

JEZIKOVNI VIRI CJVT IN NJIHOVA RABA V IZOBRAŽEVALNE NAMENE

ŠPELA ARHAR HOLDT,^{1,2} IZTOK KOSEM^{1,3} IN EVA PORI¹

¹ Univerza v Ljubljani, Filozofska fakulteta, Ljubljana, Slovenija.

E-pošta: spela.arharholdt@ff.uni-lj.si, iztok.kosem@cjvt.si, eva.pori@ff.uni-lj.si

² Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Ljubljana, Slovenija.

E-pošta: spela.arharholdt@ff.uni-lj.si

³ Institut »Jožef Stefan«, Ljubljana, Slovenija.

E-pošta: iztok.kosem@cjvt.si

Povzetek Center za jezikovne vire in tehnologije (CJVT) je interdisciplinarna raziskovalna enota Univerze v Ljubljani, katere glavni namen je vzpostavljati in vzdrževati digitalne jezikovne vire in jezikovnotehnološka orodja za sodobni slovenski jezik. Med ciljnim uporabniki izdelkov, ki vključujejo besedilne korpuse, odzivne slovarje in orodja za podporo pisanju v slovenščini, je tudi učiteljska skupnost, ki vire in orodja uporablja na področju jezikovnega izobraževanja. V prispevku pregledno predstavimo tiste vire in orodja CJVT, ki so za izobraževalne namene še posebej relevantni, nato pa umestimo učiteljice in učitelje kot pomembne sodelujoče v dveh kontekstih: pri uporabniških raziskavah in evalvacijah posameznih izdelkov ter pri usposabljanjih za rabo nove digitalne jezikovne infrastrukture pri predmetih, ki jih poučujejo. Prispevek sklenemo z napovedjo rezultatov, ki trenutno nastajajo v nacionalnem projektu Razvoj slovenščine v digitalnem okolju in bodo omogočili nadaljnji razvoj učnih gradiv in izdelkov, relevantnih za izobraževalno rabo.

Ključne besede:

jezikovni
viri,
slovenščina,
uporabniške
raziskave,
izobraževanja
učiteljev,
CJVT

CJVT LANGUAGE RESOURCES AND THEIR USE FOR EDUCATIONAL PURPOSES

ŠPELA ARHAR HOLDT,^{1,2} IZTOK KOSEM^{1,3} & EVA PORI¹

¹ University of Ljubljana, Faculty of Arts, Ljubljana, Slovenia.

E-mail: spela.arharholdt@ff.uni-lj.si, iztok.kosem@cjvt.si, eva.pori@ff.uni-lj.si

² University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia.

E-mail: spela.arharholdt@ff.uni-lj.si

³ Institut »Jožef Stefan«, Ljubljana, Slovenia.

E-mail: iztok.kosem@cjvt.si

Abstract The Centre for Language Resources and Technologies (CJVT) is an interdisciplinary unit of the University of Ljubljana whose main aim is to develop and maintain digital language resources and technologies for contemporary Slovene. Among the target users of products such as corpora, responsive dictionaries and tools that support writing tasks are teachers who use these resources in their language classes. This paper presents those CJVT resources and tools that are particularly relevant for the educational setting, followed by a description of the role of teachers in two different contexts: as subjects in user surveys and evaluations of individual resources, and as participants in training workshops on the use of new digital language infrastructure in the school subjects they teach. The paper concludes with a presentation of the resources that are currently being compiled in the national project Development of Slovene in a Digital Environment, which will facilitate the production of teaching materials and other educationally relevant products.

Keywords:

language
resources,
Slovene,
user
research,
teacher
education,
CJVT

1 Uvod

Jezikovni viri so že od nekdaj pomemben del izobraževalnega procesa, se je pa skozi leta povečala njihova raznovrstnost, tako glede same razpoložljivosti kot uporabe pri poučevanju oz. učenju jezikov. V času tiskanih medijev so jezikovno rabo predstavljali predvsem slovarji in slovnice, že v 90. letih 20. stoletja pa so se v Angliji pojavile tudi prve uporabe korpusnih konkordančnikov pri poučevanju jezika (Johns 1991, 1993). Ravno na področju poučevanja angleščine smo bili v zadnjih 30 letih priča največjemu razmahu v rabi novih metod in pristopov, ki vključujejo jezikovne vire in tehnologije.

V slovenskem prostoru se je raba jezikovnih virov in tehnologij (lahko) uveljavila šele v zadnjem desetletju, ko se je začela izdelava ustrezne (digitalne) infrastrukture, ob tem pa se je pričelo tudi izvajanje spremljajočih izobraževalnih dejavnosti, kot so npr. delavnice za različne uporabniške skupine. Prelomno vlogo pri tem je imel projekt Sporazumevanje v slovenskem jeziku (SSJ),¹ ki je potekal v letih 2008–2013 in sta ga financirala Ministrstvo za izobraževanje, znanost in šport in Evropski socialni sklad. V okviru projekta so namreč nastali številni viri in orodja, ki so močno vplivali na marsikateri danes nepogrešljiv jezikovni vir (npr. Slogovni priročnik na ePravopis, konkordančnik korpusa Gigafida na konkordančnik Gigafida 2.0 ipd.). Pri tem je treba poudariti, da so bili učitelji in učenci ciljna populacija večine projektnih aktivnosti in izdelanih virov.

Na področju izobraževanja učiteljev v delu z jezikovnimi viri in tehnologijami je bil opravljen pomemben korak s projektom Jezikovnotehnološko izpopolnjevanje učiteljev,² ki ga je financiralo Ministrstvo za kulturo in je potekal v letih 2012–2014 v izvedbi Slovenskega društva za jezikovne tehnologije in zavoda Trojina. Projektnih delavnic se je udeležilo več kot 800 učiteljev (večinoma slovenščine, ampak tudi drugih predmetov) z osnovnih in srednjih šol po vseh Sloveniji.

¹ <http://ssj.slovenscina.eu/>

² <http://ucitelji.sdjt.si/>

Nalogo vzdrževanja jezikovnih virov, nastalih v okviru projekta SSJ, izvajanja nadaljnjih izobraževanj ter izdelavo novih virov je prevzel Center za jezikovne vire in tehnologije (CJVT),³ interdisciplinarna raziskovalna enota Univerze v Ljubljani, ki deluje na petih različnih fakultetah. Glavni namen CJVT je znanstveno raziskovanje ter vzpostavljanje in vzdrževanje temeljnih digitalnih jezikovnih virov in jezikovnotehnoloških orodij za sodobni slovenski jezik, ki so praktično uporabni in preko spleta dostopni vsem uporabnikom slovenskega jezika v svetu.

Zagotavljanje uporabnosti virov in orodij CJVT za izobraževalne namene ter sodelovanje z učitelji slovenščine je že od začetka ena temeljnih dejavnosti CJVT. V pričujočem prispevku najprej predstavimo vire CJVT, zlasti tiste, ki se uporabljajo pri poučevanju oz. učenju slovenščine. Nato predstavimo relevantne študije, ki smo jih izvedli z učitelji in osvetljujejo njihove potrebe in navade pri rabi jezikovnih virov. Sledi predstavitev konkretne izkušnje izobraževanja učiteljev, del katerega je bila tudi izdelava vzorčnih učnih priprav za rabo konkretnega jezikovnega vira ali virov pri pouku slovenščine. Prispevek zaključimo s povzetkom glavnih spoznanj in opisom načrtov za prihodnost.

2 Viri CJVT

V nadaljevanju predstavljamo jezikovne vire in tehnologije CJVT, ki so med uporabniki najbolj priljubljeni in so hkrati tudi najrelevantnejši za izobraževalne namene.

Gigafida 2.0⁴ (Krek idr. 2020) je referenčni korpus pisne slovenščine in vsebuje 1,134 milijarde besed. Zajema dnevne časopise, revije, izbrani nabor spletnih besedil (ki tudi v določeni meri pokriva novice) in knjižne publikacije različnih vrst (leposlovje, učbeniki, stvarna literatura), ki tako predstavljajo vzorec sodobne standardne slovenščine. Korpus Gigafida 2.0 je temeljni gradivni vir za izdelavo slovarskih in podobnih virov sodobne slovenščine, bodisi tistih na CJVT ali tistih, ki nastajajo na drugih slovenskih ustanovah.

³ <https://www.cjvt.si/>

⁴ <https://viri.cjvt.si/gigafida/>

Slovar sopomenk sodobne slovenščine⁵ (Krek idr. 2018; Arhar Holdt idr. 2018) je s 105.473 iztočnicami in 368.117 sopomenkami najobsežnejša odprto dostopna zbirka sopomenk za slovenščino. Pomembni edinstveni elementi slovarja so možnost primerjave kolokacijske okolice sopomenk, povezave na korpus Gigafida in odzivnost: slovar je namreč prvi predstavnik t. i. odzivnih jezikovnih virov, za katere je značilno, da uporabnikom omogočajo aktivno sodelovanje pri razvoju vsebine, posledično pa hitro posodabljanje slovarskih podatkov (Arhar Holdt idr. 2018: 404).

Kolokacijski slovar sodobne slovenščine⁶ (Kosem idr. 2018a; Kosem idr. 2018b; Kosem idr. 2019b) je prvi kolokacijski vir za slovenščino in vsebuje 35.989 iztočnic in 7.338.801 kolokacij, pridobljenih iz korpusa Gigafida. Gre za še en odzivni slovar, katerega ključni namen je pomagati uporabnikom pri jezikovni produkciji. Glavni prepoznavni elementi slovarja so stopenjski prikaz gesel (z uporabo ikone piramide in časovnega žiga), predstavitev podatkov v širšem besedilnem kontekstu ter različne možnosti filtriranja in razvrščanja bogatih kolokacijskih podatkov.

Slovenski oblikoslovni leksikon Sloleks 2.0⁷ (Dobrovoljc idr. 2019) je leksikon besednih oblik za slovenski jezik. Različica Sloleks 2.0 obsega 100.802 iztočnici in 2.792.003 posamezne besedne oblike z opisanimi slovničnimi lastnostmi, besednim oblikam pa so z avtomatskimi postopki pripisani tudi naglasi in zapisi v mednarodni fonetični pisavi IPA, ki jih uporabniki lahko ocenjujejo kot ustrezne ali neustrezne. Uporabniki lahko dodajo tudi lasten posnetek izgovora iztočnice. Na ta način Sloleks sledi načelom odzivnega slovarja.

Igra besed (Arhar Holdt idr. 2020b) je jezikovnodidaktična mobilna igra (na voljo za naprave Android in iOS), v kateri se igralci lahko preizkusijo v znanju slovenskega jezika, in sicer poznavanju tipičnih kombinacij besed oz. kolokacij in sopomenk. Podatki temeljijo na bazah Kolokacijskega slovarja in Slovarja sopomenk sodobne slovenščine. Igra, ki je prva tovrstna aplikacija pri nas, je del prizadevanj uporabnikom približati prostočasno utrjevanje in izboljševanje znanja slovenskega jezika, sodelovanje uporabniške skupnosti pa kot del odzivnega koncepta pomaga izboljševati bazi obeh navedenih slovarjev (Arhar Holdt idr. 2020b: 46–47). Na podlagi slovenskega modela so bile v okviru projekta ELEXIS⁸ (Evropska

⁵ <https://viri.cjvt.si/sopomenke/slv/>

⁶ <https://viri.cjvt.si/kolokacije/slv/>

⁷ <https://viri.cjvt.si/sloleks>

⁸ <https://elex.is/>

leksikografska infrastruktura, financira EU v okviru sheme H2020) dodane različice za angleščino, estonščino in nizozemščino, kmalu naj bi sledili še drugi jeziki.

Jezikovni sledilnik 1.0⁹ (Kosem idr. 2021) je nov vir v slovenskem prostoru in prikazuje informacije o časovnih trendih rabe besed in besednih nizov. Sledilnik črpa podatke iz referenčnega korpusa Gigafida 2.0 za besedila do leta 2018 in servisa spletnih novic IJS NewsFeed¹⁰ (besedila od leta 2019 naprej), ki črpa besedila iz več kot 100 različnih slovenskih spletnih virov. Vsaka od besed, predstavjenih v Jezikovnem sledilniku, je opremljena z grafičnim prikazom rabe v izbranem časovnem obdobju, omogočeno pa je tudi primerjanje časovnih trendov več besed. Za vsako od prvih 100 ključnih besed so ponujeni tudi najpogostejši nizi, v katerih se beseda pojavlja. S Sledilnikom tako lahko spremljamo najbolj aktualno besedišče določenega obdobja, pa tudi spremembe in novosti v jeziku.

Novost med orodji CJVT so tudi Vejice 1.0¹¹ (Božič idr. 2020), spletno orodje, ki preverja postavitev vejic v slovenskih besedilih. Orodje, ki je po pričakovanjih naletelo na velik in spodbuden odziv širše javnosti, opozarja na manjkajoče vejice s sivo barvo in na odvečne vejice z modro barvo. Glede na teste program trenutno deluje uspešno v 94 odstotkih primerov.

Poleg izdelave novih jezikovnih virov in orodij CJVT skrbi tudi za vzdrževanje in nadgradnjo več obstoječih za jezikovno izobraževanje relevantnih virov in orodij. Na prvem mestu velja izpostaviti Pedagoški slovnični portal¹² (Arhar idr. 2017), pilotni projekt, ki je nastal v okviru projekta SSJ in je prvi spletni prosto dostopni didaktični vir pri nas. Projekt se posveča zadregam, s katerimi se slovenski šolarji pri pisanju v slovenščini pogosto srečujejo. Pilotni vzorec slovničnih poglavij je pripravljen na podlagi obsežnih analiz primerov avtentične jezikovne rabe iz različnih korpusov. Vsako poglavje vsebuje interaktivno razlago, kratek povzetek, rubriko za navdušene in interaktivne vaje za utrjevanje znanja.

⁹ <https://viri.cjvt.si/sledilnik>

¹⁰ <http://newsfeed.ijs.si/>

¹¹ <https://orodja.cjvt.si/vejice/>

¹² <http://slovnica.slovenscina.eu/>

Pedagoški slovnični portal podatke o pogostih jezikovnih napakah šolarjev črpa iz korpusa Šolar. Zadnja različica Šolar 2.0 (Kosem idr. 2019a) vsebuje 1,7 milijona besed oz. 5.485 besedil, od katerih številna vsebujejo tudi označene in vsebinsko kategorizirane učiteljske jezikovne popravke. Poleg Pedagoškega slovničnega portala so podatke iz korpusa Šolar uporabili številni drugi jezikvnodidaktični priročniki, od Slovnice na kvadrat (2018) ali Kratkoslovnice (2018) Kozme Ahačiča do različnih učbenikov in delovnih zvezkov za slovenski jezik.¹³

Učiteljski oz. izobraževalni skupnosti je namenjeno tudi orodje za analizo berljivosti besedil,¹⁴ ki je bilo izdelano v okviru projekta *Za kakovost slovenskih učbenikov* (KaUč; 2017–2022)¹⁵ in za katerega vzdrževanje in nadaljnji razvoj zdaj skrbi CJVT. Orodje s pomočjo različnih formul za oceno berljivosti in ostalih parametrov (Škvorc idr. 2019) ponudi oceno zahtevnosti vstavljenega besedila ter opozori na morebitne problematične dele besedila (z vidika rabe redkih besed, dolgih stavkov, besed izven učbeniškega besedišča ipd.).

Vsi jezikovni viri in tehnologije, ki nastajajo ali se nadgrajujejo na CJVT, sledijo načelom digitalne sodobnosti, uporabniške prijaznosti in intuitivnosti rabe. To pomeni, da pri predstavitvi jezikovnih podatkov v čim večji meri izkoriščamo prednosti digitalnih medijev, vmesniki virov in orodij pa so zasnovani enostavno in pregledno. V digitalnem svetu je namreč ključno zagotoviti, da uporabniki za uporabo virov ne potrebujejo obsežnih priprav ali usposabljanj. Uporabniška prijaznost virov CJVT je zagotovljena tako, da že pri zasnovi upoštevamo izsledke s področja uporabniških raziskav, ki jih po objavi nadgradimo tudi z evalvacijskimi uporabniškimi študijami. Izsledki evalvacij predstavljajo osnovo za nadaljnje izboljšave izdelkov, s čimer je njihova odzivnost zagotovljena tudi na tej ravni.

3 Učitelji in učiteljice slovenščine kot uporabniška skupina

Ob razvoju koncepta za Slovar sodobne slovenščine (Gorjanc idr. 2015) je bila prvič oblikovana shema, v kateri so natančneje opredeljene ciljne skupine uporabnikov sodobnih jezikovnih virov za slovenščino (Gorjanc idr. 2015: 143). V tej shemi učitelji/-ce tako slovenščine kot prvega in drugega/tujega jezika kot tudi učitelji/-ce

¹³ Podatki so bili uporabljeni tudi pri pripravi učnega e-okolja Slovenščina na dlani.

¹⁴ <https://orodja.cjvt.si/berljivost/>

¹⁵ <https://kauc.splet.arnes.si/>

drugih predmetov zasedajo pomembno mesto. Od takrat naprej jih na CJVT redno vključujemo v uporabniške študije, tj. raziskave, ki so namenjene identifikaciji uporabniških mnenj, težav, želja, potreb in zmožnosti. V zadnjih letih smo izvedli tako vsebinsko širše raziskave različnih vprašanj, povezanih z jezikovno opremljenostjo, kot tudi študije, usmerjene ciljno k določenemu jezikovnemu viru. Nekaj primerov študij in izsledkov, ki jih uporabljamo za načrtovanje nadaljnjega razvoja izdelkov, konceptov in rešitev, predstavljamo v nadaljevanju.

Ob izidu Slovarja sopomenk sodobne slovenščine smo raziskali, kako skupnost sprejema novosti, ki jih prinaša odzivni slovar, med katerimi so strojna priprava in izključno digitalni format slovarja, vključenost kolokacijskih podatkov, korpusnih zgledov in povezav na korpus ter možnosti za uporabniško sodelovanje pri razvoju slovarja. Anketa, ki jo je izpolnilo 671 sodelujočih, je pokazala, da so slednji novostim naklonjeni, opredelili pa so tudi nekaj predlogov za izboljšave, npr. dodatne iztočnice (tudi nestandardne), dodatek kvalifikatorjev k zaznamovanim iztočnicam in sopomenkam, nekatere izboljšave vmesnika in večjo povezljivost z drugimi jezikovnimi viri. Statistične primerjave so pokazale, da je v primerjavi z drugimi uporabniškimi skupinami osnovnošolske učitelje [N=120] redkeje motilo, da je slovar v prvem koraku pripravljen izključno avtomatsko, srednješolske učitelje [N=75] pa je pogosteje motilo, da slovar vsebuje zgolj standardno in sodobno besedišče ter da ne vsebuje kvalifikatorjev, ki bi tovrstno besedišče lahko ustrezno opredeljevali. Raziskava in izsledki so popisani v prispevku Arhar Holdt (2020).

Študija uporabniškega odnosa do podatkov v Kolokacijskem slovarju sodobne slovenščine, ki jo predstavlja prispevek Pori idr. (2020), je bila kvalitativne narave: sodelujoči [N=40] so raziskali strojno pripravljene podatke v izbranem naboru slovarskih iztočnic in ob tem poročali o pozitivnih in negativnih vtisih, nato pa so z vodjo raziskave opravili intervju o uporabnosti slovarja. V raziskavo smo vključili (tudi) 10 učiteljev slovenščine kot prvega in 10 kot drugega/tujega jezika. Posebej nas je zanimal njihov odnos do dejstva, da avtomatsko izluščeni podatki vsebujejo določeno mero šuma. Predvidevali smo, da so tovrstne napake za učne situacije lahko zelo moteče in bodo vodile v učiteljski odpor do uporabe slovarja. Učitelji so zadržke potrdili, vendar so obenem poudarili, da je ob premišljeni rabi slovar izredno uporaben in v veliko pomoč pri pripravi jezikovnih vaj, gradiv za usvajanje besedišča, kot podpora pisanju in utrjevanju jezikoslovne terminologije. Kot močno točko so

izpostavili avtentičnost vključenega jezika, povezljivost med viri in s tem možnost opazovanja jezika v širšem jezikovnem kontekstu ter različnih žanrskih okoljih.

Posebej se trudimo, da slovenske učitelje vključujemo oz. vabimo k sodelovanju tudi v mednarodnih raziskavah. V nadaljevanju navajamo dve evropski raziskavi, ki sta nastali pod okriljem evropskih mrež COST, prva v sodelovanju Evropske mreže za e-leksikografijo (ENeL) in druga v sodelovanju Evropske mreže za povezovanje jezikovnega učenja s tehnikami množičenja (EnetCollect).¹⁶

Leta 2017 je potekala evropska raziskava o odnosu jezikovnih uporabnikov do splošnih enojezičnih slovarjev (Kosem idr. 2019). Raziskava je zajela 9.099 sodelujočih, med njimi je bilo 2.457 (27 %) učiteljev jezika. Sodelujočih iz Slovenije je bilo 619, od tega 172 učiteljev jezika (28 %). Analiza slovenskega dela odgovorov (Arhar Holdt 2018) je pokazala, da sodelujoči do slovarskih podatkov najpogosteje in najraje dostopajo s pomočjo računalnika. Slovar se tipično uporablja za preverbo ustreznosti jezikovnih izbir (lastnih in tujih) in za tvorbo besedil v službenem in študijskem okolju. Simbolna vloga slovarja se v odgovorih izpostavlja redkeje, posredno pa jo nakazuje nenaklonjenost sodelujočih, da bi knjižni slovar oddali ali zavrgli, četudi uporabljajo le digitalno različico. Med vprašanimi vlada močna podpora javnemu financiranju priprave splošnih enojezičnih slovarjev, temu ustrezno pa med prioritetai navajajo brezplačnost slovarja. Primerjava slovenskih odgovorov z ostalimi državami (Kosem idr. 2018) je pokazala veliko podobnosti, pa tudi nekatere pomembne razlike. Slovenski rezultati so denimo izkazali največji razkorak med (takratno) ceno splošnega enojezičnega slovarja in odgovori, koliko so sodelujoči pripravljene za tovrstni slovar plačati.

Leta 2018 je potekala mednarodna raziskava o tem, koliko jezikovni učitelji uporabljajo množičenje (ang. *crowdsourcing*)¹⁷ in kakšno je njihovo mnenje o vključevanju množičenja v delo z učenci/dijaki, kot tudi lastne strokovne aktivnosti (Arhar Holdt idr. 2020a). Med 1.129 sodelujočimi je bilo tudi 40 slovenskih učiteljev. Raziskava je potrdila predvidevanja, da veliko sodelujočih ni seznanjenih s konceptom množičenja, vendar so ideji načeloma naklonjeni: večina je izrazila

¹⁶ Spletni strani evropskih mrež: <https://www.elxicography.eu/> in <https://enetcollect.eurac.edu/>.

¹⁷ Množičenje je organizirano, pogosto prostovoljno sodelovanje širše skupnosti za doseg določenega (skupnega, družbeno relevantnega) cilja, tipično v digitalnem okolju. Tipičen primer množičenja je priprava Wikipedije. Na področju jezikovnega učenja se lahko množičenje uporablja za pripravo sodobnih, odprto dostopnih učnih gradiv ter jezikovnih virov.

pripravljenost sodelovati pri množičenju za jezik, ki ga poučujejo, zlasti pri sodelovalnem razvoju učnih gradiv ter urejanju jezikovnih zgledov za vaje in teste. Za sodelovanje bi jih zlasti motivirale konkretne ideje, kako uporabiti množičenje pri šolskih aktivnostih, in usposabljanja o učnih metodah, ki vključujejo množičenje.

4 Učitelji in učiteljice slovenščine kot naslovniki izobraževanj in usposabljanj

Pomemben delež ciljnih naslovnikov izobraževanj in usposabljanj predstavljajo učitelji/-ce slovenščine kot prvega in drugega/tujega jezika, pa tudi učitelji/-ce drugih jezikoslovnih predmetov. Ti izpostavljajo zanimanje in potrebo po razvoju prosto dostopnih učnih e-okolij in po seznanjanju z digitalnimi jezikovnimi viri, saj lahko na ta način učinkovito osvežujejo in posodablajo fleksibilni pedagoško-didaktični proces. Pri vpeljavi jezikovnih virov in orodij v pouk jezika in razvoju inovativnih metodološko-didaktičnih pristopov pa jih je treba z novimi digitalnimi jezikovnimi orodji in viri ter inovacijami, ki jih slednji prinašajo, najprej seznaniti in jim omogočiti oz. nuditi jezikovnotehnoško podporo. Nove oblike uresničevanja predpisanih učnih ciljev in vsebin ter didaktičnih pristopov na področju jezika namreč med učitelji še niso dobro poznane, zato so izobraževanja in kreativni razmisleki o vključitvi v jezikovni pouk v sodelovanju z učitelji nujni.

Na CJVT smo se usmerili v izvajanje izobraževalnih seminarjev, ki predstavljajo novosti na področju disciplinarnega znanja, spodbujanja in razvijanja odprtosti do uporabe aktivnih metod poučevanja, IKT-tehnik učenja, novih didaktičnih pripomočkov, uporabe različnih virov učenja in vrednotenja učenja (učnih dosežkov) pri pouku slovenskega jezika in književnosti. S tem pristopom in oblikami dela sledimo aktualnim smernicam nadaljnega usposabljanja in izobraževanja pedagoških delavcev ter sodobnim priporočilom za poučevanje, učenje in vrednotenje v vzgojno-izobraževalnem procesu. S pedagoškimi seminarji in delavnicami oz. s pedagoškim programom, ki se umešča v Katalog nadaljnega izobraževanja in usposabljanja strokovnih delavcev v vzgoji in izobraževanju (KATIS),¹⁸ učitelje usposabljammo za uporabo didaktičnih pristopov, s katerimi lahko na ustrezen način znanja in metode učenja učinkovito prenesejo na učence/dijake. Usposabljammo jih za nove, drugačne oblike dela v razredu in jim nudimo podporo

¹⁸ <https://paka3.mss.edus.si/katis/Uvodna.aspx>

pri inovativnem uresničevanju ciljev, vsebin in didaktičnih priporočil za jezikovni pouk, navedenih v učnem načrtu za slovenščino.

Pri tako specifični skupini uporabnikov, kot so učitelji, je še posebej pomembno, da fokus izobraževanja usmerimo v seznanjanje s konkretnimi možnostmi rabe v razredu in s smernicami za samostojno uporabo pri pouku. Posebno pozornost usmerjamo v usklajenost gradiva z učnim načrtom, predvsem na ravni ciljev in vsebin (razporejenost vsebin in ciljev po razredih ali letnikih) za slovenščino, ter v pripravo kazalnikov za zagotavljanje kakovosti nadaljnjega učnega procesa. Smernice in priporočila za načrtovanje in izvedbo učne ure z odzivnimi jezikovnimi viri ustrezno prilagajamo glede na raven in stopnjo izobraževanja; za osnovne šole, gimnazije in druge (strokovne, poklicne) srednje šole. Učiteljem nudimo podporo pri pripravi pedagoških scenarijev (teoretični zasnovi in praktični izvedbi učnih ur), predstavimo jim teoretična izhodišča (seznanitev z viri in njihovimi funkcijami ...) in možna nova didaktična izhodišča za praktično izvedbo, pri čemer naslavljamo konkretna vprašanja, ki se navezujejo na: učni načrt, učne cilje, (inovativne) metodično-didaktične pristope k delu. Usmeritev na konkretne primere rabe, ki se umeščajo v konkretne tematske sklope učnega načrta oz. učnih ciljev in vsebin za slovenščino, učiteljem omogoča lažji prenos znanj v prakso.

Izobraževalno (predstavitveno) gradivo za učitelje pripravimo na način, da ga kasneje lahko uporabijo tudi pri neposrednem delu v razredu, npr. oblikujemo učne liste, navodila za delo s posameznim virom v obliki videoposnetkov, ki bodo zanimivi tudi za učence ipd. Učiteljem predstavimo nove možnosti, ki jih viri prinašajo, npr. možnost sodelovanja skupnosti pri razvoju slovarjev, odzivnost in možnost igrifikacije vsebin pri pouku slovenščine (kar je trenutno redka praksa), konkretne možnosti uporabe.


cjvt
UPORABA KOLOKACIJSKEGA SLOVARJA SODOBNE SLOVENŠČINE
DEJAVNOSTI:

- primerjava gesel (pomenske členitve) z drugimi slovarskimi viri
- razvrščanje kolokacij pod pomene
- določanje pomenov (napredno)
- tematsko razvrščanje kolokatorjev
- prepoznavanje kolokacijskega jedra na podlagi kolokatorjev
- prepoznavanje sinonimnih parov s pomočjo kolokacij
- predstavljanje pomena danih besed iz besedila
- prepoznavanje, utemeljevanje in odpravljanje napak
- ocenjevanje kolokacij
- naštevanje najbolj tipičnih kolokacij za dano besedo; [Igra besed](#)

CILJI:

- razvijanje metajezikovnih zmožnosti:
 - spoznajo pojme: beseda, pomen in izrazna podoba besed, tematsko polje, metaforična raba besed, slovar, jezikovni korpus
 - usvojijo pojme: enopomenke, večpomenke, besedne zveze, kolokacije, spletni slovar, jezikovni korpus, besedilni korpus, sopomenke, protipomenke, enakozvočnice, enakopisnice, enakoglasnice, nadpomenke, podpomenke, stalna besedna zveza, frazem

10x

Slika 1: Primeri dejavnosti z uporabo Kolokacijskega slovarja sodobne slovenščine, ki se umeščajo v konkretne tematske sklope učnega načrta in učnih ciljev za slovenščino.

Vir: lasten

Smiselno je, da s pomočjo učiteljev načrtujemo samostojno (akcijsko) raziskovanje in tudi implementacijo v razredu, za kar je potrebna predstavitev izkušnje in usmerjena evalvacija. Ponudimo jim smernice za izvedbo evalvacije, v sklopu refleksije jih usmerimo predvsem v (dosedanja) spoznanja in ugotovitve. Povabimo jih k razmisleku o razvoju njihovih kompetenc in kompetenc učencev. Na ravni učiteljske izkušnje nas predvsem zanima, katere nove metodično-didaktične pristope oz. možnosti za podajanje snovi so razvijali in ali so novi učni pristopi učinkovito pripomogli k uresničevanju učnih ciljev v skladu z vsebinami in učnim načrtom za slovenščino. Nadalje pa tudi, katere prednosti ali omejitvene dejavnike so zaznavali po (nekajmesečni, nekajkratni) kontinuirani uporabi digitalnih jezikovnih virov pri pouku.

Na ravni učenceve izkušnje pa je smiselno premisliti, kako se učenci odzivajo na različne možnosti načinov vpeljave in uporabe digitalnih virov pri pouku slovenščine in ali se nad novimi možnostmi podajanja snovi navdušujejo, kakšne so njihove potrebe, pričakovanja, odzivi nasploh – kako vzpostavljajo odnos do sodobnih digitalnih jezikovnih virov, ali prepoznajo, kaj jim ponujajo, kakšen je njihov odnos do digitalnega jezikovnega okolja. Evalvacija oz. skupna razprava o uporabnosti obravnavanih virov pri pouku slovenščine, vključno z evidentiranjem možnih pomislekov in zadržkov, je ključnega pomena za vrednotenje stanja in napredka. Učitelji testirane primere rabe v razredu reflektirajo in evalvirajo kot možne primere dobre prakse za nadaljnje poučevanje in vrednotenje. Pomembno pa je tudi, da so izdelani in preizkušeni primeri učnih ur prosto dostopni, rezultati pa po izobraževanju diseminirani in na voljo tudi za druge učitelje, ki bi želeli gradivo uporabiti (mreženje znanja in izkustvenega učenja).

Evalvacija uporabe novih jezikovnih virov pri pouku slovenščine je pokazala, da so odzivi učiteljev (po implementaciji učne ure v razredu) pozitivni. Učitelji menijo, da novi jezikovni viri prinašajo nove jezikovne podatke (na ravni sopomenskosti, kolokacijskosti, besednih naglasov itd.), ki ponujajo nove možnosti za podajanje snovi in učinkovito pripomorejo k uresničevanju učnih ciljev v skladu z vsebinami in učnim načrtom za pouk slovenskega jezika v srednji šoli:

»Možnosti uporabe je veliko ne samo pri jezikovnem pouku in tvorjenju zapisanih besedil, ampak tudi pri tvorjenju govornih besedil, saj s pomočjo Sopomenk in Kolokacij dijaki besedno izboljšujejo svoja besedila.« (učiteljica slovenščine na gimnaziji)

Nad novim načinom podajanja snovi pri urah slovenščine pa se navdušujejo tudi učenci oz. dijaki:

»Dijaki so večinoma zelo navdušeni za delo z digitalnimi viri in orodji, saj spoznavajo njihove prednosti, zato jih znanjo samostojno uporabljati ne samo pri pouku, ampak tudi v vsakdanjem življenju. S spoznavanjem in uporabo aplikacije Igra besed pa postanejo tudi del preživljanja prostega časa.« (učiteljica slovenščine na gimnaziji)

»Dijaki se nad rabo digitalnih virov in orodij navdušujejo, hkrati pa se razvijajo v samozavestnega bralca, poslušalca oz. govorca in pisca umetnostnih in neumetnostnih besedil ne samo pri slovenščini, ampak pri vseh predmetih.« (učiteljica slovenščine na gimnaziji)

Učitelji tudi opozarjajo na pomembnost kontinuiranega dela z novostmi in nujnost posodabljanja ustaljene pedagoške prakse, kar bi se moralo odražati tudi s premiki in posodobitvami v učnem načrtu:

»Mislim, da bi bilo seznanjanje z digitalnimi viri in tehnologijami nujno vnesti kot obvezen del v učni načrt za slovenščino v OŠ in SŠ; njihovo uporabo pa bi morali poznati tako učitelji slovenščine kot drugih predmetov.« (učiteljica slovenščine na gimnaziji)

Učiteljske izkušnje in povratna informacija kažejo, da izobraževanje o digitalnih, odzivnih jezikovnih virih predstavlja različne možnosti načinov vpeljave in uporabe digitalnih virov pri pouku slovenščine (pri sprejemanju, razčlenjevanju in tvorjenju neumetnostnih in umetnostnih besedil; kot podpora kritičnemu mišljenju, ustvarjalnosti in inovativnosti; za iskanje, zbiranje, izmenjavo in obdelavo podatkov ter njihovo sistematično rabo pri tvorjenju informacij) in omogoča nove metodično-didaktične aktivnosti, med drugim tudi sodelovanje tako učiteljev kot dijakov in učencev pri razvoju jezikovnih virov za slovenščino. Viri, temelječi na avtentični jezikovni rabi, spodbujajo razumevanje jezika kot pojava, ki nastaja v skupnosti in se z njeno pomočjo tudi razvija, obenem pa zavedanje o pomenu odprtih jezikovnih podatkov za razvoj in polnomočno življenje slovenščine v digitalnem svetu.

Izkušnja z izvedbo izobraževanja po predstavljenem modelu je razkrila, da je prednost tako zasnovanih programov v tem, da vplivajo na celovit razvoj pedagoških digitalnih kompetenc. Pedagoške delavce usposablja za (a) izdelavo, ustvarjanje, posodabljanje in objavo gradiv, s katerimi omogočijo vključenim v vzgojno-izobraževalni proces sodelovalno delo: usmerjeni so v izdelavo in objavo gradiv na temo novih jezikovnih virov in njihove rabe za reševanje jezikovnih vprašanj v sodobni slovenščini, kar omogoča dve vrsti sodelovanja: sodelovanje učiteljev pri nadaljnji rabi in razvoju gradiv; sodelovanje tako učiteljev kot učencev pri nadaljnjem razvoju jezikovnih virov za slovenščino in (b) za načrtovanje, izvedbo in evalvacijo pouka z uporabo IKT: v sklopu eksperimentalnega dela (izdelave konkretnega primera rabe vira v razredu za izbrane učne vsebine oz. priprave učne ure, ki vključuje nove jezikovne vire in njihovo testiranje v razredu) oz. implementacije v razredu, ki ji sledi evalvacija rabe novih jezikovnih virov pri pouku slovenščine. Učitelji pri tem razvijajo pedagoške kompetence, ki jih naslavlja tudi razvojna strategija MIZŠ, tj. usposabljanje učiteljev za načrtovanje, izvedbo in evalvacijo pouka z uporabo IKT.

5 Sklep in bodoče delo

Jezikovni viri CJVT prinašajo nove jezikovne podatke (na ravni sopomenskosti, kolokacijskosti, besednih naglasov itd.), ki ponujajo nove možnosti za podajanje snovi in učinkovito pripomorejo k uresničevanju učnih ciljev v skladu z vsebinami in učnim načrtom za pouk slovenskega jezika. Novosti, ki jih vpeljuje strojno podprta obdelava velike količine jezikovnih podatkov, vodijo do novih oblikovnih in tudi konceptualnih rešitev, med katerimi so odzivni viri, igrifikacija, orodja v podporo pisni produkciji (postavljanje vejic, ocenjevanje zahtevnosti besedila) in podobno. Vsi izdelki temeljijo na uporabniških raziskavah in evalvacijah, v katere je v največji možni meri vključena tudi učiteljska skupnost.

Z izobraževanju naslavljamo potrebo strokovnih delavcev po seznanjenosti z jezikovnimi priročniki, ki prinašajo sodoben, avtentičen jezik, natančneje s strokovno potrebo učiteljev slovenščine, da imajo pregled nad jezikovnimi viri, njihovim namenom rabe, prednostmi in omejitvami. Opremljeni s tovrstnimi znanji lahko učitelji bolje opolnomočijo učence za reševanje jezikovnih problemov in zagat in prek izboljšave lastne digitalne pismenosti in razgledanosti po razpoložljivih podatkih, primerljive kompetence omogočijo tudi učencem. Prenos znanj spodbuja dejstvo, da je velik del izobraževanj namenjen iskanju rešitev, ki bodo v poučevalni praksi lahko zaživele, vključno z daljšim časom, namenjenim debati o pomislekih, možnih težavah in zadržkih.

Od tekočih projektov in iniciativ, ki so relevantne za učiteljsko skupnost, gre omeniti predvsem projekt Razvoj slovenščine v digitalnem okolju,¹⁹ ki ga med leti 2020 in 2022 financirata Ministrstvo za kulturo in Evropski sklad za regionalni razvoj. Projekt razvija jezikovne vire, govorne in semantične tehnologije, strojno prevajanje in pripravo terminološkega portala. Med napovedanimi jezikovnimi tehnologijami in viri gre posebej izpostaviti projektno aktivnost, ki se posveča korpusu šolskih besedil Šolar in korpusu usvajanja slovenščine kot drugega jezika KOST (Stritar Kučuk 2020). Oba korpusa prinašata avtentično produkcijo učencev, dijakov oz. tečajnikov, pa tudi učiteljske jezikovne popravke.

¹⁹ <https://www.slovenscina.eu/>

Ker je priprava tovrstnih korpusov izredno zamudna, želimo na projektu razviti delotoke in orodja za redno zbiranje gradiva. Pripravljeno bo spletno mesto, kjer bodo na voljo vse informacije (namen gradnje korpusa, pravna soglasja za učence in šolo, informacije o točkah za napredovanje, ki jih prinaša projektno sodelovanje) in kjer bo besedila mogoče tudi hitro in pregledno oddati. Na drugi strani bomo lokalizirali in prilagodili zmogljivi program za označevanje jezikovnih napak in popravkov v šolskih besedilih Svala (Mats idr. 2019). Uredili in standardizirali bomo format korpusov z vpisanimi jezikovnimi popravki in objavili posodobljeno različico korpusa Šolar ter prvo javno dostopno različico korpusa KOST. Na ta način bodo empirični podatki o šolski pisni produkciji lažje dostopni in bolj ažurni, kar bo omogočilo nadaljnji razvoj učnih gradiv in izdelkov, relevantnih za izobraževalno rabo.

Zahvala

Raziskovalni program št. P6-0411 je sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije iz državnega proračuna.

Literatura

- Kozma AHAČIČ, 2018: *Kratkoslovnica: slovenska slovnica za osnovno šolo*. Ljubljana: Rokus Klett.
- Kozma AHAČIČ, 2018: *Slovnica na kvadrat: slovenska slovnica za srednjo šolo*. Ljubljana: Rokus Klett.
- Špela ARHAR HOLDT, Iztok KOSEM in Polona GANTAR, 2017: Corpus-based resources for L1 teaching: the case of Slovene. *Handbook on digital learning for K-12 schools*. Ur. Ann Marcus-Quinn in Triona Hourigan. 91–113.
- Špela ARHAR HOLDT, 2018: Odnos jezikovnih uporabnikov do splošnih enojezičnih slovarjev: slovenski vidik. *Slovenščina 2.0* 6/1, 1–36. Dostop 8. 6. na <https://doi.org/10.4312/slo2.0.2018.1.1-36>
- Špela ARHAR HOLDT, Jaka ČIBEJ, Kaja DOBROVOLJC, Apolonija GANTAR, Vojko GORJANC, Bojan KLEMENC, Iztok KOSEM, Simon KREK, Cyprian LASKOWSKI in Marko ROBNIK ŠIKONJA, 2018: Thesaurus of Modern Slovene: By the Community for the Community. *Proceedings of the XVIII EURALEX International Congress: Lexicography in Global Contexts*. Ur. Jaka Čibej, Vojko Gorjanc, Iztok Kosem in Simon Krek. Ljubljana: Znanstvena založba Filozofske fakultete. 401–410. Dostop 8. 6. 2021 na <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/118/211/3000-1>
- Špela ARHAR HOLDT, 2020: How users responded to a responsive dictionary: the case of the Thesaurus of Modern Slovene. *Rasprave Inštituta za hrvatski jezik i jezikoslovje* 46/2, 465–482. Dostop 3. 6. 2021 na <https://doi.org/10.31724/rihjj.46.2.1>
- Špela ARHAR HOLDT idr., 2020a: Language teachers and crowdsourcing: insights from a cross-European survey. *Rasprave Inštituta za hrvatski jezik i jezikoslovje* 46/1, 1–28. Dostop 5. 6. 2021 na <https://doi.org/10.31724/rihjj.46.1.1>
- Špela ARHAR HOLDT, Nataša LOGAR, Eva PORI in Iztok KOSEM, 2020b: "Game of Words": play the game, clean the database. *Lexicography for inclusion: EURALEX XIX: Congress of the European Association for Lexicography*. Ur. Zoe Gavrilidou, Maria Mitsiaki in Asimakis Flaitouras. 41–49. Dostop 31. 5. 2021 na

- https://euralex2020.gr/wp-content/uploads/2020/11/EURALEX2020_ProceedingsBook-p041-049.pdf
- Martin BOŽIČ, Marko ROBNIK ŠIKONJA, Simon KREK, Špela ARHAR HOLDT, Iztok KOSEM, Polona GANTAR, Jaka ČIBEJ, Kaja DOBROVOLJC, Bojan KLEMENC in Cyprian Adam LASKOWSKI, 2020: *Vejice 1.0: orodje za strojno postavljanje vejic* (Orodja CJVT). Ljubljana: Center za jezikovne vire in tehnologije Univerze v Ljubljani. Dostop 8. 6. 2021 na <https://orodja.cjvt.si/vejice/home>
- Kaja DOBROVOLJC, Simon KREK, Peter HOLOZAN, Tomaž ERJAVEC, Miro ROMIH, Špela ARHAR HOLDT, Jaka ČIBEJ, Luka KRŠNIK in Marko ROBNIK ŠIKONJA, 2019: *Morphological lexicon Sloleks 2.0*. Ljubljana: Centre for Language Resources and Technologies, University of Ljubljana. Dostop 9. 6. 2021 na <https://www.clarin.si/repository/xmlui/handle/11356/1230>
- Vojko GORJANC (ur.) idr., 2015: *Slovar sodobne slovenščine: problemi in rešitve*. Prevodoslovje in uporabno jezikoslovje. Ljubljana: Znanstvena založba Filozofske fakultete.
- Tim JOHNS, 1991: From printout to handout: grammar and vocabulary teaching in the context of data-driven learning. Classroom Concordancing. *English Language Research Journal* 4, 27–45.
- Tim JOHNS, 1993: Data-driven learning: an update. *TELL&CALL* 2, 4–10.
- Iztok KOSEM, Simon KREK, Polona GANTAR, Špela ARHAR HOLDT, Jaka ČIBEJ in Cyprian Adam LASKOWSKI, 2018a: Collocations dictionary of modern Slovene. *Proceedings of the 18th EURALEX International Congress*. Ljubljana: Ljubljana University Press, Faculty of Arts. 989–997. Dostop 8. 6. 2021 na <https://e-knjige.ff.uni-lj.si/znanstvena-zalozba/catalog/view/118/211/3000-1>
- Iztok KOSEM, Sascha WOLFER, Robert LEW in Carolin MÜLLER-SPITZER, 2018b: Attitudes of Slovenian language users towards general monolingual dictionaries: an international perspective. *Slovenščina 2.0* 6/1, 90–134. Dostop 8. 6. 2021 na <https://doi.org/10.4312/slo2.0.2018.1.90-134>
- Iztok KOSEM, Špela ARHAR HOLDT, Nikola LJUBEŠIČ idr., 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. Dostop 3. 6. 2021 na <https://doi.org/10.1093/ijl/icy022>
- Iztok KOSEM, Špela ARHAR HOLDT, Mojca STRITAR KUČUK, Simon KREK, Irena KRAPŠ VODOPIVEC, Marko STABEJ, Eva PORI, Teja GOLI, Polona LAVRIČ, Cyprian Adam LASKOWSKI, Polona KOCJANČIČ, Bojan KLEMENC in Tadeja ROZMAN, 2019a: *Developmental corpus Šolar 2.0* (CLARIN.SI data & tools). Ljubljana: Trojina, Institute for Applied Slovene Studies: University of Ljubljana, Centre for Language Resources and Technologies. Dostop 8. 6. 2021 na <https://www.clarin.si/repository/xmlui/handle/11356/1214>
- Iztok KOSEM, Polona GANTAR, Simon KREK, Špela ARHAR HOLDT, Jaka ČIBEJ, Cyprian Adam LASKOWSKI, Eva PORI, Bojan KLEMENC, Kaja DOBROVOLJC, Vojko GORJANC in Nikola LJUBEŠIČ, 2019b: *Collocations dictionary of modern Slovene KSSS 1.0* (CLARIN.SI data & tools). Ljubljana: Centre for Language Resources and Technologies, University of Ljubljana. Dostop 8. 6. 2021 na <https://www.clarin.si/repository/xmlui/handle/11356/1250>
- Iztok KOSEM, Jaka ČIBEJ, Polona GANTAR, Špela ARHAR HOLDT, Simon KREK, Cyprian Adam LASKOWSKI, Marko ROBNIK ŠIKONJA, Bojan KLEMENC, Kaja DOBROVOLJC, Vojko GORJANC, Andraž REPAR in Nikola LJUBEŠIČ, 2021: *Sledilnik 1.0: jezikovni sledilnik* (Orodja CJVT). Ljubljana: Center za jezikovne vire in tehnologije, Univerze v Ljubljani. Dostop 8. 6. 2021 na <https://viri.cjvt.si/sledilnik/slv/>
- Simon KREK, Cyprian Adam LASKOWSKI, Marko ROBNIK ŠIKONJA, Iztok KOSEM, Špela ARHAR HOLDT, Polona GANTAR, Jaka ČIBEJ, Vojko GORJANC, Bojan KLEMENC in Kaja DOBROVOLJC, 2018: *Thesaurus of modern Slovene 1.0*. Ljubljana: Slovenian Language Resource Repository CLARIN.SI. Dostop 8. 6. 2021 na <http://hdl.handle.net/11356/1166>
- Simon KREK, Špela ARHAR HOLDT, Tomaž ERJAVEC, Jaka ČIBEJ, Andraž REPAR, Polona GANTAR, Nikola LJUBEŠIČ, Iztok KOSEM in Kaja DOBROVOLJC, 2020: Gigafida 2.0:

- the reference corpus of written standard Slovene. *LREC 2020: Twelfth International Conference on Language Resources and Evaluation*. Paris: ELRA – European Language Resources Association. 3340–3345. Dostop 8. 6. 2021 na <http://www.lrec-conf.org/proceedings/lrec2020/LREC-2020.pdf>
- Wirén MATS, Arild MATSSON, Dan ROSEN in Elena VOLODINA, 2019: SVALA: Annotation of Second-Language Learner Text Based on Mostly Automatic Alignment of Parallel Corpora. *CLARIN-2018 post-conference volume*. LiUP Press. Dostop 8. 6. 2021 na <https://ep.liu.se/ecp/159/023/ecp18159023.pdf>
- Eva PORI, Jaka ČIBEJ, Špela ARHAR HOLDT in Iztok KOSEM, 2020: The attitude of dictionary users towards automatically extracted collocation data: a user study. *Slovenščina 2.0* 8/2, 168–201. Dostop 3. 6. 2021 na <https://doi.org/10.4312/slo2.0.2020.2.168-201>
- Mojca STRITAR KUČUK, 2020: Modul Leto plus – prvi korak do korpusa slovenščine kot tujega jezika. *Zbornik konference Jezikovne tehnologije in digitalna humanistika*. Ur. Darja Fišer in Tomaž Erjavec. Ljubljana: Inštitut za novejšo zgodovino. 131–135. Dostop 9. 6. 2021 na http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_StritarKucuk_Modul-Leto-plus%e2%80%93prvi-korak-do-korpusa-slovenscine-kot-tujega-jezika.pdf
- Tadej ŠKVORC, Simon KREK, Senja POLLAK, Špela ARHAR HOLDT in Marko ROBNIK ŠIKONJA, 2019: Predicting Slovene text complexity using readability measures. *Prispevki za novejšo zgodovino* 59/1, 198–220. Dostop 8. 6. 2021 na <http://ojs.inz.si/pnz/article/view/323>