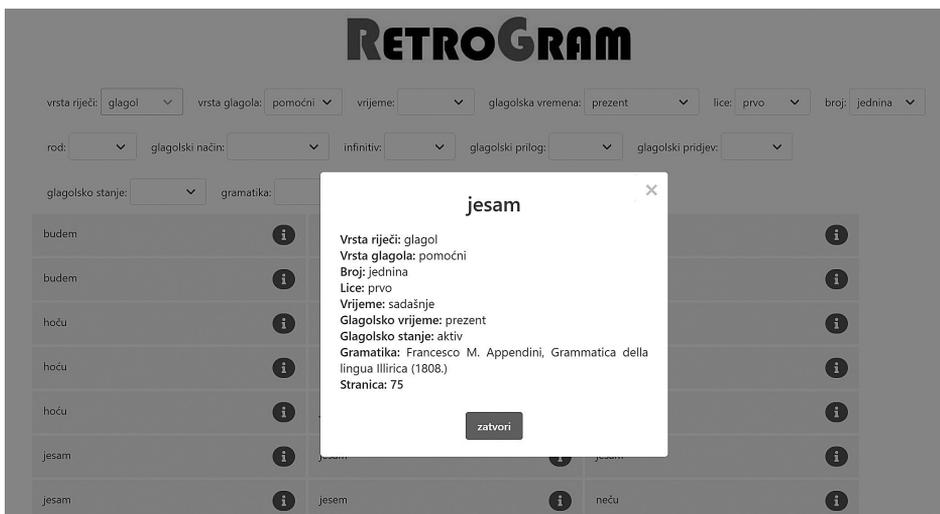
Slika 14: Proces vađenja podatka iz TEI dokumenata s pomoću Python-skripte

Tablice iz baze podataka učitavaju se izravno na mrežnoj stranici na kojoj se nalazi tražilica. Rezultati nakon pretrage ispisuju se ispod polja tražilice te je, kad se prijede mišem iznad oznake ⓘ, moguće dobiti informacije o tome iz koje je gramatike koji podatak.



Slika 15: Izgled tražilice za pretraživanje morfoloških kategorija

Pritiskom na oznaku ⓘ otvara se infopropzor koji daje dodatne informacije povezane s dobivenim rezultatima pretrage.



Slika 16: Prikaz informacija u infoprozoru za dobiveni rezultat pretrage

Također je u infoprozoru dodana i mogućnost da se otvori stranica gramatike unutar sustava za listanje na stranicama Portala na kojima se može vidjeti primjer u kontekstu.

4. Zaključak

Uspostavom *Portala hrvatskih gramatika do ilirizma* kao središnjim rezultatom rada na projektu *RETROGRAM* hrvatska je filologija, među ostalim, dobila digitalizirane i, u okvirima morfoloških paradigma promjenjivih vrsta riječi i morfološkoga nazivlja, pretražive odabrane dopreporodne gramatike. Svakako je bitna činjenica da su slike izvornika gramatika povezane s pripadajućim stranicama transkripcija ili prijevoda gramatika, što proširuje broj korisnika *Portala* i omogućuje različita istraživanja, posebice komparativna i terminološka, i ne samo kroatistička. *Portal* je otvorenoga pristupa, a pridonosi i očuvanju navedene građe i baštine. Kako je zamišljen za trajni razvoj nakon dovršetka projekta *RETROGRAM*, rad je od početka 2024. nastavljen u sklopu novoga institutskog projekta *Razvoj i primjena modela za normalizaciju grafije starih latiničnih tiskanih tekstova (MONOGRAF)*.

Važno je na kraju istaknuti da je razvijen i predložen model označivanja gramatika koji se može primijeniti na druge jezike, a ne samo na hrvatski. Model je također moguće nadograđivati za pretraživanje ostalih gramatičkih područja i kategorija.

Literatura

APPENDINI, FRANCESCO MARIA. 2022. *Grammatica della lingua Illirica / Gramatika ilirskoga jezika*. 2022. U: Lovrić Jović, Ivana. 2022. *Appendinijeva Gramatika ilirskoga jezika – jezičnopovijesna studija s prijevodom i transkripcijom uz pretisak*. IHJJ – NSK. Zagreb.

BAGO, PETRA. 2015. Označavanje natuknice u della Bellinu rječniku i povezivanje autora izvora citata s vanjskim resursima. *Studia lexicographica* 8/2. Zagreb. 77–102.

BAGO, PETRA. 2022. Progress of the RETROGRAM Project: Developing a TEI-like Model for Croatian Grammars Books before Illyrism. *Proceedings of the Conference on Language Technologies and Digital Humanities*. Ur. Fišer, Darja; Erjavec, Tomaž. Inštitut za novejšo zgodovino. Ljubljana. 235–238.

BARBARIĆ, VUK-TADIJA. 2023. *Nova ričoslovica ilirička između dvaju projekata. Hrvatske dopreporodne gramatike u europskome kontekstu: Zbornik radova*. Ur. Horvat, Marijana. Institut za hrvatski jezik. Zagreb. 17–26.

CVITKOVIĆ, MAJA. 2017. *Osnove rada u WordPressu*. Sveučilišni računski centar. Zagreb.

HORVAT, MARIJANA. 2020. Istraživanje povijesti hrvatskoga jezika u digitalno doba. *Rasprave: Časopis Instituta za hrvatski jezik i jezikoslovlje* 46/2. Zagreb. 635–643. doi.org/10.31724/trihjj.46.2.10.

HORVAT, MARIJANA; KRAMARIĆ, MARTINA. 2021. Retro-Digitization of Croatian Pre-Standard Grammars. *Athens Journal of Philology* 8/4. 297–310.

HORVAT, MARIJANA; PERIĆ GAVRANČIĆ, SANJA. 2023. On Historical-Grammatical Terminology in the RETROGRAM Project. *Collegium Antropologicum* 47/2. 137–144. doi.org/10.5671/ca.47.2.5.

HRVATSKA ENCIKLOPEDIJA (1941.–1945.). <https://hemu.lzmk.hr/> (pristupljeno 14. studenoga 2023.).

INTERNATIONAL IMAGE INTEROPERABILITY FRAMEWORK. <https://iiif.io/> (pristupljeno 25. kolovoza 2024.).

KATIČIĆ, RADOSLAV. 1981. Gramatika Bartola Kašića. *Rad JAZU* 388. 5–129.

MAFFULLI, STEFANO. 2008. GNU Affero General Public License version 3. <https://opensource.org/licenses/agpl-v3/> (pristupljeno 25. kolovoza 2023.).

MIHALJEVIĆ, ANA; PERIĆ GAVRANČIĆ, SANJA; TVRTKOVIĆ, TAMARA. 2023. Što su ablativ, konjunktiv i gerund u hrvatskome jeziku? – prilagodba latinskih gramatičkih kategorija opisu hrvatskoga jezika u dopreporodnim gramatikama. *Hrvatske dopreporodne gramatike u europskome kontekstu: Zbornik radova*. Ur. Horvat, Marijana. Institut za hrvatski jezik. Zagreb. 303–332.

ROBSON, GLEN. 2022. GitHub. <https://github.com/glenrobson/iiif-tiler> (pristupljeno 25. kolovoza 2023.).

ROUSE, MARGARET. 2011. *What is JavaScript Object Notation (JSON)?*. Techopedia. <https://www.techopedia.com/definition/3930/javascript-object-notation-json> (pristupljeno 25. kolovoza 2023.).

TIFY. <https://tify.rocks/> (pristupljeno 25. kolovoza 2023.).

Usage statistics and market share of WordPress. 2023. W3Techs. <https://w3techs.com/technologies/details/cm-wordpress> (pristupljeno 14. studenoga 2023.).

Portal of Croatian Grammar Books Before Illyrism: Construction Model and Usage of Multilevel Markup

Abstract

The authors analyze the results of the project *Retrodigitization and Interpretation of Croatian Grammar Books before Illyrism – RETROGRAM*, financed by the Croatian Science Foundation (IP-2018-01-3585) and the Institute for the Croatian Language. One of the project's key tasks was to label declension and conjugation paradigms in selected grammar books before Illyrism. Another important task was to label the terms in grammar books written in Croatian.

The paper consists of two major parts. The first part briefly outlines the processes of marking and compiling digital editions of grammar books, while the second part presents the outcome of these processes – the *Portal of Croatian Grammar Books before Illyrism*. This portal is the central deliverable of the *RETROGRAM* project and provides Croatian philology with digitized and searchable morphological paradigms of changeable word classes and morphological terminology from grammar books before Illyrism.

The portal connects photos of grammar book originals (facsimiles) to the corresponding pages of transcriptions or translations, enabling users to conduct a variety of research, including comparative and translational studies, beyond the field of Croatian studies.

The Portal is freely accessible and contributes to the preservation of grammar book data and heritage.

Conceived as a growing resource, the portal continues to develop within the new project of the Institute for the Croatian Language, *Development and Application of a Model for Normalizing the Orthography of Old Texts Printed in Latin Script – MONOGRAF*. Importantly, the proposed model for labeling grammar books is not limited to Croatian and can be applied to other languages. Additionally, the model is designed for potential upgrades, enabling research into other grammatical fields and categories.

Ključne riječi: digitalizacija, hrvatske dopreporodne gramatike, označivanje teksta, TEI oznake, programska rješenja, pregled digitaliziranih gramatika, pretraživanje računalnih gramatika, portal

Keywords: digitization, Croatian pre-standard grammar books, text encoding, TEI tags, software solutions, browsing digitized grammar books, grammar search, portal