



# Slovenščina 2.0

Let. 8 (2020), št. 1

## **Slovenščina 2.0**

Letnik/Volume 8, Številka/Issue 1, 2020

ISSN: 2335-2736

GLAVNA UREDNIKA/EDITORS-IN-CHIEF

Špela Arhar Holdt, Vojko Gorjanc

UREDNIŠKI ODBOR/EDITORIAL BOARD

Zoran Bosnić, Simon Dobrišek, Tomaž Erjavec, Ina Ferbežar, Darja Fišer,  
Polona Gantar, Peter Jurgec, Iztok Kosem, Simon Krek, Nina Ledinek,  
Nikola Ljubešić, Nataša Logar, Karmen Pižorn, Damjan Popič, Marko Robnik Šikonja, Amanda  
Saksida, Irena Srdanović, Mojca Šorn, Darinka Verdonik, Špela Vintar

TEHNIČNA UREDNICA/MANAGING EDITOR

Eva Pori

PRELOM/LAYOUT

Jure Preglau

ZALOŽILA/PUBLISHED BY

Znanstvena založba Filozofske fakultete Univerze v Ljubljani

IZDAL/ISSUED BY

Center za jezikovne vire in tehnologije Univerze v Ljubljani

ZA ZALOŽBO/FOR THE PUBLISHER

Roman Kuhar, dekan Filozofske fakultete

Publikacija je brezplačna./Publication is free of charge.

Publikacija je dostopna na/Avaliable at: dostopna na: <https://revije.ff.uni-lj.si/slovenscina2/index>

Revija izhaja s podporo Javne agencije za raziskovalno dejavnost Republike Slovenije./  
This journal is published with the support of the Slovenian Research Agency (ARRS).



To delo je ponujeno pod licenco Creative Commons Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna licenca (izjema so fotografije). / This work is licensed under a Creative Commons Attribution-ShareAlike 4.0 International License (except photographs).

## KAZALO

<b>Citiranje jezikovnih podatkov v slovenskih znanstvenih objavah v obdobju 2013–2019</b>	<b>1</b>
Jakob LENARDIČ, Tomaž ERJAVEC, Darja FIŠER	
<b>Czech gender linguistics: topics, attitudes, perspectives</b>	<b>35</b>
Vít KOLEK, Jana VALDROVÁ	
<b>Threat perception in online anti-migrant speech: A Slovene case study</b>	<b>66</b>
Rok CHITRAKAR	
<b>Okrogla miza »(Bližnja) srečanja oblikovalcev jezikovne politike«</b>	<b>92</b>
Ina FERBEŽAR, Igor CETINA, Alojz IHAN, Marko STABEJ, Lana ZDRAVKOVIČ in Tina ZUPANČIČ	
<b>Konferenca Jezikovne tehnologije in digitalna humanistika 2020</b>	<b>113</b>
Katja MEDEN	

# **CITIRANJE JEZIKOVNIH PODATKOV V SLOVENSКИH ZNANSTVENIH OBJAVAH V OBDOBJU 2013–2019**

**Jakob LENARDIČ**

Univerza v Ljubljani, Filozofska fakulteta

**Tomaž ERJAVEC**

Institut Jožef Stefan

**Darja FIŠER**

Univerza v Ljubljani, Filozofska fakulteta, Institut Jožef Stefan

*Lenardič, J., Erjavec, T., Fišer, D. (2020): Citiranje jezikovnih podatkov v slovenskih znanstvenih objavah v obdobju 2013–2019. Slovenščina 2.0, 8(1): 1–34.*

DOI: <https://doi.org/10.4312/slo2.0.2020.1.1-34>

Odperta znanost temelji na prosto in odprto dostopnih znanstvenih publikacijah in podatkih. Slednji omogočajo preverjanje rezultatov predhodnih raziskav in njihovo nadgrajevanje, v kontekstu jezikovnih tehnologij in ročno označenih jezikovnih virov pa tudi šolanje novih orodij za procesiranje besedil. Vendar pa je, tako kot za znanstvene objave, tudi za podatke pomembno, da so korektno citirani, saj šele to omogoča ponovljivost raziskav, citati pa so tudi najpomembnejši pokazatelj zanimivosti in koristnosti delovanja znanstvenikov ter pomembno vplivajo na njihovo priznanost in s tem možnost pridobivanja projektov ter zaposlitev. V prispevku najprej predstavimo ti. »austinska načela« citiranja jezikovnih podatkov in opišemo tovrstne aktivnosti v sklopu infrastrukture CLARIN.SI. Nato analiziramo stanje citiranja jezikovnih podatkov, predvsem korpusov, v šestih vodilnih slovenskih jezikoslovnih znanstvenih revijah (*Jezik in slovstvo*, *Slavistična revija*, *Slovenščina 2.0*, *Linguistica*, *Slovene Linguistic Studies* in *Jezikoslovni zapiski*) ter v zbornikih dveh znanstvenih konferenc z jezikoslovno tematiko (*Jezikovne tehnologije in digitalna humanistika* ter *Obdobja*) za obdobje zadnjih sedmih let, tj. 2013–2019. Pregledali smo 1.074 znanstvenih objav in kvantitativno ter kvalitativno analizirali rezultate. S kvantitativnega vidika pokažemo, da v celotnem obdobju zgoj dobra

četrtnina pregledanih člankov vključuje rabo virov ter da je v poznejšem obdobju (2018–2019) raba virov v objavah več kot dvakrat pogostejša kot v zgodnejšem obdobju (2013–2017). Načine navajanja virov razvrstimo v pet kategorij (npr. *navajanje hiperpovezave na vir v besedilu* ter *navajanje ključne publikacije o viru*); pokažemo, da je raba posameznega načina v veliki meri odvisna od navodil avtorjem za posamezno publikacijo. S kvalitativnega vidika se osredotočamo predvsem na vire z vnosom v repozitoriju raziskovalne infrastrukture CLARIN.SI, za katere pokažemo, da so z redkimi izjemami neustrezno citirani. Izsledke povzamemo in po ti. »austinskih načelih« pokažemo, kaj je bilo že narejenega v sklopu infrastrukture CLARIN.SI ter predlagamo smernice za citiranje jezikoslovnih podatkov in načine za njihovo implementacijo.

**Ključne besede:** odprta znanost, citiranje raziskovalnih podatkov, jezikovni viri, austinska načela, slovenske revije in zborniki

## 1 UVOD

Odprti dostop do znanstvenih publikacij in podatkov pospešuje inovacije, spodbuja sodelovanje, zmanjšuje podvajanje dela in omogoča dograjevanje predhodnih rezultatov raziskav ter vključevanje državljanov in družbe (European Commission, 2012). Odprti dostop do rezultatov raziskav predvidevajo *Resolucija o nacionalnem programu za jezikovno politiko 2014–2018*,<sup>1</sup> *Nacionalna strategija odprtega dostopa do znanstvenih objav in raziskovalnih podatkov v Sloveniji 2015–2020*<sup>2</sup> ter *Akcijski načrt izvedbe nacionalne strategije odprtega dostopa do znanstvenih objav in raziskovalnih podatkov v Sloveniji 2015–2020*.<sup>3</sup> Med pglavitnimi cilji *Akcijskega načrta* je bil pilotni program *Odprti dostop do raziskovalnih podatkov v letih 2017–2020*, ki sicer v tem obdobju nikoli ni bil celotno izveden, si je pa prizadeval izboljšati dostop do raziskovalnih podatkov, mdr. z uvedbo novega sistema za vrednotenje raziskovalnih podatkov, v skladu s katerim bodo raziskovalni podatki, shranjeni v pooblaščenem podatkovnem središču, ki so prestali presojo pomena za znanost, priznani kot

1 <http://www.pisrs.si/Pis.web/pregledPredpisa?id=RESO91#>

2 [http://www.mizs.gov.si/delovna\\_podrocja/direktorat\\_za\\_znanost/sektor\\_za\\_znanost/strategije\\_s\\_podrocja\\_znanosti/nacionalna\\_strategija\\_odprtega\\_dostopa\\_do\\_znanstvenih\\_objav\\_in\\_raziskovalnih\\_podatkov\\_v\\_sloveniji\\_2015\\_2020/](http://www.mizs.gov.si/delovna_podrocja/direktorat_za_znanost/sektor_za_znanost/strategije_s_podrocja_znanosti/nacionalna_strategija_odprtega_dostopa_do_znanstvenih_objav_in_raziskovalnih_podatkov_v_sloveniji_2015_2020/)

3 [http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Odprti\\_dostop/Akcijski\\_nacrt\\_-\\_POTRJENA\\_VERZIJA.pdf](http://www.mizs.gov.si/fileadmin/mizs.gov.si/pageuploads/Znanost/doc/Odprti_dostop/Akcijski_nacrt_-_POTRJENA_VERZIJA.pdf)

znanstvena objava. Dobra praksa doslednega citiranja raziskovalnih podatkov je pomembna, ker zagotavlja in spodbuja transparentnost znanstvenega dela in posledično deluje kot ključni vzvod tovrstnega sistema vrednotenja.

V Sloveniji imamo na področju jezikovnih virov že dolgo tradicijo odprtih podatkov. Že od nastanka so bili odprto dostopni npr. jezikovni viri projektov MULTEXT-East,<sup>4</sup> JOS<sup>5</sup> in SSJ,<sup>6</sup> leta 2013 pa je bila ustanovljena raziskovalna infrastruktura za jezikovne vire in orodja CLARIN.SI,<sup>7</sup> v sklopu katere je bil vzpostavljen certificiran repozitorij, ki trenutno arhivira več kot sto petdeset odprto dostopnih jezikovnih virov.

V pričujočem prispevku nas zanima, kako se uporaba jezikovnih virov citira v znanstvenih člankih vodilnih slovenskih jezikoslovnih publikacij. O stanju citiranja jezikovnih virov smo pred več kot desetimi leti zapisali naslednje:

Citiranje je še posebej pomembno, ker je merljiv kazalec raziskovalne uspešnosti, zato bi se tudi moralo dosledno izvajati. Žal pa to ni v navadi pri citiranju publikacij o jezikovnih virih: vse pre pogosto se nek vir omeni samo po imenu ali pa se v najboljšem primeru doda njegov spletni naslov, namesto da bi se v virih citiralo publikacijo, kjer je vir prvotno opisan. (Erjavec, 2009)

Od takrat se je stanje spremenilo, tako da je sedaj mogoče citirati ne samo publikacije o izdelavi nekega vira, pač pa tudi vir sam, saj repozitorij CLARIN.SI ponuja dolgoročno hrambo virov, ob tem pa za vsak vnesen vir na samem vrhu njegove spletne strani navaja, kako naj se ga citira. Dostop do podatkov v certificiranih repozitorijih, kot je CLARIN.SI, je v skladu s ti. austinskimi načeli (Berez-Kroeker idr., 2017) za ustrezno citiranje v jezikoslovju, ki so povzeti v dokumentu *The FORCE11 Joint Declaration of Data Citation Principles* (Data Citation Synthesis Group, 2017).<sup>8</sup> Poleg tega, da natančna navodila za citiranje jezikovnih virov sledijo drugi točki austinskih načel (*Credit and Attribution*, »Priznanje zaslug in avtorstva«), so transparentni metapodatki in stalni spletni identifikatorji, ki jih repozitoriji nudijo za vsak vir, ključnega pomena za

---

4 <http://nl.ijs.si/ME/>

5 <http://nl.ijs.si/jos/>

6 <http://www.slovenscina.eu/>

7 <https://www.clarin.si/repository/xmlui/>

8 <https://www.force11.org/datacitationprinciples>

zagotavljanje odprtega dostopa in s tem interoperabilnosti, trajnosti in preverljivosti podatkov.

Pričujoči prispevek je razširitev konferenčnega prispevka Fišer, Lenardič in Erjavec (2018), v katerem smo pregledali, kako so slovenski raziskovalci citirali jezikovne vire v šestih vodilnih jezikoslovnih revijah (*Jezik in slovstvo*, *Slavistična revija*, *Slovenščina 2.0*, *Slovene Linguistic Studies*, *Linguistica* in *Jezikoslovni zapiski*) ter zbornikih dveh znanstvenih konferenc z jezikoslovno tematiko (*Jezikovne tehnologije in digitalna humanistika* ter *Obdobja*), ki so izšli v obdobju med 2013 ter 2017. V tem prispevku pregled razširimo z analizo objav v zadnjih dveh letih, tj. 2018 in 2019.

Prispevek ima sledečo strukturo: v 2. razdelku podamo pregled mednarodnih načel in praks pri citiranju znanstvenih podatkov v jezikoslovju, 3. razdelek analizira stanje v izbranih slovenskih publikacijah, 4. razdelek predlaga smernice za boljšo prakso na tem področju, zadnji razdelek pa zaključí in poda smernice za nadaljnje delo.

## **2 MEDNARODNA NAČELA CITIRANJA PODATKOV V JEZIKOSLOVJU**

Odperta znanost, odprti podatki in citiranje le-teh je v svetu trenutno v središču pozornosti, saj so obstoječe prakse tudi mednarodno zastarele, manj v naravoslovju in posebej računalništvu, mnogo bolj pa v humanistiki in jezikoslovju; tako npr. relativno nova »Splošna pravila za oblikovanje jezikoslovnih prispevkov« (Haspelmath, 2014) citiranja podatkov sploh ne omenjajo.

Obširen pregled pomena odprte znanosti, odprtih podatkov in potrebe po korektnem citiranju v jezikoslovju je podan v Berez-Kroeker idr. (2018), ki je rezultat iniciative, v kateri je sodelovalo 41 jezikoslovcev in drugih znanstvenikov. Prispevek najprej osmisli odprte raziskovalne podatke in ponovljivost raziskav, tako na splošno kot v jezikoslovju, nato pa poda pregled trenutnega stanja v jezikoslovju, kar se tiče transparentnosti uporabljenih virov in raziskovalnih metodologij. Avtorji ugotavljajo, da je po eni strani nemogoče uveljaviti ponovljivost raziskav brez primerne citiranja virov, po drugi pa, da je stanje v jezikoslovju še vedno zelo nezadovoljivo. Nato sledijo ugotovitve avtorjev glede potrebe po mehanizmih, ki bi ovrednotila tudi »delo na podatkih«

pri zaposlovanju in napredovanju znanstvenikov, in nujnosti po korenitem premiku v omogočanju ponovljivosti raziskav v jezikoslovju, kar naj bi dosegli skozi izobraževanje, promocijo in razvoj ustreznih politik. Strinjajo se, da bi zbiralci podatkov za svoje delo morali dobiti primerno priznanje avtorstva, posebej takrat, ko so izdelani podatki dostopni, ponovno uporabni in jih je mogoče citirati. Prispevek zaključijo priporočila za konkretne dejavnosti, ki bi jih morali izvesti jezikoslovci, oddelki, sveti in založniki. Te dejavnosti so v veliki meri osredotočene na zagotovitev odprtih podatkov oz. izobraževanje, kako se s podatki upravlja, da sploh lahko postanejo odprti, kot tudi, kako to delo primerno ovrednotiti. Zadnje priporočilo pa je neposredno posvečeno boljšemu citiranju raziskovalnih podatkov, kjer avtorji prispevka urednikom ter založnikom znanstvenih revij in knjig svetujejo uvedbo konkretnih politik tako za izmenjavo podatkov kot za njihovo citiranje, pri slednjem tako, da izoblikujejo formate za citiranje jezikoslovnih podatkov.

### 3 ANALIZA CITIRANJA OBJAV V SLOVENSKIH ZNANSTVENIH PUBLIKACIJAH

#### 3.1 Izbor gradiva in zasnova analize

Za pričujoči prispevek smo pregledali ključne slovenske revije in zbornike za področje jezikoslovja in ugotavljali, v kolikšni meri in na kakšen način avtorji prispevkov omenjajo oz. navajajo jezikovne vire. Naj poudarimo, da nas v tej raziskavi ni zanimalo, kateri jezikovni viri so v objavljenih raziskavah uporabljeni in citirani, temveč, kako jih avtorji navajajo.

Pri revijah smo analizirali navodila za avtorje in izdane številke za zadnjih sedem let (2013–2019), pri zbornikih pa navodila za avtorje oz. predloge prispevkov ter štiri zbornike dveh konferenc z jezikoslovno tematiko. Med zborniki smo v študijo zajeli *Jezikovne tehnologije in digitalna humanistika (JTDH) 2016*, *JTDH 2018*, *Obdobja 2016* ter *Obdobja 2019*, med revijami pa *Linguistica*, *Jezik in slovstvo*, *Jezikoslovni zapiski*, *Slavistična revija*, *Slovene Linguistic Studies* in *Slovenščina 2.0*.<sup>9</sup>

---

9 V primeru revije *Linguistica* nismo upoštevali 58. zvezka iz leta 2018, saj so vsi članki v njem napisani v francoščini, nihče od avtorjev pričujočega prispevka pa ni kompetenten govorec tega jezika.



Analiza citiranja je bila opravljena dvakrat (Tabela 1). Prvi pregled, ki smo ga že predstavili v Fišer, Lenardič in Erjavec (2018), je bil opravljen za obdobje 2013–2017, v katerem smo pregledali 751 znanstvenih prispevkov, od katerih jih vire omenja 133 oz. dobrih 17 %. Temu smo nato dodali še pregled za obdobje 2018–2019, v katerem smo pregledali 323 objav, od katerih jih vire omenja 155 oziroma 48 %. Skupaj smo torej pregledali 1.074 znanstvenih objav iz obdobja 2013–2019, od katerih jezikovne vire navaja 288 član- kov (27 %).

Navedbe virov v pregledanih prispevkih ločujemo na naslednje kategorije:

**Povezava<sup>10</sup> na vir v besedilu prispevka** (največkrat v opombi). Zgled takega citiranja je v Žele (2014) v *Slavistični reviji*, kjer avtorica povezavo na korpus *Gigafida* podaja v opombi. Prispevki v tovrstni kategoriji ne navajajo ključne publikacije o viru, ki bi v tem primeru bila Logar Berginc idr. (2012).

**Povezava na vir v bibliografiji.** Zgled takega citiranja je v Trivunović (2019) v reviji *Jezikoslovni zapiski*, kjer avtorica navaja korpusa *Gigafida 2.0* in *IMP* tako, da podaja hiperpovezavo v končnem seznamu virov, ne navaja pa ključnih publikacij o viru, se pravi Logar Berginc idr. (2012) za *Gigafida 2.0* ter Erjavec (2015a) za *IMP*.

**Povezava na vir v besedilu prispevka kot tudi v bibliografiji.** Zgled takega citiranja je Žele (2018) v reviji *Jezik in slovstvo*, kjer avtorica podaja povezavo na korpus *Gigafida* v opombi<sup>11</sup> ter v končnem seznamu virov. Tudi prispevki v tej kategoriji ne navajajo ključne publikacije o viru.

**Publikacija o viru.** Zgled takega citiranja je v Ljubešić, Miličević Petrović in Samardžić (2019) v *Slavistični reviji*, kjer avtorji za orodje *TweetCat* navajajo ključno publikacijo o viru, tj. Ljubešić, Fišer in Erjavec (2014).

**Povezava na vir v besedilu prispevka in publikacija o viru.** Zgled takega citiranja je v Bálint Čeh in Kosem (2017). Avtorja podajata povezavo na

---

10 V to kategorijo vključujemo tudi navedbe stalnih spletnih identifikatorjev, kot sta handle in DOI.

11 Pri tem prispevku je zanimivo, da avtorica za korpus *Gigafida* navaja hiperpovezavo na demonstracijsko različico korpusa, tj. <http://demo.gigafida.net>, ki pa že od leta 2014 ni dostopna, medtem ko v bibliografiji navaja običajnejšo povezavo na spletno stran <http://www.gigafida.net>.

korpus *Gigafida* v opombi in navajata ključno publikacijo, tj. Logar Berginc idr. (2012).

**Kombinacija različnih načinov navajanja virov.** Zgled takega navajanja je v Žitnik in Bajec (2018) v *Slovenščina 2.0*. Avtorja navajata korpus *ssj500k 1.4* z medbesedilno navedbo avtorjev korpusa (tj. Krek idr. 2015) ter s trajnim identifikatorjem v pripadajočem zapisu v bibliografiji – v razdelku 3.4 bomo videli, da je to najustreznejši način navedbe jezikovnega vira v primeru, da ima vir vnos v jezikovnem repozitoriju. Nadalje pa Žitnik in Bajec (2018) navajata ključni publikaciji za orodji SkipCor (tj. Žitnik, Šubelj in Bajec, 2014) ter nutIE (Žitnik idr., 2017), vendar za vira ne podajata povezav, saj nobeno orodje ni prosto dostopno.

**Brez navedbe hiperpovezave na vir ali ključne publikacije o viru.** Zgled takega citiranja je v Vidovič Muha (2015). Avtorica se sklicuje na uporabo označevalnika *JOS*, vendar ne podaja niti povezave na vir niti ne navaja njegove ključne publikacije (tj. Erjavec idr., 2010).

**Tabela 1:** Pregled distribucije različnih načinov navajanja virov v analiziranih publikacijah

	2013–2017		2018–2019		Σ	
Vseh objav	751		323		1.074	
Objave z omembo vira	133	100 %	155	100 %	288	100 %
Povezava na vir v besedilu prispevka	13	10 %	11	7 %	24	8 %
Povezava na vir v bibliografiji	33	25 %	39	25 %	72	25 %
Povezava na vir v besedilu in v bibliografiji	14	11 %	17	13 %	31	11 %
Publikacija o viru	16	12 %	11	7 %	27	9 %
Povezava na vir v besedilu prispevka in publikacija o viru	8	6 %	13	8 %	21	7 %
Kombinirano	25	18 %	24	15 %	49	17 %
Brez	25	18 %	40	25 %	65	23 %

### 3.2 Pregled navodil avtorjem

Ker smo sklepali, da je od navodil avtorjem, ki so jih pripravili uredniški odbori revij in programski odbori konferenc, močno odvisno, kako bodo avtorji navajali vire, jih povzemamo v tem razdelku. Za revijo *Jezikoslovni zapiski* ter za zbornik *Obdobja* navodil avtorjem na njihovih spletnih straneh nismo našli.

Najpodrobnejša navodila za navajanje virov podaja revija *Slovenščina 2.0*,<sup>12</sup> ki ločuje navajanje korpusov, spletnih strani in spletnih virov:

Korpus:

- Gigafida. Dostopno prek: <http://www.gigafida.net> (datum dostopa).
- Cambridge English Corpus. Dostopno prek: [http://www.cambridge.org/gb/elt/catalogue/subject/item2701617/Cambridge-International-Corpus/?site\\_locale=en\\_GB](http://www.cambridge.org/gb/elt/catalogue/subject/item2701617/Cambridge-International-Corpus/?site_locale=en_GB) (datum dostopa).

Spletna stran:

- OpenWebSpider. Dostopno prek: <http://www.openwebspider.org/> (datum dostopa).
- Creative Commons. Dostopno prek: <http://creativecommons.org/> (datum dostopa).

Spletni vir:

- Pew Research Center (2010): Americans Spending More Time Following the News ? Ideological News Sources: Who Watches and Why. Dostopno prek: <http://www.people-press.org/> (datum dostopa).
- TEI Consortium, ur. (2011): TEI P5: Guidelines for Electronic Text Encoding and Interchange: Version 1.9.1. Dostopno prek: <http://www.tei-c.org/Guidelines/P5/> (datum dostopa).
- Scott, M. (2008): WordSmith Tools: Version 5. Dostopno prek: <http://www.lexically.net/downloads/version5/HTML/index.html> (datum dostopa).

*Jezik in slovstvo* avtorje poziva,<sup>13</sup> da vire in literaturo navajajo ločeno, kar se nam zdi dobra praksa, saj s tem avtorjem med drugim sporočajo, da sta uporaba in navajanje virov pomemben sestavni del znanstvenega prispevka. Dodatno velja omeniti, da poziv k ločenemu navajanju jezikovnih virov omogoča bralcem lažji dostop in preverjanje citiranih podatkov, ki podpirajo neko znanstveno trditev, kar je skladno tudi z austinskimi načeli (glej razdelek 4). Podrobneje ta revija načina

---

12 <http://slovenscina2.0.trojina.si/si/oddaja-prispevkov/>

13 <https://www.jezikinslovstvo.com/o2.php>

za navajanje jezikovnih virov sicer ne definira, iz primera za navajanje spletnih strani pa lahko sklepamo, da jezikovne vire v elektronski obliki enači s spletnimi stranmi, saj kot primer navajanja spletnih strani navaja korpus *FidaPLUS*:

- Korpus slovenskega jezika FidaPLUS: <<http://www.fidaplus.net>>. (Dostop dan. mesec. leto.)

Na podoben način jezikovne vire obravnava revija *Linguistica*:<sup>14</sup>

- Le dictionnaire de la zone. 20 May 2010. <http://www.dictionnairedelazone.fr/>.

*Slavistična revija*<sup>15</sup> v navodilih za oblikovanje seznama literature uvaja zelo nenatančno navajanje spletnih virov, brez navedbe spletnih povezav, verzij ali datuma dostopa:

- Lemma (Lexikographie). Wikipedia: Die freie Enzyklopädie.
- Primož JAKOPIN, 1980: Zgornja meja entropije pri leposlovnih besedilih v slovenskem jeziku: Doktorska disertacija. Ljubljana. Na spletu.

Pri prvem primeru ni jasno, na katero različico vira se referenca nanaša, saj je Wikipedija kolaborativen projekt, kjer uredniki gesla lahko ves čas spreminjajo, zato bi bilo nujno treba dodati datum dostopa. Pri drugem primeru pa ni jasno, ali gre za referenco na doktorsko disertacijo kot publikacijo ali za jezikovni vir, ki je bil v okviru disertacije razvit. Predvsem pa nobena od referenc ne vsebuje spletne povezave, zato bralec do vira ne more dostopati. Tovrstna praksa ne spodbuja preverljivosti in ponovljivosti raziskav ter priznavanja zaslug avtorjem virov, zato bi jo bilo pomembno čim prej izboljšati, še posebej, ker gre za jezikoslovno revijo, ki se v sistemu vrednotenja znanstvenih objav v slovenskem okolju uvršča v sam vrh.

Revija *Slovene Linguistic Studies*<sup>16</sup> posebej za navajanje elektronskih virov ne podaja navodil.

Podobno zbornik *JTDH*<sup>17</sup> v predlogi prispevkov sicer vsebuje primer dodajanja hiperpovezav v opombe in navaja načine navajanja različnih tipov enot

14 <https://revije.ff.uni-lj.si/linguistica/about/submissions#authorGuidelines>

15 [https://srl.si/navodila\\_guidelines.pdf](https://srl.si/navodila_guidelines.pdf)

16 <https://sjsls.byu.edu/guidelines-for-contributors/>

17 <http://www.sdj.si/wp/dogodki/konference/jtdh-2018/#navodila>

bibliografije, a med njimi ni primerov za citiranje jezikovnih virov. Glede na to, da gre za vodilno konferenco za področje jezikovnih virov in tehnologij, bi konferenca nujno morala posvečati več pozornosti ozaveščanju in usmerjanju avtorjev prispevkov za ustrezno citiranje jezikovnih virov.

### 3.3 Kvantitativna analiza

Glede na podatke v Tabeli 1 vsebuje 288 (27 %) pregledanih objav (vsaj eno) navedbo jezikovnega vira, načini navajanja pa so zelo raznoliki in razpršeni. Izrazito prevladuje navajanje povezave na vir v bibliografiji, česar se poslužuje četrtina vseh prispevkov, v katerih so bili viri uporabljeni. Dvakrat redkejša je praksa navajanja ključne publikacije o uporabljenem viru, ki je v trenutno veljavnem sistemu, ki seveda ni popoln in ni (primarni) cilj znanstvenega udejstvovanja, je pa kljub vsemu zelo pomemben za pridobivanje zaposlitev in projektov, za vrednotenje znanstvene uspešnosti edini način citiranja, ki avtorjem vira prinaša točke. Precej pogosto je kombiniranje več različnih načinov navajanja virov v istem prispevku (17 %), kar kaže na neupoštevanje navodil avtorjem oz. na pomanjkljiva navodila.

**Tabela 2:** Pregled praks navajanja jezikovnih virov v ključnih slovenskih znanstvenih revijah za področje jezikoslovja za obdobje 2013–2017 ter za obdobje 2018–2019

	<i>Jezik in slovstvo</i>						<i>Slavistična revija</i>					
	2013–17		2018–19		Σ		2013–17		2018–19		Σ	
Vse objave	157	100 %	52	100 %	209	100 %	180	100 %	78	100 %	258	100 %
Omemba vira	11	7 %	17	33 %	28	13 %	14	8 %	39	50 %	53	21 %
URL na vir v besedilu	0	0 %	0	0 %	0	0 %	1	7 %	2	5 %	3	6 %
URL na vir v biblio.	4	36 %	6	34 %	10	36 %	2	14 %	4	10 %	6	11 %
URL na vir v besedilu in biblio.	2	18 %	3	18 %	5	18 %	0	0 %	0	0 %	0	0 %
Publikacija o viru	2	18 %	2	12 %	4	14 %	0	0 %	3	8 %	3	6 %
URL na vir in publ.	0	0 %	2	12 %	2	7 %	0	0 %	0	0 %	0	0 %
Kombinirano	0	0 %	1	6 %	1	4 %	0	0 %	2	5 %	2	4 %
Brez	3	27 %	3	18 %	6	21 %	11	79 %	28	72 %	39	73 %

	<i>Slovenščina 2.0</i>						<i>Slovene Linguistic Studies</i>					
	2013–17		2018–19		Σ		2013–17		2018–19		Σ	
Vse objave	45	100 %	16	100 %	61	100 %	26	100 %	9	100 %	35	100 %
Omemba vira	34	76 %	9	56 %	43	70 %	8	31 %	5	56 %	13	37 %
URL na vir v besedilu	5	15 %	0	0 %	5	12 %	2	25 %	0	0 %	2	15 %
URL na vir v biblio.	5	15 %	1	11 %	6	14 %	2	25 %	3	60 %	5	38 %
URL na vir v besedilu in biblio.	3	9 %	1	11 %	4	9 %	3	37 %	0	0 %	3	23 %
Publikacija o viru	6	18 %	1	11 %	7	16 %	0	0 %	0	0 %	0	0 %
URL na vir in publ.	6	18 %	1	11 %	7	16 %	0	0 %	0	0 %	0	0 %
Kombinirano	8	23 %	4	45 %	12	28 %	1	13 %	1	20 %	2	15 %
Brez	1	3 %	1	11 %	2	5 %	0	0 %	1	20 %	1	8 %
	<i>Linguistica</i>						<i>Jezikoslovni zapiski</i>					
	2013–17		2018–19		Σ		2013–17		2018–19		Σ	
Vse objave	134	100 %	24	100 %	158	100 %	115	100 %	52	100 %	167	100 %
Omemba vira	6	4 %	12	50 %	18	11 %	20	17 %	17	33 %	37	22 %
URL na vir v besedilu	0	0 %	0	0 %	0	0 %	3	15 %	2	12 %	5	13 %
URL na vir v biblio.	4	66 %	3	25 %	7	39 %	7	35 %	10	58 %	17	46 %
URL na vir v besedilu in biblio.	1	17 %	5	42 %	6	33 %	2	10 %	2	12 %	4	11 %
Publikacija o viru	0	0 %	0	0 %	0	0 %	1	5 %	0	0 %	1	3 %
URL na vir in publ.	0	0 %	0	0 %	0	0 %	1	5 %	1	6 %	2	5 %
Kombinirano	0	0 %	0	0 %	0	0 %	0	0 %	1	6 %	1	3 %
Brez	1	17 %	4	33 %	5	28 %	6	30 %	1	6 %	7	19 %

V Tabeli 2 navajamo rezultate analize za posamezne revije, ki smo jih vključili v raziskavo. Za celotno obdobje pregleda 2013–2019 vsebuje najvišji delež prispevkov, ki omenjajo jezikovne vire, revija *Slovenščina 2.0* (70 %), najnižjega pa revija *Linguistica* (11 %). Pomembno je, da je pri vseh revijah (razen pri *Slovenščina 2.0*) v drugem obdobju pregleda – tj. 2018–2019 – bistveno večji delež objav, ki temeljijo na rabi jezikovnih virov, kot v prvem obdobju pregleda 2013–2017. Največjo rast opazimo pri reviji *Linguistica*, v kateri v

prvem obdobju 2013–2017 zgolj 4 % vseh objav vključuje rabo jezikovnih virov, medtem ko v drugem obdobju 2018–2019 ta delež zraste na polovico vseh prispevkov. Podobno opazimo pri *Slavistični reviji*, kjer v obdobju 2013–2017 rabo virov vključuje zgolj 8 % objav, medtem ko je ta delež v 2018–2019 tudi tu polovica vseh objav. Čeprav se zdi, da je ta dvig vsaj delno posledica dejstva, da so posamezne številke pregledanih revij v drugem obdobju v splošnem bolj osredotočene na empirično jezikoslovje kot v prvem,<sup>18</sup> je porast verjetno vseeno indikator vedno močnejše težnje k rabi jezikovnih virov, saj smo ga identificirali pri skoraj vseh vodilnih slovenskih jezikoslovnih revijah, ki so programsko precej raznolike.

Najbolj homogeno navajanje virov je v reviji *Linguistica*, kjer smo identificirali zgolj dva različna načina citiranja (od šestih možnih): povezava na vir zgolj v bibliografiji ter povezava na isti vir tako v besedilu kot v bibliografiji. Najbolj heterogeni načini citiranja so v revijah *Jezikoslovnih zapiskih* ter *Slovenščina 2.0*, kjer najdemo vseh pet načinov navajanja virov. Najvišji delež navedbe vira v obliki hiperpovezave na spletno stran vira, podane v bibliografiji, najdemo v reviji *Jezikoslovni zapiski* (46 %), najvišji delež citiranja ključnega prispevka o viru pa pripada reviji *Slovenščina 2.0* (16 %). Od posameznih načinov navajanja virov je navajanje povezav na vir v bibliografiji prispevka najpogostejši način navajanja virov v vseh revijah, razen v reviji *Slovenščina 2.0*, kjer je nekoliko pogostejše kombinirano citiranje.

Po nenavajanju uporabljenih virov izrazito izstopa *Slavistična revija*, v kateri avtorji pri skoraj tri četrtini (73 %) prispevkov, ki rabo virov omenjajo, teh virov nikjer ne citirajo. Zanimivo je tudi, da delež člankov v *Slavistični reviji* brez ustrezne navedbe ostaja približno enak v obeh analiziranih obdobjih, kar poleg tega, da skoraj vsaka revija dosledno ohranja svoj prevladujoči način citiranja v obeh obdobjih (npr. kombinirano citiranje pri *Slovenščina 2.0* ter navedbe povezav zgolj v bibliografiji in nič primerov objav z navedbo med besedilom pri *Jeziku in slovstvo*), daje misliti, da so načini citiranja

---

18 Na primer 59. številka revije *Linguistica* iz 2018, ki je bila edina pregledana številka te revije za obdobje 2018–2019, se je osredotočala na nemško aplikativno jezikoslovje, ki pogosto temelji na rabi jezikovnih virov, medtem ko so se domala vse številke v obdobju 2013–2017 osredotočale na različne veje teoretičnega jezikoslovja (npr. generativna slovnica), ki so inherentno manj vezane na preveritvi jezikovnih podatkov v sodobnih korpusih.

precej odvisni od navodil avtorjem v posamezni reviji. V prejšnjem razdelku smo na primer videli, da *Slavistična revija* podaja zelo nespecifična navodila za navajanje jezikovnih virov, ki zahtevajo precej neuporabno oznako *Na spletu* namesto dejanskih hiperpovezav, kar tako najbrž botruje visokemu deležu neustreznih oziroma manjkajočih navedb (o problematiki hiperpovezav glej tudi razdelek 3.4). Glede na to, da gre za vodilno jezikoslovno revijo v našem prostoru, ki je uvrščena tudi na seznam ARRS revij posebnega pomena, bi še posebej to uredništvo revije moralo skrbeti za visok nivo raziskovalne kulture v slovenskem jezikoslovju in ustrezno citiranje raziskovalnih podatkov od avtorjev izrecno zahtevati v navodilih za avtorje.

**Tabela 3:** Pregled praks navajanja jezikovnih virov v ključnih konferenčnih zbornikih za področje jezikoslovja

	Obdobja						JTDH					
	2016		2019		Σ		2016		2018		Σ	
Vse objave	64	100 %	56	100 %	120	100 %	30	100 %	36	100 %	66	100 %
Omemba vira	12	19 %	23	41 %	35	29 %	28	93 %	33	92 %	61	92 %
URL na vir v besedilu	0	0 %	1	4 %	1	3 %	2	7 %	6	18 %	8	13 %
URL na vir v biblio.	7	58 %	13	56 %	20	57 %	2	7 %	0	0 %	2	3 %
URL na vir v besedilu in biblio.	0	0 %	2	9 %	2	6 %	3	11 %	3	9 %	6	10 %
Publikacija o viru	1	8 %	1	4 %	2	6 %	6	21 %	4	12 %	10	16 %
URL na vir in publ.	0	0 %	2	9 %	2	6 %	1	4 %	7	21 %	8	13 %
Kombinirano	2	17 %	2	9 %	4	11 %	13	46 %	13	39 %	26	43 %
Brez	2	17 %	2	9 %	4	11 %	1	4 %	0	0 %	1	2 %

V Tabeli 3 navajamo rezultate za konferenci, ki smo ju vključili v raziskavo. V zborniku *JTDH 2016* so viri omenjeni v 93 % vključenih prispevkov, v zborniku *JTDH 2018* pa v 92 % prispevkov, kar je glede na področje konference razumljivo. V obeh edicijah zbornika naletimo na izrazito velik delež prispevkov (vse skupaj 43 %), v katerih avtorji uporabljajo različne kombinacije navajanja virov. To je verjetno odraz heterogene raziskovalne skupnosti, ki se predstavlja na tej konferenci, ter heterogenosti dostopnosti virov – npr., nekateri viri so deponirano v trajnih repozitorijih, ki beležijo tudi ključne



publikacije, medtem ko drugi niso dostopni in so tako navedeni zgolj s ključno publikacijo.

V zborniku *Obdobja* je prispevkov z omembo jezikovnega vira leta 2016 19 %, leta 2019 pa 41 %. Glede na to, da sta bili ediciji simpozija iz 2016 ter 2019 tematsko vezani na jezikovni opis sodobne slovenščine, se zdi ta rezultat nizek. Vendar sta odstotka kljub temu tako v 2016 kot tudi v 2019 občutno višja od večine programsko sorodnih revijah v istem obdobju, predstavljenih v Tabeli 2, kar morda nakazuje spremembe sestave oz. praks tudi v tej skupnosti, saj so revije tradicionalno konzervativnejše in spremembe, do katerih v raziskovalni skupnosti prihaja, absorbirajo nekoliko kasneje od konferenc.

### **3.4 Kvalitativna analiza**

V tem razdelku navajamo zanimivejše pojave, na katere smo naleteli pri kvalitativnem pregledu gradiva. Najprej predstavljamo nekatere primere dobrih praks, nato pa analiziramo identificirane problematične primere navajanja virov. Kot zgleden primer citiranja virov navajamo Logar Berginc, Gantar in Kosem (2014) v reviji *Slovenščina 2.0*, ki za isti vir navaja tako ključno publikacijo o viru v bibliografiji kot tudi povezavo na vir v besedilu prispevka v sprotnih opombah, ki so prikazane na dnu relevantne strani prispevka. Na ta način bralcu omogočimo, da neposredno dostopa tako do vira kot tudi do publikacije o njem, prav tako pa avtorjem vira ustrezno priznamo zasluge in avtorstvo ter zagotovimo citiranost.

Naslednji zgleden primer citiranje virov, ki prav tako prihaja iz revije *Slovenščina 2.0*, je Arhar Holdt in Dobrovoljc (2016), ki v bibliografiji za vir navede stalni spletni identifikator handle v repozitoriju CLARIN.SI:

- Krek, S., Erjavec, T., Dobrovoljc, K., Može, S., Ledinek, N. in Holz, N. (2015): Training corpus ssj500k 1.4. Dostopno prek: <http://hdl.handle.net/11356/1052>.

Navajanje handlov je pomembno, saj bralcu zagotavlja, da bo lahko dostopal do vira, četudi se naslov njegove spletne strani spremeni. Prav tako pa handle bralcu omogoča dostop do podrobnejšega opisa jezikovnega vira, ki je bil uporabljen v raziskavi, do njegovih metapodatkov, za prosto dostopne vire pa tudi do vira samega. S tem je močno izboljšana preverljivost in ponovljivost

raziskav, spodbuja pa tudi nadaljnje razširitve in izboljšave raziskav ter maksimizira izrabo jezikovnega vira, izdelava katerega je zahtevala finančni in časovni vložek.

Za obdobje 2018–2019 smo v vseh šestih revijah v Tabeli 2 posebej opazovali, kako avtorji navajajo jezikovne vire, ki imajo vnos v repozitoriju CLARIN.SI, ki avtorjem omogoča, da se na vire sklicujejo z navedbo handlov. Tovrstnih člankov je 19 (od 155, se pravi dobrih 12 %), pri čemer gre v večini primerov za avtocitate. Na tem mestu jih za vsako revijo posebej navedemo skupaj z viri, ki so citirani:

- *Jezik in sloustvo* (2 objavi):
  - o Rozman idr. (2018), ki se sklicujejo na korpus *Šolar* (Rozman idr., 2013);
  - o Zwitter Vitez (2018), ki se sklicuje na korpus govornjene slovenščine *GOS* (Zwitter Vitez idr., 2013).
- *Slavistična revija* (3 objave):
  - o Dobrovoljc (2018a), ki se sklicuje na korpus *GOS*;
  - o Orel (2019), ki se sklicuje na korpus *GOS* ter na družino korpusov stare slovenščine *IMP* (Erjavec, 2014);
  - o Marvin idr. (2019), ki se sklicujejo na leksikalno bazo *SNABI* (Kačič idr. 2002) ter na korpus *ccKres* (Logar Berginc idr., 2013).
- *Slovene Linguistic Studies* (2 objavi):
  - o Jelovšek in Erjavec (2019), ki navajata referenčni korpus stare slovenščine *goo300k 1.2* (Erjavec, 2015b);
  - o Krvina (2019), ki uporablja korpus *IMP* ter korpus *ssj500k 2.2* (Krek idr., 2019).
- *Slovenščina 2.0* (5 objav):
  - o Žitnik in Bajec (2018), ki navajata korpus *ssj500k 1.4* (Krek idr., 2015);

- o Arhar Holdt in Čibej (2018), ki navajata morfološki leksikon *Sloleks* (Dobrovoljc idr., 2019);
- o Dobrovoljc (2018b), ki uporablja korpus *GOS*;
- o Pisanski Peterlin in Mikolič Južnič (2018), ki uporabljata *GOS*;
- o Pori in Kosem (2018), ki uporabljata vir *Leksikalna baza za slovenščino* (Gantar idr., 2012).
- *Jezikoslovni zapiski* (6 objav):
  - o Stramljič Breznik (2018), ki uporablja leksikon *Sloleks*;
  - o Furlan (2018), ki uporablja korpus *IMP*;
  - o Uhlik in Žele (2018), ki uporabljata *IMP*;
  - o Atelšek (2019), ki uporablja *IMP*;
  - o Hudeček in Mihaljević (2019), ki uporabljata spletni hrvaški korpus *hrWaC* (Ljubešić in Klubička, 2016);
  - o Trivunović (2019), ki navaja korpus *IMP*.
- *Linguistica* (1 objava):
  - o Petrič (2019), ki navaja korpus *GOS*.

Zgolj v treh (16,7 %) od zgoraj navedenih 18 objav je jezikovni vir naveden s povezavo handle na CLARIN.SI vnos, v katerem so, kot rečeno, sistematsko beleženi pomembni metapodatki, kot je avtorstvo vira. Te tri objave so:

- Jelovšek in Erjavec (2019) v *Slovene Linguistic Studies*;
- Žitnik in Bajec (2018) v *Slovenščina 2.0* ter
- Arhar Holdt in Čibej (2018) v *Slovenščina 2.0*.

Preostali avtorji običajno navajajo omenjene jezikovne vire zgolj s hiperpovezavo na projektno stran (namesto na trajni vnos v repozitoriju) ali na konkordančnik, s katerim so iskali po viru. Najpogosteje je neustrezno naveden korpus *IMP*, ki je namesto z ustreznejšo navedbo na Erjavec (2014) običajno citiran zgolj s podano hiperpovezavo na (ne nujno trajno) projektno stran, ki

jo gosti Inštitut »Jožef Stefan«. Variante slednjega (manj ustreznega) citiranja v bibliografiji med drugim opazimo pri Furlan (2018), Uhlik in Žele (2018) ter Atelšek (2019) v *Jezikoslovnih zapiskih*:

- IMP: korpus starejše slovenščine <<http://nl.ijs.si/imp/>, dostop xx.yy.201z>

Že prej smo omenili, da neoptimalno navajanje tovrstnih virov verjetno nastaja kot delna posledica obstoječih navodil avtorjem. To je predvsem razvidno iz prispevka Marvin idr. (2019) v *Slavistični reviji*, v katerem se avtorji sicer sklicujejo na leksikalno bazo *SNABI* z ustrežno navedbo Kačič idr. (2002) v repozitoriju CLARIN.SI, vendar je v bibliografiji namesto ustrezne povezave handle (tj. <http://hdl.handle.net/11356/1051>) navedena že prej omenjena oznaka *Na spletu* na sledeč način (Marvin idr., 2019, str. 549):

- Zdravko Kačič, Bogomir Horvat, Aleksandra Markuš Zögling, Robert Veronik, Matej Rojc, Andrej Žgank, Mirjam Sepesy Maučec in Tomaž Rotovnik, 2002: *SNABI Database for Continuous Speech Recognition 1.2. Slovenian language resource repository CLARIN.SI. Na spletu.*

Da tovrstna navedba brez handle nastaja neposredno zaradi navodil avtorjem, potrjuje konferenčna objava Marvin idr. (2018) iz zbornika *JTDH 2018*, ki je predhodna različica objave Marvin idr. (2019) v *Slavistični reviji*. Namreč, konferenčna objava Marvin idr. (2018) vsebuje isto referenco na jezikovni vir avtorjev Kačič idr. (2002), citiran na enak način kot v zgornjem zgledu iz *Slavistične revije*, vendar tokrat z navedbo ustreznega handle, tj. <http://hdl.handle.net/11356/1051>. Zaradi tovrstnih zgledov predlagamo, da bi bilo v obstoječih revijah treba navodila avtorjem razširiti – zdi se, da bi do velikega napredka prišlo že samo, če bi uredništvo nekaj besed namenilo dobri praksi citiranja, predvsem pomembnosti navajanja virov s trajnimi identifikatorji (če le-ti obstajajo) in hkratno navedbo avtorjev vira (o pomembnosti vključitve te informacije glej razdelek 4.2), saj tovrstna praksa trenutno ni opisana v navodilih nobene izmed pregledanih publikacij.

Poleg problematičnega citiranja virov s trajnim vnosom v repozitoriju smo naveli tudi na druge težave, ki jih uvrščamo v naslednje kategorije:

- Nekonsistentno navajanje istega vira: V *Slavistični reviji* je isti vir navajan zelo različno. Npr. v Meterc (2013) in Jakop (2014):

- o Gigafida, korpus slovenskega jezika. Ur. Filozofska fakulteta Univerze v Ljubljani. Ljubljana: FF. Splet.
- o Korpus GigaFida. Na spletu.
- Nekonsistentno navajanje različnih virov istega tipa: V reviji *Slovene Linguistic Studies* je v Štumberger (2015) za *Sloleks* navedena hiperpovezava v opombi, nemška leksikalna vira pa sta vključena v bibliografijo:
  - o OWID, Online–Wortschatz-Informationssystem Deutsch des Instituts für deutsche Sprache, Mannheim, (13. 3. 2008).
  - o Klappenbach, Ruth, Steinitz, Wolfgang (ur.). 1967 (1964). Wörterbuch der deutschen Gegenwartssprache. 1. Band. Berlin: Akademie-Verlag. <http://www.dwds.de/> (1. 7. 2008, 27. 3. 2015).
- Neustrezne hiperpovezave: V reviji *Jezikoslovni zapiski* smo opazili neustrezno navajanje povezav na vire. Npr. v Polajnar (2013) ni hiperpovezave na osnovno stran vira, ampak na podstran:
  - o Gigafida: <http://www.Gigafida.net/Support/About>

Zastareli viri: Rath (2019) v *Slovene Linguistic Studies*, Stopar in Ilc (2019) v *Slavistični reviji* ter Kulčar (2018) v *Jezikoslovnih zapiskih* navajajo jezikovne zglede iz korpusa FidaPLUS, kar je nenavadno, saj je ta korpus že od leta 2012 v celoti vključen v Gigafida. Rath (2019) za ta korpus navaja hiperpovezavo <http://www.fidaplus.net/>, ki pa ne deluje že od leta 2016.

Ker je korpuse mogoče naložiti na različne konkordančnike, kar lahko privede tudi do razlik v rezultatih, je za zagotavljanje preverljivosti in ponovljivosti raziskav v referenci nujno potrebno navesti natančno povezavo, ki je bila v raziskavi uporabljena. Velja poudariti, da repozitorij *CLARIN.SI* rešuje ta problem tako, da je v navodilih za navajanje virov, ki so podana kot prva informacija v glavi vnosa za posamezen vir, jasno izpostavljeno, za katero različico vira gre in ali je ta različica dostopna preko spletnega konkordančnika. Za starejše različice repozitorij opozori o morebitni zastarelosti podatkov. Kot primer navajamo vnos za drugo različico korpusa *Gos VideoLectures (Transcriptions)* (Verdonik idr., 2017), ki je dostopna preko

konkordančnika *KonText*,<sup>19</sup> medtem ko prva različica (Verdonik idr., 2016) preko tega konkordančnika ni dostopna.

V reviji *Slovenščina 2.0* smo opazili nenatančno navajanje hiperpovezave do korpusa. Npr. v Arias-Badia, Bernal in Alonso (2014), kjer je za španski korpus navedena generična povezava na konkordančnik SketchEngine:

- SWC = Spanish Web Corpus. Available at: [www.sketchengine.co.uk](http://www.sketchengine.co.uk) (20 October 2014).

Tovrstno navajanje referenc na korpus je med jezikoslovci precej razširjeno, je pa problematično iz več razlogov. Ne samo, da ne priznava avtorstva korpusa, temveč resno zavira preverljivost in ponovljivost raziskav, saj iz reference sploh ni razvidno, za katero različico korpusa konkretno gre, saj po eni strani obstaja več spletnih korpusov španščine, ki so jih ustvarili različni avtorji, po drugi pa so bili številni med njimi izdelani v več različicah in vsebujejo različno gradivo. Ko smo korpus želeli preveriti v konkordančniku SketchEngine, na katerega nas referenca napoti, ga nismo našli, saj konkordančnik na dan preverjanja<sup>20</sup> ponuja dva španska korpusa tega tipa: Spanish Web Corpus oz. Spanish WaC (Sharoff, 2006) in Spanish Web 2011 oz. esTenTen11 (Kilgarriff in Renau, 2013). Tu je potrebno poudariti, da odgovornost za ustrezno navajanje virov ne leži samo na strani avtorjev prispevkov, temveč tudi avtorjev virov, ki bi vsem uporabnikom prvi morali zagotoviti ustrezno spremno dokumentacijo o korpusu, vključno z navodili za citiranje, tako ključnega prispevka o viru kot tudi navajanje korpusa v konkordančniku in korpusa kot podatkovne zbirke. Veliko razvijalcev virov tega še vedno ne omogoča, zato je ozaveščanje nujno potrebno tudi pri tej ciljni skupini.

#### 4 DISKUSIJA

Kot je pokazala analiza, je trenutno stanje na področju navajanja virov v slovenskem jezikoslovju vse prej kot idealno, saj so navodila avtorjem za področje elektronskih jezikovnih virov zelo raznolika, ponekod zastarela, pri precejšnjem številu revij in zbornikov pa celo manjkajo. Posledično so tudi prakse navajanja virov tako med kot tudi znotraj posameznih znanstvenih publikacij

---

19 [https://www.clarin.si/kontext/first\\_form?corpname=gos\\_vl](https://www.clarin.si/kontext/first_form?corpname=gos_vl)

20 <https://www.sketchengine.eu> (15. 4. 2018)

zelo heterogene. Še bolj pa je zaskrbljujoč podatek, da skoraj petina objavljenih prispevkov v uglednih znanstvenih revijah in zbornikih uporabljenih virov sploh ne navaja.

Da bi skušali prispevati k izboljšanju stanja, v nadaljevanju prispevka oblikujemo priporočila, ki temeljijo na mednarodnih iniciativah in predlogih, kako izboljšati citiranost raziskovalnih podatkov. Konkretno sledimo osmim »austinskim načelom« citiranja podatkov v jezikoslovju (Berez-Kroeker idr., 2017). Za vsako od načel podamo ime in prevod definicije, nakar ga umestimo v Slovenijo z analizo stanja in predlogi za ukrepe, kako jih realizirati.

#### 4.1 Pomembnost

*Podatki bi morali biti legitimen rezultat raziskav in jih je obvezno citirati. Citati podatkov bi za merjenje raziskovalčeve znanstvene uspešnosti morali biti enako pomembni, kot so citati objav.*

Rezultati analize so pokazali, da je to načelo v Sloveniji z manjšimi izjemami zelo slabo zastopano. Za njegovo udejanjanje sta ključna dva ukrepa. Prvi je izobraževanje, predvsem študentov, kjer bi njihovi profesorji oz. mentorji morali vztrajati pri korektnem citiranju podatkov v seminarskih nalogah, zaključnih delih in znanstvenih objavah. Drugi ukrep bi, kot predlagajo Berez-Kroeker idr. (2018), morali izvesti uredniški odbori revij in programski odbori konferenc tako, da bi v navodila za avtorje dodali navodila za ustrezno citiranje jezikoslovnih podatkov, tako kot so jih predhodno za spletne vire. Posebej poudarjamo, da je dobrim praksam navajanja raziskovalnih podatkov v slovenskih publikacijah že posvečen priročnik *Priprava raziskovalnih podatkov za odprt dostop* (Štebe, Bezjak in Vipavc Brvar, 2015). Avtorji priporočajo, »da se v seznamu uporabljene literature podatke navaja s polno navedbo avtorja oz. avtorjev, naslova, mesta dostopa do podatkov in stalnega identifikatorja, skladno z oblikovnimi zahtevami znanstvene revije« (2015, str. 13; naš poudarek).

Tu so ključni naslovniki *Slavistična revija* kot revija s posebnega seznama ARRS, konferenca oz. monografija *Obdobja*, kot tudi konferenca *Jezikovne tehnologije in digitalna humanistika*, ki bi na tem področju v skladu s svojim poslanstvom morala orati ledino, podobno, kot to že vrsto let počne mednarodna konferenca LREC v okviru iniciative *Identificiraj, opiši in deli*

*svoj jezikovni vir*, ki je od 2010 obvezni del postopka oddaje konferenčnega prispevka in so jo že prevzele tudi nekatere druge mednarodne konference. Postopek identifikacije poteka s pomočjo uporabe trajnega enkratnega identifikatorja International Standard Language Resource Number (ISLRN, [www.islrn.org](http://www.islrn.org)). Od 2014 imajo avtorji prispevkov na konferenci LREC poleg objave prispevka tudi možnost objave razvitega jezikovnega vira v repozitoriju »LRE Map«. <sup>21</sup> Ni odveč omeniti, da je citiranje obvezna sestavina v uvodu omenjeni Nacionalni strategiji odprtega dostopa in njenem Akcijskem načrtu, izpostavljeno pa je tudi v raznih drugih razpravah o odprtih podatkih v Sloveniji, vključno z nalogami financerja in uredništev revij.

#### **4.2 Priznanje zaslug in avtorstva**

*Citiranje podatkov bi moralo služiti priznavanju znanstvenih zaslug, normativnega in pravnega avtorstva vsem, ki so prispevali k njihovi izdelavi.*

Za priznavanje znanstvenih zaslug je v Sloveniji merodajen SICRIS, ki se za štetje citatov zanaša na Web of Science in SCOPUS. Vplivanje na štetje citatov znanstvenih podatkov je tako izven dometa pričujočega članka.

Lahko pa v Sloveniji vplivamo na to, kako se točkujejo objave znanstvenih podatkov, in sicer prek predlogov raziskovalni infrastrukturi Osrednjih specializiranih informacijskih centrov (OSIC). <sup>22</sup> Trenutno v sistemu COBISS že obstaja rubrika »2.20 Zaključena znanstvena zbirka podatkov ali korpus«, vendar ima takšen vnos priznanih samo 5 točk. Bistveno bolje so lahko točkovane objave pod to rubriko v primerih, ko je vir podatkov naveden v seznamu »Zaključene znanstvene zbirke podatkov, ki se upoštevajo pri kategorizaciji znanstvenih publikacij (BIBLIO-D)«. <sup>23</sup> Trenutno je na tem seznamu samo Arhiv družboslovnih podatkov (ADP). Pomembne objave v ADP tako dobijo 30 točk (Vončina, 2016), če so deponirani podatki s strani komisije ADP ocenjeni kot zelo pomembni.

Za jezikoslovne podatke bi bilo potrebno tudi repozitorij CLARIN.SI uvrstiti na seznam BIBLO-D, kar pa bi poleg samega predloga komisiji OSIC zahtevalo

---

<sup>21</sup> <http://lremap.elra.info/>

<sup>22</sup> <https://www.arrs.si/sl/infra/osic/predstavitev.asp>

<sup>23</sup> <http://home.izum.si/COBISS/bibliografije/Kateg-znan-zbirke.html>



tudi podrobnejša navodila za vnašanje virov, kot tudi ustanovitev komisije za vrednotenje vnesenih virov. Vse to pa seveda tudi zahteva precejšen vložek dela in s tem financiranje CLARIN.SI.

#### 4.3 Dokazi

*V znanstvenih objavah bi morali biti podatki ustrezno citirani poudarjeno, kjer neka trditev sloni na podatkih.*

Podobno kot za 1. načelo (pomembnost) je tudi tu ključno izobraževanje, navodila za avtorje in uredniška politika publikacij.

#### 4.4 Nedvoumna identifikacija

*Citiranje podatkov naj bi vsebovalo trajno metodo identifikacije, primerno za strojno obdelavo, mednarodno edinstveno in široko sprejeto v skupnosti.*

Ta pogoj je v veliki meri že realiziran v sklopu repozitorija CLARIN.SI. Vsak vir ima trajni identifikator PID (*persistent identifier*) po sistemu »handle«, na vrhu strani pa je jasno napisano, kako naj se vir citira, pri čemer navedek vsebuje tudi identifikator handle. Repozitorij CLARIN.SI prav tako podpira izvoz metapodatkov po shemi Dublin Core, ki jih žanje več agregatorjev: CLARIN VLO,<sup>24</sup> OpenAIRE,<sup>25</sup> re3data<sup>26</sup> in OAI.<sup>27</sup>

Večina ostalih ponudnikov jezikovnih virov v Sloveniji, kot npr. slovarski portal Fran<sup>28</sup> na ZRC SAZU, Termania<sup>29</sup> podjetja Amebis, d.o.o, ali stran z viri CJVT<sup>30</sup> ne ponujajo trajnih identifikatorjev. Izjema tu sta digitalna knjižnica dLib<sup>31</sup> NUK, kjer je vsaki publikaciji pripisan trajni identifikator po sistemu URN, ter digitalna knjižnica Sistory<sup>32</sup> INZ, ki, tako kot CLARIN.SI, tudi uporablja sistem handle.

---

24 <https://vlo.clarin.eu/>

25 <https://www.openaire.eu/>

26 <https://www.re3data.org/repository/r3d100011922>

27 <http://www.language-archives.org/archive/clarin.si>

28 <http://www.fran.si/>

29 <https://www.termania.net/>

30 <https://viri.cjvt.si/>

31 <http://www.dlib.si/>

32 <https://www.sistory.si/>

V jezikoslovju je poleg navajanja vira kot podatkovne zbirke pomembno tudi navajanje poizvedbe v konkordančniku. Konkordančniki CLARIN.SI, kot tudi konkordančniki projekta Sporazumevanja v slovenskem jeziku (torej konkordančnik za Gigafido,<sup>33</sup> Kres,<sup>34</sup> itd.) so vsi narejeni po principu REST, da torej URL poizvedbe zadošča za ponovno in enako poizvedbo.<sup>35</sup> Z drugimi besedami, URL, ki ga dobimo po poizvedbi in prikazu rezultatov, je mogoče shraniti in prek njega ponovno dobiti iste rezultate. Tu velja še opomba, da so takšni URL-ji tipično zelo dolgi in zato neprimerni ali vsaj težavni za citiranje. Vendar pa za krajšanje URL-je obstaja več spletnih storitev, od katerih je posebej zanimiva shortref.org,<sup>36</sup> ki jo ponuja češki LINDAT/CLARIN. Za razliko od drugih krajševalnikov ponuja shortref.org opis poizvedbe, kot skrajšani URL pa vrne trajni identifikator po sistemu handle.

#### 4.5 Dostop

*Citiranje podatkov naj bi pripomoglo k dostopu do samih podatkov in do povezanih metapodatkov, dokumentacije, programske opreme in drugih materialov, ki so potrebni, da tako ljudje kot računalniki te podatke lahko informirano uporabljajo.*

Ta zahteva je tudi že v veliki meri realizirana v sklopu repozitorija CLARIN.SI, saj vsak vnos vsebuje tako metapodatke kot tudi same podatke, ki so pred vključitvijo v repozitorij preverjeni s strani urednikov.

#### 4.6 Trajnost

*Enoznačni identifikatorji in metapodatki, ki opisujejo podatke, morajo biti trajni, celo bolj kot sami podatki.*

Repozitorij CLARIN.SI je del slovenske in evropske infrastrukture, domuje pa na Institutu »Jožef Stefan«, ki ima visoko razvito računalniško infrastrukturo. Oboje v največji možni meri ponuja garancijo za dolgotrajnost (meta)podatkov, deponiranih v repozitoriju. K trajnosti metapodatkov pa prispeva tudi že omenjeno dejstvo, da se le-ti redno izvažajo v več spletnih

---

33 <http://www.gigafida.net/>

34 <http://www.korpus-kres.net/>

35 Seveda, če se medtem ni spremenil korpus.

36 <http://shortref.org/>

agregatorjev. CLARIN.SI izvaja tudi redno testiranje skladnosti in povezljivosti podatkov.

#### **4.7 Specifičnost in preverljivost**

*Citiranje podatkov naj bi pripomoglo identifikaciji, dostopu in preverjanju specifičnih podatkov, ki podpirajo neko trditev. Citiranje ali metapodatki citiranja naj bi vsebovali podatke o izvoru in stabilnosti v zadostni meri, da omogočijo preverbo, da je specifičen časovni okvir, različica ali del podatkov, ki so bili naknadno prevzeti, enak kot podatki, ki so bili izvorno citirani.*

Tudi tu CLARIN.SI v veliki meri zadošča temu načelu. Vnosi v repozitorij se ne spreminjajo, v primeru dopoljenih ali popravljeni podatkov se ti vpišejo v nov vnos, vendar z medsebojno povezavo med starim in novim vnosom. Posebej velja poudariti, da je nadzor nad različnimi verzijami, ki ga repozitorij CLARIN.SI omogoča, eno izmed priporočil ustreznega digitalnega skrbništva jezikovnih podatkov (npr. Štebe, Bezjak in Vipavc Brvar, 2015, str. 6). Številni viri so zapisani po priporočilih TEI, ki tipično vsebujejo bogate metapodatke, s katerimi je mogoče podrobno določiti zelene izseke virov.

#### **4.8 Interoperabilnost in fleksibilnost**

*Metode za citiranje podatkov naj bi bile fleksibilne v zadostni meri, da omogočajo različne prakse med skupnostmi, vendar se ne smejo razlikovati v tolikšni meri, da bi to ogrozilo interoperabilnost praks citiranja podatkov med skupnostmi.*

Repozitorij CLARIN.SI mdr. navaja naslov in avtorje vsakega vira ter na vrhu dostopne strani vira točno definira, kako je vir potrebno citirati.

## **5 ZAKLJUČEK**

V prispevku smo predstavili rezultate razširjene študije, s katero smo preverjali stanje citiranja jezikovnih podatkov, predvsem korpusov, v najpomembnejših slovenskih znanstvenih revijah in zbornikih, ki so bili objavljeni v dveh različnih obdobjih, ki skupaj zajemata zadnjih sedem let. V raziskavi smo kvantitativno in kvalitativno analizirali obseg in način navajanja virov, izvedli pa smo tudi pregled navodil za avtorje znanstvenih revij in zbornikov, vključenih v raziskavo. Rezultati študije kažejo, da stanje ni zavidljivo

in si je zato potrebno prizadevati za ozaveščanje, izobraževanje in podporo v skupnosti.

Po opravljeni analizi ugotavljamo, da na jezikovnih virih temelji manj kot petina vseh objavljenih prispevkov, kar je glede na stopnjo razvitosti in razpoložljivosti jezikovnih virov za slovenščino malo in kaže na ne vključenost skupnosti, ki vire razvija, v osrednjo jezikoslovno raziskovalno skupnost pri nas. V prispevkih, ki temeljijo na uporabi jezikovnih virov, pa viri v skoraj petini primerov sploh niso ustrezno citirani. To kaže na pomanjkanje ozaveščenosti jezikoslovcev o pomenu navajanja vseh virov v znanstvenem publiciranju.

S prispevkom, v katerem smo predlagali načela za ustrezno citiranje digitalnih jezikovnih virov, ki temeljijo na mednarodnih poročilih, smo storili korak v tej smeri. Brez tega onemogočamo preverljivost, ponovljivost in nadgrajevanje prejšnjih raziskav, ki so osnovni temelji odprte znanosti. Korektno citiranje jezikovnih virov pa je pomembno tudi zato, ker je v njihov razvoj potrebno vložiti izjemno veliko truda in časa, znanstveni citati pa ostajajo daleč najpomembnejši indikator znanstvene uspešnosti.

Z ozaveščanjem in izobraževanjem bi bilo potrebno začeti že v okviru univerzitetnih študijskih programov in poskrbeti za ustrezne smernice za navajanje virov tudi v tem kontekstu. Na področju ozaveščanja skupnosti aktivnih raziskovalcev pa bi z izobraževalnimi dogodki in spletnimi gradivi veliko lahko pripomogla nacionalna raziskovalna infrastruktura CLARIN.SI.

Videli smo, da trenutna navodila avtorjem v vseh pregledanih revijah bistveno premalo natančno opisujejo dobro prakso citiranja jezikovnih virov in da v veliki meri ravno to vodi do neoptimalnega stanja, ki smo ga opisali za obdobje 2013–2019. Posledično bi bilo potrebno vzpostaviti dialog z uredništvom, ki imajo neposreden stik z raziskovalci in tako tudi veliko moč pri promoviranju dobrih praks citiranja jezikovnih virov, zaradi česar so eni najpomembnejših akterjev pri vzpostavljanju in zagotavljanju dobrih praks za citiranje. Ta proces se je že začel v okviru RDA Node Slovenia.<sup>37</sup> Projekt vodi ADP, ki se mdr. trudi za »razvoj pilota politike znanstvenih založb glede obveznosti predaje raziskovalnih podatkov k objavljenim znanstvenim člankom in aktivna promocija obveze za dostop, ocenjevanje in citiranje raziskovalnih podatkov«. Za

---

<sup>37</sup> <https://www.rd-alliance.org/groups/rda-slovenia>

te namene je projekt že vzpostavil smernice (Štebe, Dolinar in Bezjak, 2019), kot tudi dialog z nekaterimi raziskovalnimi revijami v Sloveniji z namenom, da služijo kot piloti za vzpostavljanje boljših navodil avtorjev, ki bodo mdr. poudarjale pomembnost navajanja virov s trajnimi identifikatorji, kot je povezava handle, ter navedbo avtorjev, pri čemer je pričujoča revija eden od teh pilotov, predvidene spremembe v navodilih za avtorje pa so bile tudi predstavljene na konferenci »Znanstvene revije Slovenije in raziskovalni podatki«.<sup>38</sup>

Poleg sprememb navodil za avtorje revij in konferenčnih zbornikov bi bilo prav tako nujno poskrbeti tudi za ozaveščanje razvijalcev virov, ki lahko k ustreznemu citiranju veliko pripomorejo tako, da ustrezno deponirajo in dokumentirajo svoje vire. Za odprto znanost namreč še zdaleč ni dovolj, da nek vir obstaja in je dostopen, temveč mora biti tudi opremljen z vso potrebno spremno dokumentacijo, med katero vključujemo tudi navodila za citiranje. Opuščanje teh praks že na prvem koraku zavira ustrezno citiranje, avtorji prispevkov, ki tovrstne nepopolne vire uporabljajo, pa so pri tem nemočni. Tudi k temu bi lahko z nudenjem ustrezne dokumentacije, izobraževanj in tehnične podpore veliko doprinesla nacionalna raziskovalna infrastruktura CLARIN.SI.

V prihodnje bi bilo raziskavo zanimivo razširiti na jezikovne vire s področja eksperimentalnega in računalniškega jezikoslovja, ki jezikovne vire uporabljajo kot podatkovne množice, zaradi česar se njihovi interesi, pa tudi potrebe, razlikujejo od skupnosti, ki smo se jim posvetili v tej raziskavi.

### **Zahvala**

Avtorji se zahvaljujejo anonimnima recenzentoma za koristne pripombe in nasvete. Raziskava, opisana v prispevku, je bila opravljena v okviru raziskovalnih infrastruktur za jezikovne vire in orodja CLARIN.SI in CLARIN ERIC.

### **LITERATURA**

- Arhar Holdt, Š. in Dobrovoljc, K. (2016). Vrednost korpusa *Janes* za slovensko normativistiko. *Slovenščina 2.0*, 4(2), 1–37. doi: 10.4312/slo2.0.2016.2.1-37
- Arhar Holdt, Š. in Čibej, J. (2018). Morphological Patterns in the Sloleks Lexicon of Slovene: An Initial Set of Patterns for Nouns. *Slovenščina 2.0*, 6(2), 33–66. doi: 10.4312/slo2.0.2018.2.33-66

---

38 <https://www.adp.fdv.uni-lj.si/dogodki/znanstvene-revije-slovenije-raziskovalni-podatki/>

- Arias-Badia, B., Bernal, E. in Alonso, A. (2014). An online Spanish Learners' dictionary: the Daele project. *Slovenščina 2.0*, 2(2), 53–71. doi: 10.4312/slo2.0.2014.2.53-71
- Atelšek, S. (2019). Navajanje prevzetih jezikoslovnih terminov in celovitost pojmovnih skupin v Cigaletovi *Znanstveni terminologiji* (1880). *Jezikoslovní zapiski*, 25(1), 67–82. doi: 10.3986/jz.v25i1.7566
- Bálint Čeh, J. in Kosem, I. (2017). Prvi koraki do novega velikega slovensko-madžarskega slovarja: analiza relevantnih dvojezičnih virov. *Slovenščina 2.0*, 5(2), 113–150. doi: 10.4312/slo2.0.2017.2.113-150
- Berez-Kroeker, A. L., Gawne, L., Holton, G., Smythe Kung, S., Pulsifer, P. in Collister, L. B. (2017). The Data Citation and Attribution in Linguistics Group, & the Linguistics Data Interest Group. The Austin Principles of Data Citation in Linguistics (Version 0.1). Dostopno prek <http://site.uit.no/linguisticsdatacitation/austinprinciples>
- Berez-Kroeker, A. L., Gawne, L., Smythe Kung, S., Kelly, B. F., Heston, T., Holton, G., Pulsifer, P., Beaver, D. I., Chelliah, S., Dubinsky, S., Meier, R. P., Thieberger, N., Rice, K. in Woodbury, A. C. (2018). Reproducible research in linguistics: A position statement on data citation and attribution in our field. *Linguistics*, 56(1), 1–18. doi: 10.1515/ling-2017-0032
- Data Citation Synthesis Group. (2014). *Joint Declaration of Data Citation Principles*. Martone, M. (ur.). San Diego CA: FORCE11. doi: 10.25490/a97f-egyk
- Dobrovoljc, K. (2018a). Raba tipično govorjenih diskurzivnih označevalcev na spletu. *Slavistična revija*, 66(4), 497–513. Dostopno prek <https://srl.si/ojs/srl/article/view/2018-4-1-6>
- Dobrovoljc, K. (2018b). Formulaicity in Slovene. *Slovenščina 2.0*, 6(2), 67–95. doi: 10.4312/slo2.0.2018.2.67-95
- Dobrovoljc, K., Krek, S., Holozan, P., Erjavec, T., Romih, M., Arhar Holdt, Š., Čibej, J., Krsnik, L. in Robnik-Šikonja, M. (2019). *Morphological lexicon Sloleks 2.0*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1230>
- Erjavec, T. (2009). Odprtost jezikovnih virov za slovenščino. V M. Stabej (ur.), *Simpozij OBDOBJA 28*. Dostopno prek <http://centerslo.si/wp-content/uploads/2015/10/28-Erjavec.pdf>

- Erjavec, T., Fišer, D., Krek, S. in Ledinek, N. (2010). The JOS Linguistically Tagged Corpus of Slovene. V *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Dostopno prek <http://www.lrec-conf.org/proceedings/lrec2010/summaries/139.html>
- Erjavec, T. (2014). *Digital library and corpus of historical Slovene IMP 1.1*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1031>
- Erjavec, T. (2015a). The IMP historical Slovene language resources. *Language Resources and Evaluation*, 49, 753–775. doi: 10.1007/s10579-015-9294-7
- Erjavec, T. (2015b). *Reference corpus of historical Slovene goozook 1.2*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1025>
- European Commission. (2012). Towards better access to scientific information: Boosting the benefits of public investments in research. Dostopno prek [http://ec.europa.eu/research/science-society/document\\_library/pdf\\_06/era-communication-towards-better-access-to-scientific-information\\_en.pdf](http://ec.europa.eu/research/science-society/document_library/pdf_06/era-communication-towards-better-access-to-scientific-information_en.pdf)
- Fišer, D., Lenardič, J. in Erjavec, T. (2018). Citiranje jezikoslovnih podatkov v slovenskih znanstvenih objavah: stanje in priporočila. V D. Fišer in A. Pančur (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika 2018* (str. 77–84). Univerza v Ljubljani, Filozofska fakulteta.
- Furlan, M. (2018). O govejem lastnem imenu Hrdagata in kletvici (h)ardigata. *Jezikoslovni zapiski*, 24(1), 131–141. doi: 10.3986/JZ.24.1.6938
- Haspelmath, M. (2014). The Generic Style Rules for Linguistics. *Zenodo*. doi: 10.5281/zenodo.253501
- Hudeček, K. in Mihaljević, M. (2019). Hrvatsko mocijsko nazivlje. *Jezikoslovni zapiski*, 25(1), 107–126. doi: 10.3986/jz.v25i1.7569
- Jakop, N. (2014). Leksikalizacija prostorskih razmerij v slovenščini: jezikovnopragmatični vidik. *Slavistična revija*, 62(3), 353–362. Dostopno prek [https://srl.si/sql\\_pdf/SRL\\_2014\\_3\\_08.pdf](https://srl.si/sql_pdf/SRL_2014_3_08.pdf)
- Jelovšek, A. in Erjavec, T. (2019). A corpus-based study of 16th-century Slovene clitics and clitic-like elements. *Slovene Linguistic Studies*, 12, 3–19. Dostopno prek <http://hdl.handle.net/1808/29671>
- Kačič, Z., Horvat, B., Zögling Markuš, A., Veronik, R., Rojc, M., Žgank, A., Sepesy Maučec, M. in Rotovnik, T. (2002). *SNABI database for continuous*

- speech recognition 1.2*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1051>
- Kilgarriff, A. in Renau, I. (2013). esTenTen, a vast webcorpus of Peninsular and American Spanish. *Procedia-Social and Behavioral Sciences*, 95, 12–19. doi: 10.1016/j.sbspro.2013.10.617
- Krek, S., Erjavec, T., Dobrovoljc, K., Holz, N., Ledinek, N. in Može, S. (2015). Training corpus ssj500k 1.4 Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1052>
- Krek, S., Dobrovoljc, K., Erjavec, T., Može, S., Ledinek, N., Holz, N., Zupan, K., Gantar, P., Kuzman, T., Čibej, J., Arhar Holdt, Š., Kavčič, T., Škrjanec, I., Marko, D., Jezeršek, L. in Zajc, A. (2019). *Training corpus ssj500k 2.2*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1210>
- Krvina, D. (2019). Zaporednost dejanj in njen vpliv na rabo glagolskega vida v slovenščini. *Slovene Linguistic Studies*, 12, 75–83. doi: 10.3986/sjls.12.1.05
- Kulčar, M. (2018). Povezanost vida in vezljivosti pri netvorjenih in predponskoobrazilno tvorjenih glagolih. *Jezikoslovni zapiski*, 24(1), 45–62. doi: 10.3986/JZ.24.1.6932
- Ljubešič, N., Fišer, D. in Erjavec, T. (2014). TweetCaT: A tool for building Twitter corpora of smaller languages. V N. Calzolari (ur.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)* (str. 2279–2283). Reykjavik, Islandija.
- Ljubešič, N. in Klubička, F. (2016). *Croatian web corpus hrWaC 2.1*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1064>
- Ljubešič, N., Miličević Petrović, M. in Samardžić, T. (2019). Jezična akomodacija na Twitteru: primjer Srbije. *Slavistična revija*, 67(1), 87–106. Dostopno prek <https://srl.si/ojs/srl/article/view/2019-1-1-6>
- Logar Berginc, N., Grčar, M., Brakus, M., Erjavec, T., Arhar Holdt, Š. in Krek, S. (2012). *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in cc-KRES: gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko, Fakulteta za družbene vede. Dostopno prek <https://www.fdv.uni-lj.si/docs/default-source/zalozba/pages-from-logar-et-al---korpusi.pdf?sfvrsn=2>



- Logar Berginc, N., Erjavec, T., Krek, S., Grčar, M. in Holozan, P. (2013). *Written corpus ccKres 1.0*. Slovenian language resource repository CLARIN. SI. Dostopno prek <http://hdl.handle.net/11356/1034>
- Logar Berginc, N., Gantar, P. in Kosem, I. (2014). Collocations and examples of use: a lexical-semantic approach to terminology. *Slovenščina 2.0*, 2(1), 41–61. doi: 10.4312/slo2.0.2014.1.41-61
- Marvin, T., Derganc, J., Beguš, S. in Battelino, S. (2018). Word Selection in the Slovenian Sentence Matrix Test for Speech Audiometry. V D. Fišer in A. Pančur (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika 2018* (str. 181–187). Univerza v Ljubljani, Filozofska fakulteta.
- Marvin, T., Battelino, S., Beguš, S. in Derganc, J. (2019). Porazdelitev fonemov v slovenščini in izdelava matričnega testa za govorno avdiometrijo. *Slavistična revija*, 67(4), 537–550. Dostopno prek <https://srl.si/ojs/srl/article/view/2019-4-1-1>
- Meterc, M. (2013). Antonimija enako motiviranih paremioloških enot (primeri iz slovenščine in slovaščine). *Slavistična revija*, 61(2), 361–376. Dostopno prek [https://srl.si/sql\\_pdf/SRL\\_2013\\_2\\_02.pdf](https://srl.si/sql_pdf/SRL_2013_2_02.pdf)
- Orel, I. (2019). Ženske dvojninske glagolske oblike v starejšem slovenskem knjižnem jeziku. *Slavistična revija*, 67(2), 273–280. Dostopno prek <https://srl.si/ojs/srl/article/view/2019-2-1-15>
- Petrič, T. (2019). Modal Particles in German Declarative Sentences and their Slovenian Counterparts. *Linguistica*, 59(1), 235–251. doi: 10.4312/linguistica.59.1.235-251
- Pisanski Peterlin, A. in Mikolič Južnjič, T. (2018). Subject Personal Pronouns in Slovene: Pragmatic Aspects of a Grammatical Category. *Slovenščina 2.0*, 6(2), 127–153. doi: 10.4312/slo2.0.2018.2.127-153
- Polajnar, J. (2013). Neprodani in trdni. Ja, seveda, potem pa svizec ... Osamosvajanje oglasnih sloganov v slovenskem jeziku. *Jezik in slovastvo*, 58(3), 3–19. Dostopno prek <https://www.jezikinslovstvo.com/pdf.php?part=2013|3|3%E2%80%9319>
- Pori, E. in Kosem, I. (2018). In the Search of Lexicographically Relevant Collocation: The Example of Grammatical Relations Containing Adverbs. *Slovenščina 2.0*, 6(2), 154–185. doi: 10.4312/slo2.0.2018.2.154-185
- Rath, A. (2019). Anmerkung zur slowenischen Klitikakette (naslonski niz). *Slovene Linguistic Studies*, 12, 95–112. doi: 10.3986/sjls.12.1.06

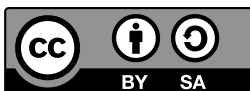
- Rozman, T., Stritar Kučuk, M., Kosem, I., Krek, S., Krapš Vodopivec, I., Arhar Holdt, Š. in Stabej, M. (2013). *Learners' corpus Šolar 1.0*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1036>
- Rozman, T., Arhar Holdt, Š., Pollak, S. in Kosem, I. (2018). Kolokacije v korpusu Šolar. *Jezik in slovnstvo*, 63(2–3), 117–128. Dostopno prek <https://www.jezikinslovnstvo.com/pdf.php?part=2018|2-3|117-128>.
- Sharoff, S. (2006). Creating general-purpose corpora using automated search engine queries. *WaCky! Working papers on the Web as Corpus*. Dostopno prek <http://wackybook.sslmit.unibo.it/pdfs/sharoff.pdf>
- Stopar, A. in Ilc, G. (2019). Stilistična (ne)zaznamovanost moških in ženskih poimenovalnih parov za poklice v angleščini in slovenščini. *Slavistična revija*, 67(2), 333–342. Dostopno prek <https://srl.si/ojs/srl/article/view/2019-2-1-21>
- Stramljič Breznik, I. (2018). Ženske ne povedo nič pametnega: jezikovnokorpusna analiza stereotipa. *Jezikoslovni zapiski*, 24(1), 27–44. doi: 10.3986/JZ.24.1.6931
- Štebe, J., Bezjak, S. in Vipavc Brvar, I. (2015). *Priprava raziskovalnih podatkov za odprto dostop. Priročnik za raziskovalce*. Ljubljana: Založba FDV. Dostopno prek <https://www.dlib.si/details/URN:NBN:SI:DOC-o6SLBVXX>
- Štebe, J., Dolinar, M. in Bezjak, S. (2019). *Smernice za oblikovanje politik znanstvenih založb glede navajanja raziskovalnih podatkov v znanstvenih publikacijah in zagotavljanja dostopa do primarnih podatkov, uporabljernih v člankih* (Verzija 2.3.). Dostopno prek [https://www.rd-alliance.org/system/files/documents/Smernice\\_za\\_razvoj\\_politike\\_zalo%C5%BEB\\_RDA\\_Slovenija\\_V2\\_3.pdf](https://www.rd-alliance.org/system/files/documents/Smernice_za_razvoj_politike_zalo%C5%BEB_RDA_Slovenija_V2_3.pdf)
- Štumberger, S. (2015). Slovaropisna obravnava novejšje leksike. *Slovene Linguistic Studies*, 10, 153–166. Dostopno prek <https://ojs.zrc-sazu.si/sjls/article/view/7365>
- Trivunović, E. (2019). Diahrono raziskovanje biblijskih in izbiblijskih frazemov. *Jezikoslovni zapiski*, 25(2), 47–61. doi: 10.3986/JZ.25.2.3
- Uhlik, M. in Žele, A. (2018). Brezosebne zgradbe v slovenščini: kontrastiva z drugimi južnoslovanskimi jeziki in ruščino. *Jezikoslovni zapiski*, 24(2), 99–112. doi: 10.3986/jz.v24i2.7112
- Verdonik, D., Potočnik, T., Sepesy Maučec, M. in Erjavec, T. (2016). *Spoken corpus Gos VideoLectures 1.0 (transcription)*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1069>

- Verdonik, D., Potočnik, T., Sepesy Mauček, M. in Erjavec, T. (2017). *Spoken corpus Gos VideoLectures 2.0 (transcription)*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1158>
- Vidovič Muha, A. (2015). Propozicija v funkcijski strukturi stavčne povedi – vprašanje besednih vrst (poudarek na povedkovniku in členu). *Slavistična revija*, 63(4), 389–406. Dostopno prek [https://srl.si/sql\\_pdf/SRL\\_2015\\_4\\_04.pdf](https://srl.si/sql_pdf/SRL_2015_4_04.pdf)
- Vončina, M. (2016). Zaključena znanstvena zbirka podatkov – primeri katalogizacije in Sicris vrednotenja. [Delavnica ADP, 26. 10. 2016.] Dostopno prek [https://www.adp.fdv.uni-lj.si/adp\\_delavnica\\_okt2016/presentations/2016\\_Mira-Voncina\\_Znanstvena\\_zbirka\\_podatkov.pdf](https://www.adp.fdv.uni-lj.si/adp_delavnica_okt2016/presentations/2016_Mira-Voncina_Znanstvena_zbirka_podatkov.pdf)
- Zwitter Vitez, A., Zemljarič Miklavčič, J., Krek, S., Stabej, M. in Erjavec, T. (2013). *Spoken corpus Gos 1.0*. Slovenian language resource repository CLARIN.SI. Dostopno prek <http://hdl.handle.net/11356/1040>
- Zwitter Vitez, A. (2018). Enota analize spontanega govora: interakcija prozodije, pragmatike in skladnje. *Jezik in slovastvo*, 63(2–3), 157–175. Dostopno prek <https://www.jezikinslovstvo.com/pdf.php?part=2018|2-3|157-175>
- Žele, A. (2014). Členki tudi kot vnašalniki novih prostorskih razmerij v obstoječe sporočilo. *Slavistična revija*, 62(3), 321–330. Dostopno prek [https://srl.si/sql\\_pdf/SRL\\_2014\\_3\\_05.pdf](https://srl.si/sql_pdf/SRL_2014_3_05.pdf)
- Žele, A. (2018). O aktualnostnočlenitveni stavi v slovenščini. *Jezik in slovastvo*, 63(2–3), 59–73.
- Žitnik, S., Šubelj, L. in Bajec, M. (2014). SkipCor: Skip-mention coreference resolution using linear-chain conditional random fields. *PloS one*, 9(6), e100101. doi: 10.1371/journal.pone.0100101
- Žitnik, S., Draskovic, D., Nikolić, B. in Bajec, M. (2017). nutIE—A modern open source natural language processing toolkit. *Proceedings of the 25th Telecommunication Forum (TELFOR)*, 1–4. doi: 10.1109/TELFOR.2017.8249486
- Žitnik, S. in Bajec, M. (2018). Coreference Resolution for Slovene on Annotated Data from coref149. *Slovenščina 2.0*, 6(1), 37–67. doi: 10.4312/slo2.0.2018.1.37-67

## LINGUISTIC DATA CITATION IN SLOVENE SCIENTIFIC PUBLICATIONS: ANALYSIS AND RECOMMENDATIONS

Open science is based on freely and openly available scientific publications and data. The latter enable the verification and improvement of previous research. In the context of language technologies and manually annotated language resources, they also enable training of new text processing tools. However, just like scientific publications, research data need to be properly cited, as only this makes reproducibility of experiments possible and is the most important indicator of how interesting and useful researchers' work is in the community and plays a major role in their success with research grant proposals and career trajectory. In this paper, we survey the landscape of linguistic data, mainly (mainly language corpora) citation in six leading Slovene scientific journals (*Jezik in slovastvo*, *Slavistična revija*, *Slovenščina 2.0*, *Linguistica*, *Slovene Linguistic Studies* and *Jezikoslovni zapiski*) and in the proceedings of two scientific conferences focused on linguistics (*Jezikovne tehnologije in digitalna humanistika* and *Obdobja*) for the period of the last seven years, i.e. from 2013 to 2019. We consider 1,074 papers and analyse the results both quantitatively and qualitatively. From the quantitative perspective, we show that, overall, only about a fourth of the papers includes the use of language resources, and that in the later period (2018–2019) the use of language resources is over twice as frequent as it is in the earlier period (2013–2017). We classify the manner of language resource citation into five categories (e.g. *citing the hyperlink in the texts* or *citing the key paper about the resource*) and show that how a resource is cited is, to a large extent, dependent on the instructions for authors of the particular publication. Our qualitative analysis focuses mainly on resources deposited in the repository of the CLARIN.SI research infrastructure, where we show that they are, with few exceptions, incorrectly cited. We summarise the finding using the so-called Austin principles, show what has already been achieved in the scope of the CLARIN.SI infrastructure and propose guidelines for citing linguistic research data and how to implement them.

**Keywords:** Open Science, Research Data Citation, Language Resources, Austin Principles, Slovenian Journals and Conference Proceedings



To delo je ponujeno pod licenco Creative Commons: Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna. / This work is licensed under the Creative Commons Attribution-Share-Alike 4.0 International.

<https://creativecommons.org/licenses/by-sa/4.0/>

## CZECH GENDER LINGUISTICS: TOPICS, ATTITUDES, PERSPECTIVES

Vít KOLEK, Jana VALDROVÁ

Palacký University Olomouc

*Kolek, V., Valdrová, J. (2020): Czech gender linguistics: topics, attitudes, perspectives. Slovenščina 2.0, 8(1): 35–65.*

DOI: <https://doi.org/10.4312/slo2.0.2020.1.35-65>

The aim of the present article is to analyse the existing publications dealing with Czech gender linguistics, and attitudes of their authors. Firstly, the authors of the article present an overview of the development of the discipline, beginning with its rejection to today's acceptance. They then deal with the various ways to label people with regard to gender, and consider the formation of nouns labelling women as one of the strategies to make women visible in the language. The authors also inform readers about the psycholinguistic testing of the “generic” masculine in the Czech language and about the possibilities of Czech gender-fair language. The authors discuss the proper names of persons from the perspective of gender and personal identities, while they briefly mention translatology publications which reflect the gender perspective. Additionally, the authors present a call for further research in certain areas.

**Keywords:** Czech language, gender, generic masculine, feminization, gender-fair language, word formation, translation studies.

### 1 INTRODUCTION

In the present article, the authors introduce the existing issues related to Czech gender linguistics and highlight the individual publications and attitudes of their authors from a metalinguistic perspective. This discipline (the name of which has stabilized to the Czech term *genderová lingvistika*, ‘gender linguistics’) is called by various names. The first publications which explicitly<sup>1</sup>

---

1 Publications that are of interest for gender-linguistic research, however, were published even before this, and the authors note them in the following sections.

thematized this linguistic field (Čmejrková, 1995; Hoffmannová, 1995) used the term *feministická lingvistika*, ‘feminist linguistics’. S. Čmejrková used the term *lingvistika rodu*, ‘linguistics of (grammatical) gender’ which was supposed to compete with the term *genderová lingvistika*, already used by J. Valdřová (1997). J. Valdřová (2018a) also uses the term *lingvistika genderových a sexuálních identit* ‘linguistics of gender and sexual identities’ as an overarching term for feminist, gender and queer linguistics. Since this article is devoted to the topic of the relation between the language and gender identities, the authors refer to this field in the rest of this text as *gender linguistics* (hereinafter referred to as *GL*).

The first part of the present text introduces the general publications dealing with the field and the attitudes of individual authors. The second part focuses on the labelling of persons in the Czech language, mainly on the usage of the “generic” masculine, its empirical testing and on the suggestions for gender-fair language use within the Czech language, while the next part deals with personal names and then finally the issue of gender in translation is addressed as well.

In the framework of these thematic fields, it is possible to see theoretical approaches to the relation of language and gender, although these are often not concretized explicitly in the Czech GL discourse. Mainly in the period up to 1945, the structural linguistic analysis perspective dominated without any special regard to the social or situational context – this was *inter alia* caused by the strong influence of the Prague linguistic circle on the related scientific research. The socialist period was typical with its ideological (Marxist) approach, which was often necessary in order to have a text published. In this context, the promotion of female personal names was “explained” as an achievement of socialism (e.g. Dokulil below).

The year 1989 enabled a broader feminist reflection on language, although the lack of information on the development and state of feminist research into language is obvious up to the end of 20<sup>th</sup> century, together with the misapprehension of the analytic category of *gender* (cf. Čmejrková, 1995; Hoffmannová, 1995). Its application to empirical research brought findings regarding the connections between verbal behaviour and the perception of reality. At present queer linguistic research underdeveloped, and the status of Czech queer linguistics is researched by V. Kolek (2019).

## 2 FROM REJECTION TO ACCEPTANCE: THE CRUCIAL CZECH PUBLICATIONS

As already mentioned above, the first texts reflecting foreign feminist linguistics (hereinafter referred to as *FL*) were published by S. Čmejrková (1995) and J. Hoffmannová (1995). Both authors considered the topic of the use of gender-fair language irrelevant for Czech— they reduce the topics of FL to the making of women visible within the language, while ironizing or questioning the whole field in general.<sup>2</sup> They also purposely create inappropriate Czech equivalents to English gender-fair formulations, disqualify FL with the use of negatively expressive labellings and do not see its relevance within the Czech language environment. It is easy to imagine the effects of these ideas, coming from respected linguists, on the possible GL-related interests of the linguistic community. An argumentative analysis of both texts is presented by Kolek and Valdřová (2017).

The two articles motivated J. Valdřová to perform gender-linguistic research. Her first texts (1996, 1997) introduced the proposed gender-fair language to the public, while the author then elaborated on the topic in further works (e.g. Valdřová, 2001, 2005a, 2010, 2013). F. Daneš (1997) reacted to J. Valdřová, considering the term *gender* a feminist construct, and the proposed gender-fair language “*násilné brusičství*” (‘forced purism’), with an appeal to common sense (Daneš, 1997, p. 258). In the same year, S. Čmejrková published another FL-related article; she, once again, explicitly and ironically distances herself from FL while rejecting the thesis that language not only reflects reality, but also constructs it. Apart from the topic of the generic masculine and the subsequent “*revize patriarchálního paradigmatu*” (‘revision of patriarchal paradigm’; Čmejrková, 1997, p. 147), the author also mentions the issue of the formation of female surnames (see Chapter 4.2.) and the analysis of female magazines and advertisements as FL topics.

In 1998, the first dissertation dealing with Czech GL was defended (Valdřová, 1998) at Masaryk University. Based on the contrastive German-Czech comparison, the author dealt with the possibilities of making women visible in the language. The validity of gender-fair language was then confirmed by Czech law

---

2 It is, however, necessary to consider the lack of information in times before the spread of the Internet.



– Act no. 167/1999 Coll., which, *inter alia*, demands the offering of job opportunities without sex and gender discrimination, and this resulted in more gender-fair formulations of jobs. After numerous language analyses and publications on gender-fair language principles (see below), J. Valdrová (2015a) asked whether it is possible to speak about the existence of FL in Czech linguistics: she briefly summarizes the previous GL developments and deals with the language constructs of gender, the consequences of using the “generic” masculine in oral presentations and its alternatives, proper names from the perspective of gender (see below), and so on. She states that Czech FL (or GL) is still in its early days in 2015, but notes that interest in this field is growing, as manifested in the interest of the public in gender-fair language and in the increasing number of students’ theses dealing with the topic of *gender and language*.<sup>3</sup>

This demand was met by the *Nový encyklopedický slovník češtiny* (*New Czech Encyclopaedic Dictionary*’, eds. Karlík, Nekula and Pleskalová, 2016) which contains entries such as *biologický rod X mluvnický (gramatický) rod, feministická lingvistika, gender, genderová lingvistika, genderově vyvážené vyjadřování, generické maskulinum* (*‘biological sex X grammatical gender, feminist linguistics, gender, gender linguistics, gender-fair language, generic masculine’*), authored by J. Valdrová (together with M. Nekula for the entry for *feministická lingvistika, ‘feminist linguistics’*). The entries inform readers about the historic development of FL and GL in both foreign countries and in the Czech Republic, while they focus on the description of the main topics and objects of research, including the most topical elements.

In addition to the *Nový encyklopedický slovník* (2016), Kolek and Valdrová (2017) also provide a summary of FL and GL. They analyse the various approaches to the making of women visible in the language used in the most important journals of Czech studies, *Naše řeč* (*‘Our Speech’*) and *Slovo a slovesnost* (*‘Word and Word Art’*)<sup>4</sup> before the year 1995, and reach the surprising conclusion that the linguistic community was already more inclined to making of women more visible in the language before 1989, and the fall of socialism in the Czech Republic, than it is today.

---

3 E.g. in the [www.theses.cz](http://www.theses.cz) database.

4 The editorial board does not translate the name *Slovo a slovesnost*, cf. <http://sas.ujc.cas.cz/?lang=en>.

In 2018, J. Valdřová published a monograph *Reprezentace řenství z perspektivy lingvistiky genderovřch a sexuálnřch identit*, (*Representations of femininity from the perspective of linguistics of gender and sexual identities*). The monograph summarizes the state and trends of GL research in the Czech Republic, it supports the language construction of gender and gender hierarchy using examples from the public sphere and media, it reviews the possibilities of gender-fair language in Czech, and it informs readers of approaches used by the Czech general public as well as the linguistic community. Additionally, the author analyses the causes of the slow development of GL in the Czech Republic, which include: non-acceptance of the field by the linguistic community, its unwillingness to admit the interdisciplinary overlaps, and the ignorance of the results of the development and application of gender-fair language in foreign languages (Valdřová, 2018a, pp. 75–81). J. Valdřová notes that perhaps two research generations have neglected such issues, although adds that the situation has been improving, as those individuals who are still studying encounter issues regarding gender-fair language when coming into contact with foreign countries, and thus have to compare the foreign customs with Czech ones; moreover, the general public also perceives the issue of sexism within the language and discusses it. Additionally, another gender-linguistic dissertation is currently being prepared.

### 3 LABELLING OF PEOPLE REGARDING THEIR GENDER

#### 3.1 Approaches to the formation of nouns labelling women

The formation of the gender opposites of nouns (e.g. *učitel<sub>m</sub>* -> *učitelka<sub>f</sub>*, *vdova<sub>f</sub>* -> *vdovec<sub>m</sub>*, *teacher<sub>m</sub>* -> *teacher<sub>f</sub>*, *widow* -> *widower*) is a natural element of the Czech language system. The formation of nouns labelling women from the nouns labelling men is more common (i.e. female gender inflection, similarly to the surnames below). The Czech language features several suffixes to form a noun to label a woman – the most productive is the suffix *-ka* (see the examples above). The formation of feminine noun forms is easy (except for rare exceptions), and is among the main ways of creating gender-fair language.

The formation of nouns labelling the professions of women was actively supported by the editorial board of *Naše řeč*,<sup>5</sup> with a significant influence on

---

5 Available at <http://nase-rec.ujc.cas.cz/>

public opinion. The article *Slečna doktor* ('Miss Doctor<sub>m</sub>', 1918) criticizes the usage of masculine forms of nouns to label women as "nesmysl do nebe volající a ohavnost" ('an utter nonsense and atrocity'; p. 156). The article *Paní poslanec* ('Mrs. Deputy<sub>m</sub>', 1920) contains a short message that the labelling of women by masculine noun forms is "nesprávné a nečeské" ('incorrect and not Czech', p. 312). Moreover, in the article *Paní doktor* ('Mrs. Doctor<sub>m</sub>', 1922) the masculine noun form *doktor*, 'doctor' is seen as incorrect. With the note that "...v úředním seznamu osob působících na universitě Karlově čtete v oddílu nadepsaném »lektori« (ne »lektori a lektorky«)" ('...in the official list of people working at the Charles University we may read in the part which is labelled "lecturers<sub>m</sub>" (not "lecturers<sub>m</sub> and lecturers<sub>f</sub>"), p. 265), the editors address the issue of the lack of feminine noun forms if there are any women within the group. In the article *Z našich časopisů* ('From our magazines', 1922), the usage of the "generic" masculine is, however, called a *fashion* (nevertheless, FL is, paradoxically, called a *fashion* in subsequent articles). In the article *Paní ministryně* ('Mrs Minister<sub>m</sub>', 1924), the editors support the feminine noun form *ministryně*, 'minister<sub>f</sub>' as an analogue to *paní poslankyně*, 'Mrs Deputy<sub>f</sub>'. In the article *Magister, soudce, mistr* ('Magister, judge, foreman', 1929) the masculine noun forms which are used to label women are called "nečeským modernismem" ('non-Czech modernism', p. 166). The two articles *Slečna doktor* ('Miss Doctor<sub>m</sub>', 1930, 1931) give information about the recommendations made by the Academic Senate of Charles University and Czech Technical University in Prague to use feminine noun forms in both the professional and private spheres. The editors of *Naše řeč* welcomed that decision, adding "těší se, že nepřirozené a nečeské tituly (...) přece jednou ustoupí způsobu správnému." ('they look forward to the fact that the unnatural and non-Czech degrees will yield to the correct ones.', p. 199). The article *Člen, členka* ('Member<sub>m</sub>, member<sub>f</sub>', 1932) shows many examples of women labelled by a masculine noun form which was, again, seen as "not Czech" by the editors of the journal, while they were also surprised by the fact that "na věc tak samozřejmou je třeba stále ještě ukazovat, ačkoli nejen *Naše řeč*, ale i úřady a noviny už několikrát zavrhly toto nepřirozené zaměňování rodu." ('despite that this is self-evident, it is necessary to point this out, although not only *Naše řeč*, but also the offices and newspaper have rejected this unnatural mixing up of the gender', p. 124). F. Oberpfalcer (1932a, 1932b, 1932c, 1932d,

1933a) introduced the suffixes of nouns labelling women and presented many examples. Additionally, the article *Paní poslanec* ('Mrs. Deputy<sub>m</sub>', 1935) points out to the unnaturalness of the labelling of women with masculine noun forms, and it contains many then-topical feminine noun forms.

The text *Akademické tituly žen* ('Academic degrees of women', 1938) welcomes the decision of Charles University in the case of the formation of feminine forms of academic degrees, although the author of the text admits that female academics prefer the masculine form; that is, however, "*dokonalý komplex méněcennosti ženského pohlaví*" ('an absolute inferiority complex of the female gender', p. 221). J. Příkryl (1938) deals with "*módní zanedbávání ženského tvaru členka*" ('the fashionable negligence of the feminine noun form *členka*, member', p. 288). The then unusual sounding feminine *náměstkyně* 'deputy' is promoted by J. Příkryl (1945): it is formed "*v duchu jazyka*" ('in the spirit of the language', p. 215) and, according to the author of the article, there is no reason to avoid using the term, similar to the feminine noun form *ministryně* 'minister' in contrast to the then rival form *ministra* (Šmilauer, 1946). The frequency of articles dealing with nouns labelling women and the opinions they express in the first half of the 20<sup>th</sup> century show the considerable support for making women more visible in the language among both the professional and general public.

In the period of socialism in Czechoslovakia (1948–1989), several articles in *Naše řeč* dealt with nouns labelling women; however, they presented their arguments in the spirit of the political system of the day. K. Hausenblas (1950) states, for example, that "*...zrovnoprávnění žen v socialistické společnosti se obráží – jak to ani jinak nemůže být – i v našem jazyce*" ('... emancipation of women in the socialist society is reflected – and it cannot be otherwise – even in our language', p. 159). The article *Učednice* ('Apprentice', 1951) reports on the successful penetration of women into the factories, workshops and companies. "*Jazyk tu byl postaven před úkol poskytnout pro tato ženská povolání náležité názvy.*" ('The language was faced with a problem to provide appropriate labels for those female occupations', p. 77). M. Dokulil (1951) paraphrases Stalin's thoughts and states that "*teprve socialistická společnost zrovnoprávnila ženu*" ('it was only the socialist society that emancipated the woman completely', p. 127) – as evidenced by the formation of nouns labelling female occupations.

The tendency to protect the language against the foreign influences manifested in the Czechization of given names and surnames (see Chapter 4.2). With regard to common names, the article by V. Kondrová (1960) reflects the uncertainty towards foreign elements in the Czech language, using the example of the hybrid form *hosteska* (the English feminine noun form *hostess* and the Czech feminine suffix *-ka*). According to the author, the form *hostesa* or Czech equivalents *informátorka*, ‘*informér<sub>f</sub>*’ or *průvodkyně*, ‘*guide<sub>f</sub>*’ would better suit the rules of Czech word formation. However, Z Hrušková (1970) does not agree with the proposed alternatives, and states that “*méně obvyklý způsob tvoření jména hosteska nemůže být na překážku jeho užívání, je-li to slovo opravdu potřebné*” (‘even the less common way of formation of the word *hosteska* does not have to be an obstacle in its usage, if the word is really necessary’, p. 190).

J. Petr (1980) mentions the change of the political system which enabled the penetration of women into many occupations. The author allows the usage of masculine word forms if “*půjde o apoziční vyjádření obecného označení funkce nebo hodnoty, které bude stát oddělené čárkou za osobním jménem*” (‘it is an apposition expressing of a general label of a function or a rank which would be separated from the personal name by a comma’, p. 267). That is, apparently, the first mentioned possible use of (generic) masculine for labelling women in *Naše řeč*.

The formation of nouns labelling women outside *Naše řeč* is dealt with by, for example, J. Schwarz (1999) or T. Dickins (2001), and the latter, a British researcher of Czech studies, criticizes *Slovník spisovného jazyka českého* (*The Dictionary of the Official Czech Language*) in which many nouns labelling women (mainly occupations) are missing.

S. Čmejrková (2002) presents a theoretical analysis of the formation of nouns labelling women while using the structuralist theory of markedness (supposedly, the masculine noun gender represents both men and women, the feminine noun gender represents women only). However, she admits that “*užití generických maskulin může v některých kontextech vytvářet dojem, že to jsou hlavně (spíše, pouze) muži, o kom je řeč*” (‘the usage of generic masculine noun forms may, in some contexts, make an impression that it is (rather, only) the men who is talked about’; p. 279). Nevertheless, S. Čmejrková (2003) does

not step out of the shadow of structuralism or give up warning against gender-fair language, even in her similar article published in English. She strives again to trivialize the Czech gender-fair language by using heavy-handed, artificially created example sentences.

From the brief review presented above it may be assumed that, similar to the first half of the 20<sup>th</sup> century, the period of socialism can be characterized by the considerable support for using gender-inflected noun forms among of both the professional and general public. Naturally there were no feminist arguments present here, but instead appeals were made to the naturalness of the language, clarity of conveyance, and rightness, along with analogies to already existing and common words. During the period of socialism, however, fewer articles dealing with the formation of feminine forms of common nouns were published, and the authors used the lens of ideology as a means of argumentation. It was thus only the fall of the Iron Curtain which enabled the feminist reflection of the language, i.e. more than 20 years after this occurred in many other democracies.

### 3.2 “Generic” masculine noun forms and the gender-fair language

J. Valdřová (2017) defines the term “generic” masculine as “[n]ázev osoby v mužském rodě, který je míněn jako neutrální z hlediska rodu biologického, neboť mluvčí biologický rod nezná, nebo ho v daném kontextu považuje za méně důležitý než jiné sociální charakteristiky” (‘an appellation of a person in masculine noun gender which is seen as neutral from the point of biological gender since the speaker does not know the biological gender themselves or that they consider it less important than other social characteristics’). J. Valdřová (1996, 1997) was the first to point out the problematic nature and discriminative potential of the “generic” masculine in Czech linguistics: the overuse of the “generic” masculine evokes the image of men as the sole actors in public events. Czech language thus works in favour of men – the gender-specific inflection may (apart from the nouns) also be performed also on pronouns, adjectives, some numerals and some verbs forms: *Oba<sub>m</sub> naši<sub>m</sub> známí<sub>m</sub> umělci<sub>m</sub> zazpívali<sub>m</sub> hymnu* (‘Both<sub>m</sub> of our<sub>m</sub> known<sub>m</sub> artists<sub>m</sub> sang<sub>m</sub> the anthem’). Despite the fact that this sentence may also “include” a woman, she is not even mentioned, and therefore there is no need to look for her. In Valdřová (1997) as well as in many other articles after that (Valdřová, 2001, 2005a, 2010, 2013,

etc.), she formulated gender-fair alternatives to the “generic” masculine. Based on an analysis of 1,000 newspapers’ headlines (2001), she showed that the “generic” masculine may be replaced in as many as nine different ways (e.g. *výletní lodě*, ‘cruise ships’ instead of *lodě s výletníky*, ‘ships with shightseers’). In an analysis of communication at one conference (2005a), she illustrated the reproduction of the gender hierarchy – the men negotiate, women are the subjects of the negotiating.

The first psycholinguistic test of “generic” masculine perception was carried out by J. Valdová (2008). The research sample consisted of 572 respondents between 10–18 years of age. The task of the respondents was to think up funny surnames for ten people in a list according to their occupations (such as *a scientist*). The experimental group completed a test presented in the “generic” masculine form, while the control group completed a test with both feminine and masculine noun forms. The first version produced an overwhelming prevalence of masculine surnames, while in the second version the associations were significantly more balanced between masculine and feminine ones, and they reflected the reality better. This study was then criticized by J. Chromý (2008).<sup>6</sup>

In the same year, S. Čmejrková published the results of a questionnaire, in which the respondents were given a statement, “*Práce učitele<sub>m</sub> je hůř ohodnocená než práce policisty<sub>m</sub>*” (*The job of a teacher<sub>m</sub> is worse paid than the job of a police officer<sub>m</sub>*), and she asked them how they understood the sentence while giving the following options:

- 1) the author means both men and women,
- 2) the author means men only,
- 3) is it necessary to point out that both men and women are meant? (*učitele<sub>m</sub> a učitelky<sub>f</sub>*, ‘teacher<sub>m</sub> and teacher<sub>f</sub>’, *policisty<sub>m</sub> i policistky<sub>f</sub>*, ‘police officer<sub>m</sub> and police officer<sub>f</sub>’),
- 4) write down your own opinion.

More than a half of respondents selected the first option, and the second and third were each selected by more than one fifth of the respondents. However,

---

6 Chromý criticizes the author for the choice of research method, although he himself suggests a method that would only emphasize the contemporary usage.



this research method may not be seen as a truly associative experiment, but rather a purposeful interpretation of the given statement without the context.

The psychologist I. Smetáčková (2016) dealt with the “generic” masculine and the double forms from the perspective of their relation to the perception of prestige. By doing so, she discovered that the feminine noun form has a lower prestige in male-dominated fields and the masculine noun form has a lower prestige in female-dominated ones, and this may be caused by gender stereotypes. Other associative experiments can only be found in students’ theses. V. Kolek and D. Scheller-Boltz (in press) provide an overview together with a critical analysis of those works, even with Kolek’s own (comparative German-Czech) associative experiment, which verifies the conclusions made by J. Valdřová (2008). Tests that are interdisciplinary, methodologically varied, well captured and carried out in laboratories (working with a text and pictures, measuring of reaction times, etc.) are desired for the Czech GL. Such tests are conducted by researchers with various professional competences from the fields of sociology, psychology, linguistics, etc.

The meaning and aim of gender-fair language is to support the equal naming of women, men and other genders in the language. In Czech, gender-fair language concerns mainly a) the way of labelling persons (including overusing the “generic” masculine), and b) language sexism. The authors of the current work provide a list of available alternatives to the “generic” masculine below. The issue of language sexism is a matter of choice of language means and style, similar to that seen in English and other languages.

In 2010, the Ministry of Education, Youth and Sport of the Czech Republic published a guide to Czech gender-fair language on their website (Valdřová et al., 2010). In this manual, the authors initially deal with discrimination and language discrimination in general; subsequently, they discuss both the “generic” masculine and methods to achieve gender-fair Czech language, and gender-fair visualisations. The last part of the manual presents an overview of the methods of gender-fair language in English and German. In response to this, R. Adam organized a protest with a petition involving nearly all departments of Czech studies in the Czech Republic, and despite the fact that he did not propose any professional arguments the manual was withdrawn from the website and the professional discussion on the topic of gender-fair language



ended before it had really started. Media responses to this affair were analysed by Z. Mađarová (2015), and she gave evidence of fallacies involving *ad hominem*, *ad populum*, *causing of fear* and *downplaying of the stated issue*.

J. Valdřová (2013 in German, 2016 in English) reacts to the often used reproaches that the diversity of Czech inflection makes gender-fair language too difficult and the resulting texts non-transparent. She supports the functionality of Czech gender-fair language based on German-Czech analogies. The argument of the typological difference of the Czech language as an obstacle to using gender-fair language seems to be false, and only the willingness of speakers to express themselves in a gender-fair way is decisive.

The latest overview of Czech gender-fair language methods supplemented by suggestions of how to label non-binary persons is presented by V. Kolek (in press; cf. Valdřová, 2018a, pp. 401–407). These methods include:

- using feminine noun forms – when a woman is labelled: učitel<sub>m</sub> -> učitelka<sub>f</sub>, ‘teacher<sub>m</sub> -> teacher<sub>f</sub>’;
- using both feminine and masculine noun forms: učitelky<sub>f</sub> a učitelé<sub>m</sub>, učitelé<sub>m</sub> a učitelky<sub>f</sub>, ‘teachers<sub>f</sub> and teachers<sub>m</sub>, teachers<sub>m</sub> and teachers<sub>f</sub>’;
- using forms with slashes: učitel<sub>m</sub>/ka<sub>f</sub>, or parentheses: učitel<sub>m</sub>(ka<sub>f</sub>);
- verbal adjectives: vyučující, ‘teaching’;
- epicenes: osoba, tým, personál, ‘person, team, staff’;
- general names, names of the titles and institutions instead of “generic” masculines: rektorát, Německo, ‘rectorate, Germany’;
- nouns with attributes: učitelský sbor místo učitelé<sub>m</sub>, ‘teaching staff instead of teachers<sub>m</sub>’;
- a simple omission of the generic masculine: dopravní inspektoráty místo zaměstnanci<sub>m</sub> dopravních inspektorátů, ‘traffic inspectorates instead of employees<sub>m</sub> of the traffic inspectorates’;
- deictic labels: Vaše adresa místo adresa žadatele<sub>m</sub>, ‘your address instead of the address of the applicant<sub>m</sub>’;

- conversion to passive verb voice: *Formulář A musí být kompletně vyplněn místo Žadatel<sub>m</sub> musí kompletně vyplnit formulář A*, ‘form A has to be filled in completely instead of the applicant<sub>m</sub> has to fill in the form A completely’;
- when labelling (or including) the non-heteronormative persons, it is also possible to use a star – asterisk: *učitel<sub>m</sub>\*ka<sub>f</sub>*, or an underscore: *učitel<sub>m</sub>\_ka<sub>f</sub>*.

Studies and tests of the “generic” masculine and research into GL concepts have been slow in coming in the Czech Republic compared to other countries with stable democratic systems. This may be caused by the general aversion of both the general and professional public with regard to feminist topics across the scientific fields, and also by the unwillingness of middle-aged and older researcher to deal with the gender analyses of language. Many opponents of GL research hold important posts at universities, editorial boards or grant committees. Nevertheless (compared to the situation in the 1990s and despite the petition from 2010), the situation has changed, the issue of the relation between gender and language is gaining ground, and the gender-fair language is gradually respected by more individuals and institutions.

#### 4 THE NAME AS A GENDERED AND GENDERING PHENOMENON

The relation between a name, sex and gender is dealt by the scientific field *gender onomastics*. It is a very young scientific discipline, however “*kaum eine andere sprachliche Einheit ist so eng mit Geschlecht assoziiert wie der Personennamen*” (‘hardly any linguistic unit is that closely bound to the gender as the personal name’; Kotthoff and Nübling 2018, p. 191). Personal names are a popular subject of specialized Czech treatises, while the bibliographic basis of the area of personal names is very extensive.<sup>7</sup> However, the gender analysis of personal names is almost unknown in the field of Czech onomastics. The first text dealing with this issue in Czech was published by J. Valdřová (2019). Let the present chapter therefore be mainly the call for further research on this topic.

---

7 The lists of publications are available i.a. at <http://www.caslin.cz/caslin/homepage-caslin> (14. 6. 2020). The bibliography of entries *onomastika*, *proprium*, *přijmení* (‘onomastics, proper noun, surname’) *Nový encyklopedický slovník* include them as well; those are available at <https://bibliografie.ujc.cas.cz/search?type=global&q=propria> (14. 6. 2020), etc.

#### 4.1 Given name

Czech newborns get one or two given names.<sup>8</sup> In history, the formation of names was influenced by non-linguistic factors; apart from the parents' ideas and wishes, it also included the then value ladder of the specific community which included the feminine and masculine models. Female names often accented tenderness and beauty, while male names verbalized the image of bravery, strength, wisdom, etc. For example, the Czech names *Blanka* ("snow-white") and *Svatopluk* ("strong in the army"). In the course of time, however, semantics lost its primary function and a convention of approaches started to be applied. Female counterparts were created from the male names (*Daniel-a*, *Petr-a*, etc.). The system differences between the male and female names may also be seen (besides the meaning differences) in the phonetic form, length and accent, and these proprieties constitute the gender index of names (Nübling et al., 2015, pp. 131–137).<sup>9</sup>

In the Czech Republic, the given names have mostly been researched by M. Knappová (1978–2017,<sup>10</sup> 1989, etc.). The lists of given names are presented in encyclopaedias (*inter alia* Kopečný, 1974; Rameš, 2000). During the registration of given names and surnames, Czech register offices observe the current law on such offices and also the instructions and information given by the Ministry of the Interior, which follow from the statements of the Ústav pro jazyk český, 'Czech Language Institute' (hereinafter referred to as ÚJČ) and its long-term associate – the previously mentioned M. Knappová. Her influence on the practice of given names' and surnames' registration was called *completely essential* by R. Šrámek (2006, p. 104), and her manuals shape the work of register offices in a significant way.<sup>11</sup>

---

8 The often used term *křestní jméno*, 'Christian name' should be a sign that the baptism was performed.

9 There are, however, no studies on this issue in the Czech language.

10 The manual was published in years 1978–2017 with a slightly changed title and it was gradually updated in six editions.

11 In the practice, it is often possible to encounter the resistance of the register offices to less common given names: "*Není to v Knappové, takže to nejde.*" ('It is not in Knappová, therefore it is not possible.'). Knappová (2009, p. 84) states that the (given) names of film characters are chosen for the children of those with lower education, not-very-much occupied with work, constantly watching TV, etc. I. Lutterer (1990, p. 254) classes them explicitly "*k oné svérázné etnické skupině, která je u nás nechvalně*

In the field of gender onomastics research into given names the following issues may seem to be of most interest:

- the diachronous and synchronous research into persons' labelling from the perspective of gender criticism of the language; the name as a socionym;
- the semantic principles of the construction of gender in the field of names;
- the gender index of Czech personal names and its perception by society;
- the addressing and labelling of women and men, girls and boys with personal names, the pragmatics of diminutives;
- the gender factors underlying the motivation to choose specific given names, trends;
- the quality, accessibility and function of scientific literature and information sources for the public;
- gender neutral given names, their development, trends, changes, index of options.

The last item on the list seems particularly urgent in the present day, when register offices give only limited information on gender neutral given names. The only source (Knappová, 2017, pp. 88–92) deals with them in a short chapter with an unfortunate title *Osobní jména transsexuálů*, 'Personal names of transsexuals'; the common name *transsexuál*, 'transsexual' in the title needlessly accentuates the issue of sexuality, unlike the labelling *trans lidé*, 'trans people'. Supposedly, the chapter "*není určena rodičům vybírajícím jméno pro očekávané dítě*" ('is not intended to parents choosing the name for their expected baby'; Knappová, 2017, p. 88), however, the author does not clarify why a neutral name could not be borne by any child if it is a wish of their parents. The content of the chapter aims at the tokenization of gender difference: e.g. outside the law No. 301/2000 Coll. in force, which reserves the

---

*známá zvýšeným sklonem k zločinnosti.*" ('to the peculiar ethnic group which is here ingloriously known by their higher tendency to commit crime').

domestic forms of given names (*Míťa, Zbyňša*, etc.) for trans people, and the choices among the surnames are restricted as well. It is no wonder that trans people make their requests for equality in terms of names, as referred to by J. Valdrová (2018a, 2018b) and V. Kolek (2019).

#### 4.2 Surnames

In the Czech language area, the majority of people have one family or married name. Moreover, the origin of surnames was influenced by non-linguistic factors, and in the context of the current study these include gender hierarchies. The strong motivation to mark the proprietary dependence of women on their fathers, husbands or legal guardians led to the difference of female Czech surnames by a formant (suffix) *-ová* (masculine possessive suffix *-ov-* and an adjective ending *-á*, or once *-a*): *Jahoda* (masculine surname) – *Jahodova, Jahodová* (feminine surnames).<sup>12</sup> M. Knappová (1992, p. 78) states that the difference between *a* and *á* differentiates an unmarried woman from married ones; however, as previously mentioned by F. Oberpfalcer (1933b) and F. Cuřín (1936), among others, there is no evidence for this. In Old Czech, the suffix *-ka* was used to label women: J. Kouba (1983) researched the official data of women's surnames with this suffix (mainly from the 15th century), and postulated a thesis that those women were mainly unmarried ladies, including widows.<sup>13</sup>

Up to the Second World War, the formant *-ová* was used non-systematically for both Czech and foreign women's surnames – certainly due to the influence of a multilingual Czech-German-Jewish environment, the mutual respect for names' forms in various languages, and sometimes even due to a speculation about the advantages of claiming allegiance to this or that ethnicity (cf. Valdrová, 2019, p. 454). The xenophobic speeches of then Czechoslovak president E. Beneš, who demanded complete de-Germanization, including of names, sped up the expulsion of minority German Bohemians in 1945. A year later, the national committees (contemporary municipal authorities) performed an

---

12 Among adjective surnames, the gender is marked among both men and women by an ending: male surname – *Novotný*, female surname – *Novotná*.

13 In the contemporary Czech interdialect, the surnames ending with *-ka* are, in an unofficial usage, marked: e.g. *Bohdalka* (derived from *Bohdal*) may, according to the context, express the popularity of a particular person as well as the critical distance.

unprecedented Czechization of German personal names: e.g. *Hans Schmied* became *Jan Šmíd*, etc.<sup>14</sup>

The feminization of women's surnames was performed in even a more thorough way by using the *-ová* suffix (see Valdrová, 2019, pp. 454–455). Since it was not possible for the country to see the feminization of surnames on the radio, TV and in newspapers, J. Kuchař (1956) made the call “*Přechylujme ženská příjmení cizího původu!*” (‘Let’s use the gender inflection on the women’s surnames of foreign origin!’) and demanded changing, for example, Hungarian women’s surnames according to Czech grammar: *Matayová* instead of the original *Matayné*. The suffix *-né*, which marks the marital status, “*stejně pro nás nemá významovou hodnotu*” (‘does not have the meaningful value for us anyway’; *ibid.*, p. 304). The instructions for the Czechization of the surnames are examined in a book by M. Knappová (1992); according to her, the desirable forms are the following ones *Indira Gándhí* -> *Gándíová* etc.

In summary, this means that blanket feminization at this time was performed for political motives. During the socialist period, it was promoted by M. Knappová (1979, p. 225), a member of ÚJČ regardless of the medium used and the nationality and origin of the women concerned. The surnames of female foreigners who lived and worked in Czechoslovakia were given a formant on official documents without their consent.<sup>15</sup>

The formant *-ová* would not be so widespread if it was not justified by quasi-linguistic arguments. In the internet advisory centre of ÚJČ that is aimed at the public, it is possible to read even now that if the speaker does not use the formant it may lead to misunderstandings. In sentences e.g. *Susan Sonntag navštívila Shirley Temple* (‘*Susan Sonntag visited Shirley Temple*’), it is supposedly not possible to find out the gender of the person without the formant, and mistaking the subject for the object may occur since Czech word order is allegedly not as fixed as in other languages. Therefore, the “correct”

---

14 Matúšová (2003). Some people, however, resisted the Czechization of their names; they pointed out that even some members of the government (such as Gottwald) had a German surname.

15 The Ministry of Interior of the Czech Republic is still basically following the recommendations made by ÚJČ and Knappová. JUDr. Kateřina Guluškinová, letter MV ČR č. MV-25510-2/VS-2009, 14. 4. 2009).

form of the sentence should be *Susan Sontagová navštívila Shirley Temple-ovou*.<sup>16</sup> Some linguists (R. Adam, M. Pravdová, etc.) cling to the necessity to express the feminine gender in the form of one's surname, and even confess that they cannot deal with non-feminized surnames themselves.<sup>17</sup> In this they differ remarkably from actual practice, as lay speakers do not have any issues with the usage of original forms in this context (a number of examples are given by Valdová, 2015b).

At first, the foreign women who were affected by the change of their official documents protested against the compulsory feminization of surnames. They argued for the right to inviolability of the surname, the problems they encountered in their daily lives anytime the form of their surnames in various official documents differed, even by just the semantics of the formant as a sign of the allegiance of a woman to a man. The peak of their protests was a complaint at the Czech Helsinki Committee in 2000; in 2001, an amendment brought some unclearly formulated relief to foreign women, allowing the surname to stay in its original form if it assigned by an internationally recognized convention; however, the type of convention was not specified (Valdová, 2002). In the meantime, even Czech women started to apply for a non-feminized surname.

Both the general and professional public still keep silent about the issue of the patrilineality of surnames, the fact that women abandon their maiden names and adopt their husbands' surnames. In contrast, only very rarely do husbands adopt their wives' surnames, and this asymmetry is analysed by D. Komanická (2016). As yet, however, there is no research dealing with same-sex couples in this matter. Over the last decade, the number of women accepting their husband's name in the original form (e.g. Eva *Havel*, i.e. without the formant) has risen sharply.

---

16 *Přechylování příjmení ve veřejné komunikaci*. Available at <https://prirucka.ujc.cas.cz/?id=700&dotaz=přechylovánípříjmení> (14. 6. 2020). As the only concrete source, Knappová has been recommended here for more than 15 years, although others, such as Moldanová (2015) and Matúšová (2003), have also dealt with surnames.

17 M. Pravdová in a TV talk, available at [https://www.idnes.cz/zpravy/domaci/ova-prijmeni-marketa-pravdova-ustav-pro-jazyk-cesky-akademie-ved-rozstrel.A190909\\_152128\\_domaci\\_rko](https://www.idnes.cz/zpravy/domaci/ova-prijmeni-marketa-pravdova-ustav-pro-jazyk-cesky-akademie-ved-rozstrel.A190909_152128_domaci_rko) (K. Bulisová; 14. 6. 2020). R. Adam admitted problems with the usage of the original form of women's surnames in a radio talk, available at <https://vltava.rozhlas.cz/prezije-cestina-bez-prechylovani-8084719> (T. Samek and L. Matoška; 14. 6. 2020).

The existing distinguishing features between so-called masculine and feminine surnames is also increasingly unsustainable as it produces and legitimizes heteronormativity. Therefore, when it comes to the choice of surnames for trans people there is not much room to manoeuvre, as the manual prepared by M. Knappová (2017) for register offices does not meet today's naming needs.

The topical themes for gender onomastics research into surnames are as follows:

- the formation, development and usage of surnames as elements of one's identity both in the past and in the present; the disrespect of surnames and denying of surnames as an instrument of symbolic violence;<sup>18</sup>
- the feminization of surnames as a means of the grammaticalization of gender, as well as sexual and gender hierarchies, history, development and trends;
- the formation and official registration of surnames from the perspective of patronymy and patrilinearity;
- both critical diachronic and synchronic revision of the scientific authorities' arguments for and against the feminization of surnames, and the role of ÚJČ;
- the contemporary practice of the process of surname registration, its development and trends;
- a critical reflection of the possibilities and limits of the choice of surnames from the perspective of trans people and their needs.

As for now, the research into contemporary anthronomastics (Kopecký, 2014; Tušková and Žižková, 2016, etc.) ignores the gender aspects of the issue even, in such cases in which the choice of the form of a surname is the subject of the

---

18 In the time of the refugee crisis, there were cases when people were labelled with numbers rather than names, cf. Netrvalová (2. 9. 2015), and as in the 1930s, the world criticized Czechia for this practice. Available at [https://www.idnes.cz/zpravy/zahranicni/cisla-na-rukou-uprechliku-pobourila-aktivisty.A150902\\_140434\\_zahranicni\\_san](https://www.idnes.cz/zpravy/zahranicni/cisla-na-rukou-uprechliku-pobourila-aktivisty.A150902_140434_zahranicni_san) (14. 6. 2020).



research.<sup>19</sup> The lens of gender onomastics is applied by publications dealing with surnames from the perspective of human rights (Valdrová, 2002), the usage and attitudes of expert authorities (Valdrová, 2015b), the taking over of surnames in the patronymic and patrilinear line (Komanická, 2016), names as elements of the identity (Valdrová, 2018a, pp. 343–368), and trans surnames (Valdrová, 2018b, 2019; Kolek, 2019).

From the perspective of gender onomastics, it is possible to see the surprisingly close bond between personal names and the political situation of the day. The pre-war multinational society was characterized by a variety of given names and surnames. From 1946, both names and surnames were Czechized. The surnames were thoroughly feminized and the related political measures were (and are, even now) quasi-justified by the language system, word order and the demand to express the gender of the person. With the gradual emancipation of women (including language emancipation), however, forms without the formant *-ová* are increasingly common.

## 5 GENDER ASPECTS IN TRANSLATOLOGY

The lack of publications thematizing the aspects of gender is, in our opinion, caused by the generally low sensibility of the Czech public towards of its importance in oral presentation, despite the fact that everyone who is able to use a foreign language had to encounter at least some recommendations for the use of non-sexist language.

The first Czech translatology text was, apparently, published by E. Věšíňová (1998). In this the author deals with general questions such as the lack of discussion about gender linguistics, including the male and female genderlect, or the overuse of the “generic” masculine in the Czech language.

An empirical comparison of the two Czech translations of the book *Lady Chatterley’s Lover* from two different periods was carried out by S. Širokovská

---

19 It is possible to object to research in terms of its methodology, e.g. the issue of researchers “...zda respondentky vědí o možnosti používat v neoficiální komunikaci svá příjmení bez *-ová*” (‘whether the female respondents know that there is a possibility to use their surnames without the formant *-ová* in the informal communication’). This gives an impression that the informal usage is set by non-specified authorities (Tušková and Žižková, 2016, p. 129).

(2004), with the first translation done by a female translator (S. Jílovská, 1932) and the second by a male one (F. Vrba, 1987). In her article, the author aimed to present the translation differences caused by the gender of the translators. V. Janiš (2004) reacted critically to this while, noting the inconsistencies and paradoxes raised by S. Širokovská, and overall the connection of such an analysis with the translators' gender. V. Janiš also suggests a better way to carry out such research, and notes the insignificant extent size of the sample studied by S. Širokovská (just two translations).

E. Věšínová elaborated her idea in an article published under the name Kalivodová in 2012. In this she deals with the concept of so-called feminist translation, outlines in detail the relations between the translated and translation on the one hand, and the gender category on the other hand. She uses good and relatively bad (students') translations which are subjected to analyses, and assesses them from the gender perspective based on various theoretical approaches. She also points out the changes in the impressions the texts give if translated insensitively with regard to the gender. The author also describes her own experience of the related educational activity at the translatology institute in another publication (Kalivodová, 2017).

J. Valdřová (2005b) criticizes the lack of gender-fair language and distorting of foreign female surnames with a Czech formant (cf. above) when translating and interpreting. The lack of gender-fair language may misrepresent the reality, e.g. *hrdinný policista*, 'heroic<sub>m</sub> police officer<sub>m</sub>', in a Czech news story regarding a terrorist attack in Barcelona, when the police officer was in fact a policewoman (Valdřová, 2019, p. 163).

Apart from English, which the authors of the present paper do not consider quite appropriate to compare with the Czech language in the area of gender (due to the typological differences), the authors of the present article suggest comparing the gender dimension in Czech to that in other languages, e.g. German. In that case, it would be possible to follow the various levels of gender-fair language being enforced in German-speaking countries and consider their use in Czech, while also possible to deal with the texts from the perspective of their acceptability and adequacy.

## **6 CONCLUSION**

Gender linguistic research has seen considerable development in Czech studies. The pre-feminist articles on the topic of the formation of female profession names contributed significantly to the support for the labelling of women by a feminine marker. After 1989, Czech linguistics opened up to information from the field of foreign feminist linguistics, although some linguists protested against this. In the 1990s, J. Valdřová started discussions on gender linguistic topics among both the professional community and general public, such as on the unreliability of the generic function of a masculine form when labelling women, language gender stereotypes in the spoken language and their influence on the formation and understanding of reality and attitudes to it. After the analyses of gender-fair alternatives used in the spoken language in German, English and Russian, J. Valdřová formulated language recommendations for Czech. Today's social situation and the demands of the public have been dealt with by J. Valdřová carrying out research into the labelling needs, possibilities and limits in the area of first names and surnames, including trans names. V. Kolek follows on from her work and broadens it with additional topics, e.g. the labelling of non-binary genders within the texts.

J. Valdřová and V. Kolek deal with the language aspects of gender sensitive education. It is possible to recommend various works on the gender fairness of textbooks (Valdřová et al., 2005), methods of gender sensitive educational work (Smetáčková and Vlková, 2005; Smetáčková, 2007; Babanová and Miškolci, 2007; Babanová, 2019) and other issues. The gender critical analyses concern the communications among teachers and pupils and students in a gendered school environment, the role of gender stereotypes in the educational work, assessment and study results of pupils and students, the gender burden of text books, work with class dynamics, and so on.

Every social issue has its language side. The language either limits its speakers and misrepresents reality (when it functions as a medium for the transfer of stereotypes), or – to paraphrase J. Butler (2004) – when it sharpens the perception of the gendered world and makes us free, since it teaches us to better deal with the gendered reality and shows (mainly young people) the way to free life choices.

### Acknowledgements

This paper and its underlying research were made possible with financial support of the Ministry of Education, Youth and Sports of the Czech Republic (MŠMT ČR) granted to Palacký University Olomouc (IGA\_FF\_2019\_005).

### REFERENCES

- Akademické tituly žen (1938). *Naše řeč*, 22(6–7), 220–221. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=3314>
- Babanová, A. (2019). *Dvanáct ověřených metod pro rozvoj kolegiální spolupráce*. Praha: Gender studies. Retrieved from [https://aa.ecn.cz/img\\_upload/8b47a03bf445e4c3031ce326c68558ae/dvanact\\_overenych\\_metod\\_a4\\_web.pdf](https://aa.ecn.cz/img_upload/8b47a03bf445e4c3031ce326c68558ae/dvanact_overenych_metod_a4_web.pdf)
- Babanová, A., & Miškolci, J. (2007). *Genderově citlivá výchova: Kde začít? Příručka pro vyučující základních a středních škol*. Praha: Žába na prameni.
- Butler, J. (2004). *Undoing Gender*. London: Routledge
- Chromý, J. (2008). K článku Jany Valdové „Žena a vědec? To mi nejde dohromady.“. *Naše řeč*, 91(4), 197–200. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=8007>
- Člen, členka (1932). *Naše řeč*, 16(4), 124–125. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2600>
- Čmejková, S. (1995). Žena v jazyce. *Slovo a slovesnost*, 56(1), 43–55. Retrieved from <http://sas.ujc.cas.cz/archiv.php?art=3614>
- Čmejková, S. (1997). Jazyk pro druhé pohlaví. In F. Daneš (Ed.), *Český jazyk na přelomu tisíciletí* (pp. 145–160). Praha: Academia.
- Čmejková, S. (2002). Rod v jazyce a komunikaci: specifika češtiny. *Slovo a slovesnost*, 63(4), 263–286. Retrieved from <http://sas.ujc.cas.cz/archiv.php?art=4086>
- Čmejková, S. (2003). Communication Gender in Czech: Repertoire of Nouns and their Distribution and Reference in Text. In H. Bussmann & M. Helinger (Eds.), *Gender across Languages* (pp. 27–57). Amsterdam/Philadelphia: Benjamins.
- Čmejková, S. (2008). Jak rozumíme kategorii rodu? Kdy je Češka Čech? *Slavia – časopis pro slovanskou filologii*, 77(1–3), 41–54.

- Cuřín, F. (1936). O přechylovací příponě -ova, -ová. *Naše řeč*, 20(4–5), 106–111. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=3028>
- Daneš, F. (1997). Ještě jednou „feministická lingvistika“. *Naše řeč*, 80(5), 256–259. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7412>
- Dickens, T. (2001). Gender Differentiation and the Asymmetrical Use of Animate Nouns in Contemporary Czech. *The Slavonic and East European Review*, 79(2), 212–247.
- Dokulil, M. (1951). Nová skutečnost v zrcadle slovní zásoby češtiny. *Naše řeč*, 35(7–8), 121–131. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=4246>
- Hausenblas, K. (1950). Pověřenkyňe a stachanovka. *Naše řeč*, 34(7–8), 159. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=4199>
- Hoffmannová, J. (1995). Feministická lingvistika? *Naše řeč*, 78(2), 80–91. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7234>
- Hrušková, Z. (1970). Hosteska. *Naše řeč*, 53(3), 188–190. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=5506>
- Janiš, V. (2004). Jak nesrovnávat překlady. *Naše řeč*, 87(4), 205–207. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7792>
- Kalivodová, E. (2012). Genderová slepota, citlivost nebo provokace? Polohy a možnosti překladu. *Gender, rovné příležitosti, výzkum*, 13(2), 77–85. Retrieved from [https://www.genderonline.cz/uploads/b48098405a472611d-805517288c26cc30bde2061\\_kalivodova-genderova-slepota.pdf](https://www.genderonline.cz/uploads/b48098405a472611d-805517288c26cc30bde2061_kalivodova-genderova-slepota.pdf)
- Kalivodová, E. (2017). The Challenges of Teaching Gender in Translation in a Contemporary Czech Context. In J. Santaemilia (Ed.), *Traducir para la igualdad sexual = Translating for sexual equality* (pp. 149–156). Granada: Comares.
- Karlík, P., Nekula, M., & Pleskalová, J. (Eds.). (2016). *Nový encyklopedický slovník češtiny*. Praha: NLN. Retrieved from <https://www.czechency.org/>
- Knappová, M. (1978–2017). *Jak se bude jmenovat?* Praha: Academia.
- Knappová, M. (1979). Přechylování příjmení v češtině. *Naše řeč*, 62(5), 225–233. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=6153>
- Knappová, M. (1989). *Rodné jméno v jazyce a společnosti*. Praha: Academia.
- Knappová, M. (1992). *Příjmení v současné češtině*. Liberec: AZ KORT.
- Kolek, V., & Valdřová, J. (2017). Die tschechische sprachwissenschaftliche Geschlechtsforschung im Spiegel der bohemistischen Fachzeitschriften *Naše*

- řeč und *Slovo a slovesnost*. In M. Reisigl and C. Spieß (Eds.), *Osnabrücker Beiträge zur Sprachtheorie 91. Sprache und Geschlecht. Band 2: Empirische Analysen* (pp. 147–165). Duisburg: Universitätsverlag Rhein-Ruhr.
- Kolek, V. (2019). Research of the Czech language from the perspective of Queer Linguistics. *Slovensčina 2.0: empirical, applied and interdisciplinary research*, 7(1), 113–125. Retrieved from <https://revije.ff.uni-lj.si/slovenscina2/article/view/8655/8735>
- Kolek, V. (in press). Options for Labelling Non-heteronormative People: a German-Czech Comparison. In D. Scheller-Boltz (Ed.), *Language Policies in the Light of Anti-Discrimination and Political Correctness: Tendencies and Changes in the Slavonic Languages (Wiener Slawistischer Almanach, Sonderband 84)*, pp. 303–314). Frankfurt a.M.: Peter Lang.
- Kolek, V., & Scheller-Boltz, D. (in press). The Use and Perception of Masculine Generics in Czech, German, and Polish: a Cognitive Study. In D. Scheller-Boltz (Ed.), *Language Policies in the Light of Anti-Discrimination and Political Correctness: Tendencies and Changes in the Slavonic Languages (Wiener Slawistischer Almanach, Sonderband 84)*, pp. 235–249). Frankfurt a.M.: Peter Lang.
- Komanická, D. (2016). Patronymia a patrilinearita vo vzťahu k nerovnému postaveniu žien v spoločnosti. In L. Molnár Satinská & I. Valentová (Eds.), *Prechýľovanie: Áno – nie?* (pp. 55–67). Bratislava: Veda.
- Kondrová, V. (1960). Hosteska. *Naše řeč*, 43(9–10), 306–307. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=4799>
- Kopecký, J. (2014). Přechylování příjmení v češtině jako případ jazykového managementu. *Slovo a slovesnost*, 75(4), 271–293. Retrieved from <http://sas.ujc.cas.cz/archiv.php?art=4258>
- Kopečný, F. (1974). *Průvodce našimi jmény*. Praha: Academia.
- Kotthoff, H., & Nübling, D. (2018). *Genderlinguistik. Eine Einführung in Sprache, Gespräch und Geschlecht*. Tübingen: Narr.
- Kouba, J. (1983). Pojmenování žen s formantem -ka ve staré češtině. *Naše řeč*, 66(5), 246–249. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=6418>
- Kuchař, J. (1956). Přechylujme ženská příjmení cizího původu!. *Naše řeč*, 39(9–10), 302–304. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=4519>

- Lutterer, I. (1990). Monografie o postavení rodného jména v jazyce a společnosti. *Naše řeč*, 73(5), 252–254. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=6968>
- Maďarová, Z. (2015). Analýza mediálnej odozvy na príručku o rodovo vyváženom vyjadrovaní v Českej republike. In J. Cviková (Ed.), *Analýza významu a možností používania rodovo vyváženého jazyka* (pp. 111–132). Bratislava: Centrum vzdelávania Ministerstva práce, sociálnych vecí a rodiny SR. Retrieved from [http://www.ruzovyamodrysvet.sk/chillout5\\_items/1/6/0/1/1601\\_5755fo.pdf](http://www.ruzovyamodrysvet.sk/chillout5_items/1/6/0/1/1601_5755fo.pdf)
- Magister, soudce, mistr (1929). *Naše řeč*, 13(7), 166. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=1587>
- Matúšová, J. (2003). K vývoji příjmení u Čechů v prvních poválečných letech. *Naše řeč*, 86(5), 251–256. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7747>
- Moldanová, D. (2015). *Naše příjmení*. Praha: Pankrác.
- Nübling, D., Fahlbusch, F., & Heuser, R. (2015). *Namen: Eine Einführung in die Onomastik*. Tübingen: Narr.
- Oberpfalcer, F. (1932a). Přechylování jmen jako výraz rozdílu v pohlaví, I. *Naše řeč*, 16(7), 193–199. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2627>
- Oberpfalcer, F. (1932b). Přechylování jmen jako výraz rozdílu v pohlaví, II. *Naše řeč*, 16(8), 225–232. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2633>
- Oberpfalcer, F. (1932c). Přechylování jmen příponou -ice, I. *Naše řeč*, 16(9), 257–263. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2642>
- Oberpfalcer, F. (1932d). Přechylování jmen příponou -ice, II. *Naše řeč*, 16(10), 295–300. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2653>
- Oberpfalcer, F. (1933a). Přípona -yni, -yně při přechylování. *Naše řeč*, 17(1), 5–11. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2656>
- Oberpfalcer, F. (1933b). O příponě -ová. – Přechylování jmen značících mužská zaměstnání. – Přechylování názvů příbuzenských. – Přechylování mužských příjmení. – Přípona -ovka. – Přípona -ovna. *Naše řeč*, 17(3), 72–76. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2678>
- Paní doktor. (1922). *Naše řeč*, 6(9), 265–266. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=1874>
- Paní ministryně. (1924). *Naše řeč*, 8(6), 167. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=1135>



- Paní poslanec. (1920). *Naše řeč*, 4(10), 312–313. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=728>
- Paní poslanec. (1935). *Naše řeč*, 19(3), 94–96. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2962>
- Petr, J. (1980). Rada, nebo radová? *Naše řeč*, 63(5), 265–267. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=6223>
- Příkryl, J. (1938). Členka a j. *Naše řeč*, 22(9), 288. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=3339>
- Příkryl, J. (1945). Náměstkyně. *Naše řeč*, 29(9–10), 214–215. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=3927>
- Rameš, V. (2000). *Po kom se jmenujeme? Encyklopedie křesťanských jmen*. Praha: Libri.
- Schwarz, J. (1999). *Die Kategorie der Weiblichkeit im Tschechischen: die Verwendung femininer und maskuliner Personenbenennungen für Frauen im Vergleich zum Deutschen*. Praha: Filozofická fakulta.
- Slečna doktor. (1918). *Naše řeč*, 2(5), 156–157. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=271>
- Slečna doktor. (1930). *Naše řeč*, 14(5), 108. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=1718>
- Slečna doktor. (1931). *Naše řeč*, 15(5), 119. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=2333>
- Smetáčková, I., & Vlková, K. (Eds.). (2005). *Gender ve škole - příručka pro vyučující předmětů občanská výchova, občanská nauka a základy společenských věd na základních a středních školách*. Praha: Otevřená společnost. Retrieved from <https://www.otevrenaspolecnost.cz/download-publication/821?publication=2955>
- Smetáčková, I. (Ed.). (2007). *Příručka pro genderově citlivé vedení škol*. Praha: Otevřená společnost. Retrieved from <https://www.otevrenaspolecnost.cz/knihovna/otevrenka/prosazovani-genderove-rovnosti/genderove-citlive-vedeni-skol.pdf>
- Smetáčková, I. (2016). Femininní a maskulinní označení: vliv na hodnocení prestiže. *Gender, rovné příležitosti, výzkum*, 17(2), 81–92. Retrieved from [https://www.genderonline.cz/uploads/7995bdecbe1e13bfd1b917dbobb38cdef91017375\\_grpv-216-stat-smetackova-femininni-a-maskulinni.pdf](https://www.genderonline.cz/uploads/7995bdecbe1e13bfd1b917dbobb38cdef91017375_grpv-216-stat-smetackova-femininni-a-maskulinni.pdf)



- Širokovská, S. (2004). Dvojitý překlad knihy D. H. Lawrence Milenec Lady Chatterleyové z perspektivy genderu. *Naše řeč*, 87(1), 14–24. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7758>
- Šmilauer, V. (1946). Ministra. *Naše řeč*, 30(9–10), 179. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=3989>
- Šrámek, R. (2006). K jubileu Miloslavy Knappové. *Naše řeč*, 89(2), 103–104. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7888>
- Tušková, J. M., & Žižková, H. (2016). Postoje českých žen k přechylování příjmení. In L. Molnár Satinská & I. Valentová (Eds.), *Přechyľovanie: Áno – nie?* (pp. 119–133). Bratislava: Veda.
- Učednice (1951). *Naše řeč*, 35(3–4), 77–78. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=4233>
- Valdrová, J. (1996). Sexismus v německém jazyce a situace u nás. In D. Moldánová (Ed.), *Žena – jazyk – literatura. Sborník z mezinárodní konference* (pp. 288–291). Ústí nad Labem: Univerzita J.E. Purkyně.
- Valdrová, J. (1997). K české genderové lingvistice. *Naše řeč*, 80(2), 87–91. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7365>
- Valdrová, J. (1998). *Kontrastivní genderová lingvistika: téma zviditelnění ženy v současném německé a českém jazyce*. Disertační práce. Brno: FF MU.
- Valdrová, J. (2001). Novinové titulky z hlediska genderu. *Naše řeč*, 84(2), 90–96. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7614>
- Valdrová, J. (2002). Linguistische und soziopolitische Aspekte der Movieierung fremder weiblicher Familiennamen im Tschechischen. In J. van Leeuwen-Turnovcová et al. (Eds.), *Wiener Slawistischer Almanach* (Sonderband 55) (pp. 247–258). Wien: Gesellschaft zur Förderung slawistischer Studien.
- Valdrová, J., Smetáčková, I., & Knotková-Čapková, B. (2005). Příručka pro posuzování genderové korektnosti učebnic. In B. Knotková-Čapková (Ed.), *Ročenka katedry genderových studií FHS UK 03/04* (pp. 173–196). Praha: Fakulta humanitních studií UK.
- Valdrová, J. (2005a). Jak jazyk zabíjí image odbornice. *Gender, rovné příležitosti, výzkum*, 6(2), 1–3. Retrieved from [https://www.gendersonline.cz/upload-s/4bd25679c2fcaac2ca403a0395672d458dc8d941\\_jak-jazyk-zabiji-image-odbornice.pdf](https://www.gendersonline.cz/upload-s/4bd25679c2fcaac2ca403a0395672d458dc8d941_jak-jazyk-zabiji-image-odbornice.pdf)

- Valdrová, J. (2005b). Feministické vědecké aspekty překladu názvů osob. In M. Hrala (Ed.), *Český překlad II (1945–2004)* (pp. 367–379). Praha: FF UK.
- Valdrová, J. (2008). „Žena a vědec? To mi nejde dohromady.“ Testy generickeho maskulina v českém jazyce. *Naše řeč*, 91(1), 26–38. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7976>
- Valdrová, J., Knotková-Čapková, B., & Paclíková, P. (2010). *Kultura genderově vyváženého vyjadřování*. Praha: MŠMT.
- Valdrová, J. (2010). Genderová kultura projevu v českém jazykovém prostředí. In J. Valdrová, B. Knotková-Čapková & P. Paclíková (Eds.), *Kultura genderově vyváženého vyjadřování* (pp. 18–46). Praha: MŠMT.
- Valdrová, J. (2013). Typologische Unterschiede zwischen Sprachen als Argument gegen geschlechtergerechte Sprachkritik und Sprachpflege? In M. Nekula, K. Šichová & J. Valdrová (Eds.), *Bilingualer Sprachgebrauch und Typologie: Deutsch – Tschechisch* (pp. 141–158). Tübingen: Julius Groos Verlag.
- Valdrová, J. (2015a). Gibt es eine feministische Linguistik in der tschechischen Sprachwissenschaft? In D. Scheller-Boltz (Ed.), *New Approaches to Gender and Queer Research in Slavonic Studies* (pp. 293–305). Wiesbaden: Harrassowitz.
- Valdrová, J. (2015b). Angela Merkelová, Elfriede Jelinek(ová) und Barbara Coudenhoveová-Kalergiová: Zum gegenwärtigen tschechischen Umgang mit weiblichen Familiennamen In P. Anreiter, E. Mairhofer & C. Posch (Eds.), *Argumenta. Festschrift für Manfred Kienpointner zum 60. Geburtstag* (pp. 565–575). Wien: Praesens Verlag.
- Valdrová, J. (2016). Typological Differences between Languages as an Argument against Gender-fair Language Use?. In I. Jusová & J. Šiklová (Eds.), *Czech Feminism. Perspectives on Gender in East Central Europe* (pp. 270–283). Bloomington and Indianapolis: Indiana University Press.
- Valdrová, J. (2017). Generické maskulinum. In P. Karlík, M. Nekula & J. Pleskalová (Eds.), *Nový encyklopedický slovník češtiny*. Retrieved from <https://www.czechency.org/slovník/GENERICK%C3%89%20MASKULINUM>
- Valdrová, J. (2018a). *Reprezentace ženství z perspektivy lingvistiky genderových a sexuálních identit*. Praha: SLON.

- Valdrová, J. (2018b). Obourodá jména: Nenechte se odbýt! *Prague Pride Magazin, Jaro 2018*, 21. Retrieved from <https://praguepride.cz/cs/kdo-jsme/media-download/prague-pride-mag/19-prague-pride-mag-jaro-2018/file>
- Valdrová, J. (2019). K naléhavosti genderových analýz jazyka na příkladu genderonomastiky. In V. Sokolová & L. Kobová (Eds.), *Odvaha nesouhlasit. Feministické myšlení Hany Havelkové a jeho reflexe* (pp. 450–464). Praha: FHS UK.
- Věšíňová, E. (1998). Úvaha na téma „Čeština a žena“. *Naše řeč*, 81(1), 21–28. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=7418>
- Z našich časopisů (1922). *Naše řeč*, 6(1), 24–25. Retrieved from <http://nase-rec.ujc.cas.cz/archiv.php?art=1376>

## ČEŠKO JEZIKOSLOVJE DRUŽBENEGA SPOLA: TEME, STALIŠČA, PERSPEKTIVE

Namen članka je analizirati obstoječe objave, ki se ukvarjajo s češkim jezikoslovjem družbenega spola, in stališča njihovih avtorjev. Avtorja prispevka najprej pregledno prikažeta razvoj področja, začetke z zavračanjem tovrstnega pristopa in njegovo današnje sprejemanje. V nadaljevanju se ukvarjata z različnimi načini poimenovanja oseb glede na družbeni spol in se sprašujeta o možnosti tvorjenja samostalnikov za poimenovanje žensk kot eno od strategij za doseganje njihove vidnosti v jeziku. Predstavita tudi psiholinguistične teste za ugotavljanje "generičnosti" moškega spola v češčini in možnosti za spolno vključujočo češčino. Avtorja tudi razpravljata o osebnih lastnih imenih s perspective družbenega spola in posameznikove identitete, pri čemer se na kratko ustavita ob prevodnih publikacijah, v katerih se odraža specifika družbenega spola. Na koncu pozove-ta k nadaljnim raziskavam na nekaterih področjih raziskav družbenega spola.

**Keywords:** češčina, družbeni spol, generični moški spol, feminizacija, spolno vključujoči jezik, besedotvorje, prevodoslovje.

## **THREAT PERCEPTION IN ONLINE ANTI-MIGRANT SPEECH: A SLOVENE CASE STUDY**

**Rok CHITRAKAR**

Faculty of Arts, University of Ljubljana

*Chitrakar, R. (2020): Threat perception in online anti-migrant speech: A Slovene case study. Slovenščina 2.0, 8(1): 66–91.*

DOI: <https://doi.org/10.4312/slo2.0.2020.1.66-91>

The aim of this article is to describe the perception of refugees as a threat in Slovene online discourse, based on a critical analysis of commenters' responses to popular media posts at the height of the European migrant crisis. The proposition of the study is that the perception of migration as a threat is at the core of socially unacceptable discourse (SUD), portraying refugees and migrants as an undesirable and potentially dangerous presence. Within the framework of a comprehensive project examining public responses to media coverage of the arrival of migrants to Slovenia, online comments classified as SUD targeting refugees were extracted and annotated to reveal the recurring themes of threat perception. The analysis focused on describing the main categories of threat, as well as the various discursive features and strategies employed. Although the approach to observing this subject is essentially qualitative, a general case-specific overview of the frequency and distribution of identifiable categories is also given.

**Keywords:** corpus analysis, critical discourse analysis, anti-migrant speech, hate speech, European migration

### **1 INTRODUCTION**

In the wake of the war in Syria and the growing involvement of foreign armed forces in the wider regional conflict, Europe was faced with increasing numbers of migrants entering its territories in search of refuge and assistance. The arrival of migrants to Europe reached mass proportions in the second half

of 2015 and the beginning of 2016 (Kogovšek Šalamon, 2017), demanding a large-scale, concerted response from the European Union and its member states. Countries along the route leading to preferred destinations (Germany or other Western and Northern European countries) were particularly affected, having to provide safe passage as well as the necessary administrative facilities and reception infrastructure to accommodate large numbers of arrivals. For Slovenia, the turning point came when Hungarian authorities decided to prevent migrants from boarding trains and continuing their journey toward Austria and Germany and, ultimately, closed the country's borders with Serbia and Croatia (BBC, 2015). With Hungary securing its borders and Croatia unwilling to accept returns, the migrants' path was rerouted through Slovenia, placing it within the newly established humanitarian corridor, in operation from October 2015 until March the next year. The countries along the Western Balkan Route, running through Greece, Macedonia, Serbia, Croatia, and Slovenia to Austria and Germany, had to provide transport, food, shelter and medical assistance to migrants in transition, which put a significant strain on their capacities and resources.

Officials, the civil society and humanitarian organizations all contributed to managing the situation by providing staff, facilities and logistical support (Ladić & Vučko, 2016). However, due to exceptionally high numbers of migrants arriving at the time, it became evident that these countries were largely underprepared and not able to rely on an adequate response mechanism to deal with these exceptional circumstances. Migrants were often met with increasingly difficult conditions, having to face long waits in overcrowded and poorly managed ad hoc reception centers and fenced off camps. Living conditions in these areas were at times bad enough to threaten their survival en route and could, in the worst cases, be interpreted as a clear violation of human rights. The corridor continued to be in operation until March 2016, when Turkey, within the framework of the EU-Turkey Statement (European Council, 2016), started opening its labor market to Syrians under temporary protection, accepted a large investment into its refugee reception facilities, and agreed to rapid returns of migrants declared ineligible for international protection. These measures led to the decrease in the use and gradual phasing out of the corridor, with the remaining groups of migrants applying for

asylum. Nationals from Syria, Iraq and Afghanistan (SIA) were granted special status, while others, undocumented SIA citizens included, encountered less favorable conditions and were not allowed to proceed with their journey (Kogovšek Šalamon, 2017, p. 260). Because of widespread concerns over the authorities' ability to manage the migrant situation and rising tensions in refugee hotspots, at times resulting in violent altercations and material damage, the Slovene media extensively covered this issue. Continuous exposure in conjunction with Parliamentary debates and various civil initiatives, such as awareness raising campaigns, framed the general public discourse. The exchange of opinions was particularly intense in social media, where Twitter and Facebook were permeated with argumentative and oppositional discourse. The social media platforms were transformed into a repository of different forms of socially unacceptable discourse (SUD), ranging from the more implicitly offensive and discriminative comments to overtly threatening hate speech and incitation to violence.

This article aims to examine and describe some of the ways in which migrants and refugees are perceived in socially unacceptable discourse in Slovenia. The basis for this discussion will be the assumption that participants in anti-migrant online discourse mostly act and react out of a sense of fear or threat linked to the presence of migrants directly, or to the presumed negative effects of migration as a phenomenon in general. To support this claim, we will rely on the principles of critical discourse analysis (van Dijk, 2001), a theoretical framework for describing how social dominance and relations are asserted through language use. The role of discourse in establishing social inequality and maintaining the distribution and reproduction of power (Fairclough, 2001) has already been amply discussed and demonstrated. As an approach, combining corpus studies, based on extensive datasets of actual language use, with critical discourse analysis (CDA) has proven to be efficient in researching discursive features (Baker, 2008), which is especially relevant when analyzing discourse aimed at vulnerable groups, such as migrants and asylum seekers.

According to Hatebase, a large database of multilingual hate speech, extreme forms of SUD most frequently target individuals on the basis of their<sup>1</sup> 1) ethnicity, 2) nationality, and 3) religion (Sindoni, 2018, p. 270). This finding

---

1 In order of importance.

places refugees and migrants in the context of the European migration crisis (for example Arab, Syrian, Muslim) among those most likely to find themselves at the center of discrimination, exclusion, and socially unacceptable discourse. Computer-mediated communication (CMC), interactional discourse retrieved from social media genres including popular social networks (Beißwenger, Horsmann & Zesch, 2018), represents an important segment of language use that significantly contributes to the formation of new social practices and accepted concepts of normality, within the confines of which traditional meanings and established power relations are maintained and consolidated (Gorjanc & Fišer, 2018; Gorjanc & Fišer, 2020). This study will attempt to provide some insight into how online socially unacceptable discourse, as an important component of increasingly present discriminative and exclusive attitudes, forms the perceived identity of refugees and migrants in Slovene society.

## **2 RELATED RESEARCH**

Research dealing with the issue of SUD directed at migrants, either on the level of conflict-driven migrations in general (Baker & McEnery, 2005; Wodak & Boukala, 2015) or within the specific context of Slovenia and its role in the European migration crisis (Erjavec & Kovačič, 2012; Ladić & Vučko, 2016; Vezovnik, 2017; Vehovar et al., 2020), already points out some of the prevalent features of anti-migrant speech that are likely to occur during analysis. From the point of view of threat construction and reproduction, various forms of negative portrayal of migrants as a distinctive social group by means of risk-related framing, othering, dehumanization, and vilification will be at the forefront of our attention.

Framing, a process by which the media and other influential discursive agents (e.g. political actors) emphasize a certain aspect of reality (Keren, 2011), is employed to establish specific meanings and ways of interpreting issues or events, such as the arrival of migrants. This is an important factor that plays a crucial role in predetermining the tone of discussion and, as certainly is the case with migrants in Slovenia, provides fertile grounds for SUD to develop and prosper. By focusing on a certain subject and consistently framing it in terms of risk and danger, discursive agents can already set the main themes



around which SUD will revolve. Negative representations of migrants in media are common, and their portrayal as a threat to society is typically shaped along several main argumentative lines. Innes (2010), examining the framing and perception of asylum seekers in British media narratives, recognizes three distinct categories of threat: 1) physical threat to society and its members, formulated in terms of sovereignty and security, 2) economic threat, wherein migrants are shown as undermining or exploiting social welfare, and 3) cultural threat, related to endangering and eroding social and cultural values.

In order to legitimize the construction of threat in socially unacceptable discourse, a clear distinction is made between the group presented as the threat and the group believed to be threatened. This is largely achieved by establishing an *us/them* opposition, where a rhetoric of exclusion is employed to delimit a European *us* from the non-European *other* (Wodak & Boukala, 2015). The first is usually represented as the potential victim, while the latter is portrayed as the probable offender to whom an infinite number of negative qualifiers and characteristics may be ascribed. Within various mechanisms of othering, migrants and refugees are commonly conflated into the out-group, culturally dissimilar from the host society and politicized as an invasive element that should be feared and distrusted (Huysmans, 2006). In addition, the process of social categorization, reinforced through language used in social media, effectively contributes to solidifying ideas of in-group homogeneity and solidarity between those under presumed threat (Waldron, 2012).

The negative representation of migrants should also be viewed in the context of securitization (Buzan et al., 1998), where critical discourse analysis can be used as a tool to explore the textual mechanisms behind the perception of migrants as a security threat. Discussing the media representation of migrants as a security threat in Slovene public discourse, Vezovnik (2017, p. 48) points out the common phenomenon of dividing arrivals to Slovenia into the categories of the *genuine* and the *bogus*, a clear example of a discursive attempt to criminalize refugees by turning them into subjects deserving increased suspicion. She also emphasizes that references to securitization on the Slovene-Croatian border could be interpreted as an indication of how *talk about protection and defense* may lead to actual security actions (2017, p. 50).

Another function of extreme forms of socially unacceptable discourse is to degrade and dehumanize target individuals or groups, stripping them of human characteristics that would generally warrant empathy and solidarity. When, for example, migrants are portrayed not as human individuals, but as an unfeeling personification of a natural disaster on a collision course with the host society, harsh and decisive counter-measures may seem much more acceptable. Similarly, the notion of mass, as often deployed in SUD, can be used as an effective discursive strategy. Van Dijk (1991) refers to the *aggregation technique*, by way of which the targeted group of discourse participants is objectified and treated as statistics. The use of numbers and figures as a discursive tool has a double effect; it gives the impression of neutrality, of merely reporting the factual truth, while social groups represented in the form of aggregated data or numbers can easily create the sense of *too many* (Ve-zovnik, 2017, p. 47).

The examples stated above are among some of the common features present in anti-migrant discourse that this study will attempt to highlight. In the framework of critical discourse analysis, examples of SUD and hate speech offer abundant material for criticism and discussion. One of the principle goals of such research is to expose and condemn the construction of discriminative and exclusive discursive practices that can, in the worst of cases, lead to violence against targeted groups and individuals. However, an additional aim of this article is to gain a deeper understanding of the root causes for the development and reinforcement of such practices. As suggested by Erjavec and Kovačič (2012), it is also important to focus on the individuals who produce SUD and hate speech in the attempt to understand their motivations for inciting hatred. By doing so, we might be able to move away from the individual conflicts and opposition between social actors and groups participating in the argument, and gain insight into some of the broader, systemic issues that lie underneath.

### **3 RESEARCH CONTEXT AND METHODOLOGY**

Before describing the methodological approach applied to dataset selection and analysis in this study, a more general presentation of the research context should be given. This work is part of a wider, comprehensive project examining

the issue of socially unacceptable discourse in a multidisciplinary framework combining corpus linguistics, CDA, legal analysis and sociological survey methods (Fišer et al., 2017). With a view to researching the construction, use and perception of SUD in a multifaceted sociocultural context, the project partners (the Department of Knowledge Technologies at the Jožef Stefan Institute, the Faculty of Arts, the Faculty of Social Science, and the Peace Institute, all in Ljubljana) developed large corpora of computer-mediated communication (CMC) and SUD, focusing on social media and online networking sites. As textual forms of CMC offer easily accessible and downloadable language data, they were chosen as the basis for an interdisciplinary sociolinguistic analysis. In the case of the SUD corpus, a list of Facebook groups in which such forms of speech frequently occur was compiled and used as a source for harvesting targeted data that would be included in the final dataset. After identifying the most visited Slovene mainstream media websites (24ur.com, SiOL.net.Novice, and Nova24TV), their Facebook pages were selected for inclusion in the repository of SUD comments.

The crucial phase before any analysis could take place was to produce an annotation schema and guidelines, according to which the type and target of socially unacceptable discourse could be determined. A two-dimensional typology was developed, taking into account six types of SUD and five targets at which SUD is directed (Ljubešić et al., 2019, p. 4). In relation to this study, SUD could be perceived as (1) aimed at migrants, (2) aimed at individuals or groups related to migrants, (3) aimed at media or journalists, (4) aimed at other commenters participating in the discussion thread, or (5) aimed at someone else. In terms of type, the main distinctions in annotating individual comments were made when determining whether (YES/NO) SUD was aimed at background (e.g. religion, ethnicity), or other groups/individuals, and whether it contained elements of violence. A manual annotation campaign was carried out after specific training was given to a select group of annotators (Ljubešić et al., 2019, p. 5) to ensure the best possible level of consistency and consensus in the results acquired. In the next phase, an in-depth analysis of the manually annotated dataset was performed, highlighting several findings relevant to this study. Perhaps the most revealing conclusions came from an overview of annotation distribution, which showed that only about half of

the annotated comments were classified as socially acceptable, a low number in the context of a public discussion. One of the conclusions that could be considered especially alarming pointed out that SUD inciting violence toward migrants was present in seven percent of annotations (Ljubešić et al., 2019, p. 8). This confirms the potential harmfulness of SUD as a form of widespread discursive practice that needs to be thoroughly researched and continually brought to the attention of regulatory actors and the general public alike.

In the second stage of annotation, the dataset was adapted specifically to the aims of this study, which required focusing predominantly on the anti-migrant component of socially unacceptable discourse. This entailed narrowing down the initial dataset pertaining to the general topic of migrants by reducing it to those comments where migrants were directly or indirectly targeted. In other words, comments in which SUD was aimed at media, other participants in the discussion thread, or someone else entirely, were deleted from the annotated dataset to be used for subsequent analysis. As the principal subject of this study was to analyze how negative and discriminatory perception of migrants was constructed and reproduced by anti-migrant commenters, it seemed relevant to create a clear opposition between the in-group and the out-group by excluding other actors who entered the discourse dynamic in a more complementary way. The final dataset contained three possible combinations of annotated comments: (1) comments referring to background, aimed at migrants, using offensive language, (2) comments referring to background, aimed at migrants, inciting violence, and (3) comments referring to others, using offensive language, related to migrants.

#### **4 ANALYSIS AND RESULTS**

After the comments relevant to this study were selected and annotated, a preliminary qualitative analysis was performed to obtain a general sense of how SUD aimed at migrants was reflected in language use and the argumentative strategies employed. This proved to be a critical moment in the study, as it determined the direction of the adopted approach. The principal observation that became apparent during the first analysis of the selected comments was that SUD targeting migrants could indeed be identified as a form of response to various, recurring types of threat perceived by discourse participants. The

comments that contained an argumentative element, as opposed to being purely offensive or inciting violence, could all be grouped according to several major topics repeated throughout the threads. After completing the analysis which focused on the different types of threat perception reflected in the comments, the following major categories were identified: (1) physical threat, (2) threat of Islamization, (3) threat of infiltration, (4) economic threat, (5) sociocultural threat, and (6) referential threat.

Another round of annotation was now performed to indicate the category of threat perception that each of the comments belonged to. This was done by an individual annotator (the author of this paper) with the purpose of facilitating a more detailed qualitative analysis. In addition to enabling a more systematic approach to examining anti-migrant SUD by grouping the comments into more or less distinctive themes, this method also made it possible to obtain a general overview of their distribution. First, a comparison was made between comments that contained a mention of threat perception related to migrants, and comments that contained no such reference. The latter were examples of purely offensive comments, hate speech, or outright incitation to violence against migrants, and had no argumentative value. Considering the fact that SUD in which threat perception was identified is divided into six additional categories, this indicates a significant prevalence of comments (nearly half) where no reason or justification for rejecting or targeting migrants is provided or deemed necessary.

**Table 1:** *Percentage of threat perception identified in SUD*

SUD, threat perception identified	SUD, threat perception not identified
58 %	42 %

This was followed by a breakdown of different categories of threat perception in terms of distribution, yielding the following results:

**Table 2:** *Breakdown of threat perception categories*

Physical threat	33 %
Threat of Islamization	22 %
Threat of infiltration	19 %
Economic threat	14 %
Sociocultural threat	7 %
Referential threat	5 %

It should be noted that these results are of indicative nature, as additional annotation was only carried out by one person and the quantitative distribution of individual themes was, to a certain extent, predetermined by the content of posts to which commenters were responding. However, some of the topics identified in this case study have proven to be prevalent throughout the threads analyzed. Viewing migrants as a potential physical threat was clearly the main concern for most of the commenters participating in the discussion. The fear of imminent danger in the form of physical violation or even death is, perhaps not surprisingly, much greater than, for example, anxieties related to the possibility of sociocultural disruption. The threat of Islamization and the threat of infiltration, at times overlapping and appearing within the same comment, seem to carry comparable weight. On the following pages, a more detailed description of individual categories, with examples taken from the analyzed threads, will be given to highlight the diverse discursive strategies and approaches to portraying migrants as a form of threat in socially unacceptable discourse.

## **5 DESCRIPTION OF IDENTIFIED THREAT CATEGORIES**

### **5.1 Physical threat**

The most easily identifiable and largest category of threat, directly attributed to the arrival of refugees and migrants, reflected the commenters' fear of material damage, outbreaks of violence, violation of physical integrity and, ultimately, death. The potential of physical danger underlying negative attitudes to migration elicited the most emotional and violent responses. In the context of the Slovene migration crisis, commenters often alluded to the need to regain control over an out-of-hand situation to protect themselves and, perhaps more importantly, their families. Women and children, in particular, were singled out as the most vulnerable group in need of protection. To support and justify these concerns, commenters often resorted to citing recent events, such as cases of rape and violence reported in relation to migrants, mediatized across Europe. Western, non-Muslim women were claimed to be most exposed to possible physical attacks by refugees and migrants, as they were considered most likely to present an affront to the Muslim faith.

The safety of public spaces where migrants were now a growing presence was also called into question, especially in terms of limited freedom of movement

and obstacles to the established way of life (*not being able to go out at night*). Generally speaking, in a post 9/11 world and in the light of terrorist attacks occurring across Europe (*the Paris attacks, what happened in Brussels*), refugees and migrants as the out-group were identified with *terrorists, criminals, savages, aggressors, extremists, combatants, jihadists, hooligans, and ISIS*. The visual portrayal of migrants in the media also played an important role in forming the general perception of physical threat. Commenters often referred to the fact that newspaper images and televised footage showed groups of *strong, combat-ready, young, 16-45 year old men*, which likely contributed to establishing associative links with organized units, military stratagems and planned invasion.

Preference for the visual framing of refugees as a threat was most evident in graphic references to scenes of extreme violence, mentioning *decapitation, throats cut, bombs exploding, stoning, and child bombers*, among others. Extremist iconography specific to acts of terrorism and war (*black flags, Kalashnikovs, beards, shouts of Allahu Akbar*) was frequently used to evoke a sense of dread at the arrival of Muslim population. Also on the symbolic level, important landmarks of the Muslim community in Europe were given negative connotations. Mosques and Muslim cultural spaces, for example, were associated with *terrorist training centers* and images of *children cutting heads off teddy bears*. Another aspect of physical threat that was widely referred to were alleged health hazards and security risks associated with the arrival of refugees to Europe. This was in line with similar discursive approaches aimed at dehumanizing migrants who were, in this case, depicted as *carriers of disease, dirty, defecating in the streets, leaving trash everywhere*, and even as a *cancer, ready to metastasize across Europe*.

The perception of physical threat also established the parameters of internal dispute present throughout the threads analyzed. The main argumentative opposition was created between those who categorically rejected refugees and those who defended them, namely pro-migrant activists who participated in protest rallies and welcomed migrants at the borders. The latter were seen as facilitators in the process of migration, and as such an important part of the problem (*idiots, traitors*) that, much like refugees, needed to be *eradicated*. Interestingly, this opposition also encouraged the usual antagonism between

the center (Ljubljana, the capital of Slovenia) and the periphery, the center generally being seen as more liberal and supportive of migrant integration. In addition, the government and political actors were blamed for failing to adequately address national security concerns and putting the country and its citizens at risk by helping *the enemy*. Throughout the threads, commenters emphasized their concerns over the potential physical threat posed by refugees by calling on the government to *protect their own*. Along with the anti-migrant public, pro-migrant activists, and the media, government actors formed the discursive frame for the entire discussion.

### 5.2 Threat of Islamization

The next major topic that emerged from examining the comments for identifiable types of threat perception was centered on Islamization. In this case, migrants were perceived as the driving force of a potential Muslim invasion of Europe. In certain regards, this category of threat perception overlapped with the fear of physical harm, but this was conveyed indirectly, by referring to religion. The Muslim faith was portrayed as the basis for extremist violence and terrorism, giving migrants and refugees a *license to kill, break our laws because their religion tells them to, or enforce their religion*. Commenters also resorted to quoting decontextualized parts of the Qur'an to illustrate Muslim migrants' intentions. This proved to be a popular argumentative strategy, supposedly providing proof that needs no further explanation, but *speaks for itself*. Again, the threat to the *European way of life* was placed in the foreground (ex. *not being able to eat pork*), with women repeatedly mentioned as the primary victims, especially in terms of *losing rights* and diminished equality.

The concept of invasion was at the core of the discussion. Most of the fears revolved around migrants *invading our homes, taking our property, or taking the throne*, a process through which Europeans would become subjugated, even *enslaved under the Muslim rule*. This hypothetical invasion of Europe was characterized as *calculated and planned*, as a gradual transformation in a *grand scheme of global domination*, implemented methodically and systematically through the construction of mosques and Muslim schools, as well as the *spread of Muslim ghettos*. Some commenters claimed that one of the



main threats associated with migration from Muslim countries would be substitution. This would occur on several levels, a logic according to which *the Constitution would be replaced with Sharia law, universities with madrasas and women's skirts with headscarves.*

Conspiratorial undertones were present throughout this segment of discussion with theories proposed of Arab countries financing or *bribing* the authorities to allow invasion as a form of *neocolonialism* that would eventually lead to the *destruction of European cities*. Referring to historical experience was also one of the favored lines of argumentation, aimed at reducing the Muslim population to the role of invader within the context of an envisioned large-scale clash of civilizations (*they have always been the enemy, What did our forefathers fight for?*). Cautionary examples were also provided, with allusions, for instance, to the French who *opened their doors and are now the minority, with 64 % of schoolchildren being Muslim.*

This brings us to the question of statistics, often provided as argumentative evidence in favor of the validity and credibility of posted comments. Commenters opposing the arrival of refugees seemed to have a wide array of numerical data at their disposal whenever they needed to illustrate the severity of the migrant situation. However, when these statistics are not supported with a clear reference to their source, one has to wonder about their veracity. It is important to note that such data, tailored to fit a specific purpose and sometimes even altogether invented, can serve as an effective discursive device (van Dijk, 1991). When it comes to constructing threat as a means to legitimate socially unacceptable discourse, describing a situation where *10 000, 20 000, or 50 000 refugees* are crossing the border into a small and underprepared European country can paint a convincing picture.

In fact, responding to the threat or fear of invasion is often related to various ways of interpreting and expressing the central notion of mass. When discussing the issue of mass migration in terms of amount, commenters tended to highlight the omnipresence and multitude of migrants who, in their opinion, were *everywhere, on our front lawns, in the fields, around our houses*, and in great numbers, *hordes* that would eventually *run us over*. Parallels were frequently drawn between migrants and natural disasters, typically in the form of metaphors of water flow, calling our attention to the risk of *drowning* or

becoming *submerged, overflowed* and, conversely, to the necessity of *building a dam, stemming the flow*. Other dehumanizing comments, related to the threat of mass invasion, portrayed migrants as an infestation (*vermin, rats*) or an epidemic (*plague*), and even expanded to legal interpretations of the EU provisions on family reunification (Council Directive 2003/86/EC), which might enable migrants to *abuse rules in order to multiply*.

### 5.3 Threat of infiltration

The threat of invasion in the context of migration often presupposes an act of infiltration committed by members of the out-group. This was identified as the next category of threat perception within which the status of refugees is constantly questioned or, in most cases of anti-migrant speech, altogether rejected. The logic of undermining the status of migrants escaping areas of war and conflict is based on the imperative that only individuals who can be classified as refugees according to international law are deserving of help. Immediately, this precept introduces the hypothetical possibility, and consequently threat, of persons willing to abuse such legal provisions and attempt to enter a country as *fake refugees*. To add an additional layer of complexity, the issue of identifying and correctly using the notion of *refugee*, as compared to *migrant, immigrant* and *emigrant*, seems to cause a fair amount of confusion in public discussions, as these terms are often used interchangeably, without discerning relevant semantic nuances. That aside, the most frequently used discursive technique commenters applied to dealing with the threat of potential infiltrators almost invariably followed the formulaic *They are not refugees, they are...* The blank space would then be filled with any number of qualifiers readily available to discredit the status and intentions of refugees, among which the most common were *terrorists, criminals, economic migrants, opportunists, cowards, traitors*, and *fake, so-called, illegal or quasi-refugees*.

Migrants were generally subjected to a form of group profiling, especially on the grounds of the images shown in mainstream media as well as other, alternative sources of information. For example, the repeatedly emphasized image of *mostly male, young, combat-ready, healthy, strong* men was quite evidently used to insinuate that migrants were to be treated as a threatening

new presence with a hidden, sinister agenda. Determining migrants' age and sex was mentioned as the crucial stage when screening for potentially *trust-worthy* refugees. Again, the risk of persons attempting to abuse established legal provisions, such as the frequently cited unaccompanied minors clause (Directive 2011/95/EU of the European Parliament and of the Council), was brought to the foreground with accusations of migrants *lying about their age* and references to *bearded children*. Another argument that was used to undermine refugees' credibility was directed purely at personal appearance; the idea of war refugees still seems to be visually rooted in images of starvation and torn clothing, as opposed to *nice clothes and expensive smartphones*.

Throughout the timeline of the posted threads, media reports provided continuous support for rejecting refugees. The newcomers' unwillingness or inability to present acceptable identification as proof of their status, a precondition for assistance and possible entry, was a particularly salient point (*If they really are refugees, why won't they prove it?*). In general, the behavior of migrants during the reception and identification procedure at the borders was extensively covered by the Slovene media and closely scrutinized by the public. When questioning the status and eligibility of migrants as a group, commenters relied on specific, widely mediatized events showing individuals refusing assistance (Slovenske novice, 2015). Rejecting food and clothes, for example, was perceived as *ungrateful*, in turn supporting the claim that migrants did not deserve to be helped. It was also made clear that refugees were not expected to complain or make demands, and should be content with any kind of assistance offered.

The migrants' demands to choose their country of destination were heavily criticized, but sparked a curious and almost contradictory discussion in which commenters were rejecting their credibility while, at the same time, struggling with feelings of wounded pride (*We don't want them, but why won't they stay?*). This line of argumentation (Ladić & Vučko, 2016) was also complemented by the first safe country principle (Regulation 604/2013/EU of the European Parliament and of the Council), according to which refugees were expected to seek asylum as soon as they escaped danger, and not to continue travelling *in search of a better life*. Even when the status of refugees was recognized, attempts were made to undermine their character. In this case,

migrants and refugees were portrayed as *cowards, deserters, or traitors who gave up their homeland without a fight, and left their elderly and women behind*. Within this category of threat perception, a significant amount of distrust was also aimed at the media, who were accused of intentionally *portraying migrants as a vulnerable group, in need of help*. Commenters were particularly active in highlighting the media's agenda of manipulating the public opinion by *selective reporting*, only displaying certain content while *covering up the negative*. This is an interesting claim, especially considering the fact that anti-migrant commenters have a clearly ambivalent and selective relationship with media, whose information can be used or discarded at will, according to discursive needs.

#### 5.4 Economic threat

The following category of threat perception was based on rejecting the status of refugees who were generally labeled as *economic migrants*, in order to emphasize what commenters believed to be the true nature and purpose of migration. Even though migrants and refugees remained the nominal target of socially unacceptable discourse within this topic, in comparison with the other recurrent themes, the argumentation provided here was much more a reflection of the commenters' discontent with the general economic situation. The issue of poverty in Slovenia seemed to be at the core of commenters' concerns, pointing out that while *our children are hungry and parents aren't able to afford school meals*, refugees could not be treated as a priority. Personal struggles (*I can barely make it through the month myself...*) and the added fear of having to provide for refugees' subsistence could be seen as a decisive factor in refusing to accept migrants. In this context, political actors were singled out as the main culprits responsible for the failing economy, and the government was accused of *not providing for their own citizens while giving more benefits to refugees than to Slovenes*. The media attention given to refugees was also a sore spot (*When have they ever cared about us, talked about our problems?*) and, in response, vulnerable members of the in-group, i.e. Slovene citizens, were prioritized (*the children, the elderly, the unemployed, the indebted*).

The presence of refugees was perceived as a potential impediment to improving the economic situation as well as a burden to the social welfare system that

could find itself at risk if the State needed to provide for migrants. The willingness to accept and integrate refugees was generally equated with transferring funds from social welfare to migration policy, for example from *child support and schooling* to refugee reception facilities. A strong sense of ownership over the way the national budget was employed in dealing with migration was also expressed, further stressing the imminence of the supposed threat (*taxpayers' money, our money, money from our pockets*). Among the comments directly targeting migrants as agents of possible economic ruin, many were directed at portraying them as *lazy, not willing to work for their money, or not used to work*, with frequent insinuations about an alleged growing criminal presence (*thieves*). Commenters also resorted to dehumanizing strategies within this segment of anti-migrant speech, mostly by referring to the burden of having to *feed these so-called refugees or leeches*, and insisted on shifting financial responsibility for doing so to those willing to help (namely pro-migrant activists and politicians).

### 5.5 Sociocultural threat

When discussing the perception of migrants and refugees as a sociocultural threat in Slovenia, we will be focusing on how migratory flows were seen and presented as a disruptive force in the societal and cultural landscape. The starting point of argumentation within this category of threat perception was the assumption that refugees and migrants were unwilling or unable to integrate with the host society. Among the principal claims propagated by commenters was the categorical statement that migrants *are not capable of respecting our legal order and would try to enforce their order and traditions*. General judgments, such as *they're not good at integrating*, were highlighted as an accepted truth, supported by references to migrants' insistence on pursuing a *completely different lifestyle* (wearing headscarves in public, for example).

Constructing the image of out-group inferiority was another popular discursive strategy. The alleged difficulty in integrating migrants was, for instance, attributed to their *not speaking English, not being educated or not being able to understand our laws, our ways*. To reinforce the validity of this negative image, commenters referred to past experiences with refugees, some of who *have lived here for decades and still haven't learned to speak a word of*

*Slovene*. Where refugees and migrants were given an active role in terms of responsibility for integrating, the discussion was deeply anchored in the paradigm of a host-guest dynamic, according to which *they should try to adapt, should be respectful and should not get too comfortable*. In a reversal of roles, a much more positive light was shed on members of the in-group who would *accept their rules and their culture if they were in their place*, setting an example of proper behavior.

A clear conditionality was attached to the possibility of welcoming refugees and migrants, with commenters stating they *would be willing to accept them if they gave up their ways and lived according to our rules*. This could, however, prove to be a difficult condition to fulfill, as it implies the necessity of migrants discarding their identity in order to erase their otherness. Sometimes the fear of otherness, as reflected in the comments analyzed, was of purely superficial nature, and migrants were perceived as a foreign element with which contact was undesirable and difficult to imagine (*women dressed like mummies sitting in my living room*). On the other hand, the difficulty of including and integrating migrants in the host society was viewed as an inherent and unchangeable part of the problematic itself. Clear lines of separation were drawn to demonstrate the incompatibility of migrants with the Slovene society and culture, by emphasizing the inevitable clash between mentalities and, in particular, religions (*Islam does not allow for integration*).

Integration in the context of religious identity was also seen as improbable due to migrants' supposed isolationism, their tendency to function exclusively within *family units, closed communities, or Muslim ghettos*, and build their own environment (*mosques, schools, media*) instead of adapting. With this, the logic underlying the various claims of dangers contained in possible non-integration came full circle; the failure to integrate migrants and refugees could lead to their marginalization, which could in turn foster feelings of discontent and resentment, ultimately leading to physical retribution. Perhaps to alleviate the overarching sentiment of xenophobia, a curious comparison was made to Bosnian refugees that arrived to Slovenia during the Yugoslav Wars. Despite well-known discriminative and prejudiced attitudes against Yugoslav immigrants, particularly pronounced at the time, but still present today, they were now portrayed as model refugees, *acceptable due to linguistic and*

*geographic proximity, not savages but civilized like us.* As a more appropriate solution, commenters suggested that migrants who were *too different* should instead head to *neighboring countries where the way of life is similar and it's easier for them to integrate.*

### 5.6 Referential threat

As opposed to other identified groups of threat perception, the final category examined was related to migrants and refugees indirectly, usually by providing links to other articles and videos, by quoting influential figures (*Gaddafi said..., An interesting quote by Geert Wilders...*), or by explicitly referring to media as the incontestable source of information. The formulation of socially unacceptable discourse through various means of reference is a particularly interesting phenomenon. Here, the commenter is not the immediate source of the content given, nor does she explicitly agree with the content referred to. In the case of *The answer is clear ...* [https://www.youtube.com/watch?v=EZbCK95Zu\\_o](https://www.youtube.com/watch?v=EZbCK95Zu_o), or *This video says it all ...* <https://www.youtube.com/watch?v=X2rmA9A7CEM>, the analyst is required to proceed with a second phase of examination to verify the nature of the comment. Of course, other links may be explicit enough as to their aim and content without much further investigation, for example *This is what I thought:* <http://pamelageller.com/2015/09/document-muslim-migrants-raping-women-and-children-in-camp-in-germany.html/> or *And this:* <http://vladtepesblog.com/2015/09/12/rape-and-forced-prostitution-common-in-muslim-refugee-camps-in-germany/>, giving a clear enough indication of the statement contained.

Whether a link was sufficiently descriptive or required the implications of its contents to be deduced from a broader discursive context, the common denominator underlying this form of socially unacceptable discourse was most evident in the shift to a more personal level of commenting. Commenters were now referring to specific, heavily mediatized images and situations, and were speaking directly against subjects in photographs and videos, viewed as the main actors personifying the encroaching threat of migration. Audiovisual content was appended as self-evident truth, entirely exempt of critical thought (*We all saw on TV... Watch the footage...*), and provided a readily available opportunity to reinforce solidarity between like-minded commenters who could simply voice their agreement (*I agree, exactly right, well put*).



In addition to being a communicative facilitator and an effective argumentative strategy, establishing threat by way of reference also functioned as a tool of empowerment, specifically when opposing mainstream media or pro-migrant activists. When the need to bolster a personal opinion arose, commenters could readily access and select information, according to their discursive needs (*Just GOOGLE it!*). In and of itself, taking personal initiative in gathering information is by no means problematic and can indeed contribute to a better grasp of a multifaceted subject, such as migration. However, in the context of socially unacceptable discourse, the tendency to decontextualize facts and information most often only serves the creation and reproduction of potentially harmful biases.

## 6 DISCUSSION

The qualitative analysis of comments in Slovene social media, carried out within the framework of this case study, has confirmed the importance of the perception of refugees and migrants as a threat in the formation, consolidation, and reproduction of socially unacceptable discourse. Examining the discursive mechanisms and strategies underlying this practice provides insight into the various motivations for anti-migrant speech. It also reveals how key discursive actors, such as politicians and the media, contribute to the construction of the targeted out-group as a threat, by relying on narrative framing that effectively elicits feelings of fear that are often at the root of social exclusion, discrimination, and rejection.

The recurrent topics that are central to anti-migrant speech in Slovene online discourse have made it possible to identify several categories of threat perception, within which commenters typically present arguments to justify or legitimize their rejection of refugees and migrants. In order of importance, migrants and refugees were perceived as presenting a threat in the form of (1) physical danger, (2) Islamization, (3) infiltration, (4) economic burden, and (5) sociocultural disruption, with an additional category of (6) referential threat also identified to take into account arguments presented indirectly, by way of reference.

The quantitative distribution of comments according to different threat categories, although indicative and case specific, implies that the fear of physical



harm is at the forefront of motivations underlying the negative and discriminative attitudes against refugees and migrants, as reflected in SUD. This shows that the direct threat associated with the presence of migrants and refugees takes precedence over other, more abstract concerns, such as the potential for sociocultural disruption. The noticeable prevalence of the sense of physical threat among anti-migrant commenters would also be in line with the media's negative framing of the migration phenomenon, as well as the enhanced securitization at the political level.

This leads to questions about a possible feedback loop between different, multi-level types of responses that can perpetuate negative and exclusive attitudes against members of an out-group. Considering the fact that most of the examples of SUD found in online comments follow a similar argumentative logic and employ well-known and predictable discursive techniques, it is not difficult to imagine the automatization of anti-migrant speech in computer-mediated communication. As this presents an increased potential for spreading and reproducing harmful SUD and hate speech, adequate responses should be developed and implemented by policy-makers and social media platforms alike.

This is, of course, a complex legal issue, inseparably intertwined with the principles of fundamental freedoms, privacy and the rule of law, so finding a balanced solution would require a comprehensive multi-disciplinary approach. In terms of raising public awareness, CDA and other forms of sociolinguistic analysis should attempt to highlight the need to question some of the arguments presented, especially when stated as facts or accepted truths, relying on unverified statistical support and external references so freely given as "evidence". Such studies could also lead to a more thorough discussion about the issue of attributing and accepting social accountability for discriminative and exclusive practices in the context of public discourse.

In addition to the important role played by those wielding discursive power, both in pre-determining and shaping the established narratives pertaining to members of the out-groups deemed threatening, SUD could also be seen as an indirect, yet relevant, criticism of society. Even though most of the anti-migrant SUD, observed on the textual level in this case study, was aimed directly at migrants or their religious and ethnic backgrounds, many of the

arguments did, in fact, reflect Slovenes' growing dissatisfaction and frustration with the economy, politics, and the State itself. Along with the relevant work of revealing the harmful and widespread effects of socially and democratically disruptive discursive practices and highlighting the many forms in which language use allows them to manifest, an equal measure of attention should also be given to the broader context and the root systemic causes that lay beneath.

## **7 CONCLUSION**

The study confirmed the central role of threat perception in the formation of online anti-migrant speech as a subcategory of socially unacceptable discourse in Slovenia. In fact, some form of threat perception was identified in more than half (58 %) of the comments analyzed. This finding should be viewed in light of the fact that in the remaining 42 % of the cases, in which threat perception could not be pointed out, commenters gave no reason or justification for targeting migrants. These comments were, for the most part, examples of pure hate speech in the form of insults and incitation to violence. This means that whenever a comment was based on actual argumentative content, it could be linked to one of the categories of threat perception identified. The analysis of comments targeting migrants and refugees revealed six recurrent themes that permeated the debate about the European migration crisis.

The results indicated that anti-migrant speech was predominantly fear-driven, with one of the main concerns expressed related to the immediate physical threat allegedly posed by migrants arriving to Slovenia. Interestingly, the comments from this category of threat perception were the least substantiated and seemed heavily influenced by sensationalist media reporting. Arguments that provided more explanatory basis, such as those related to the fear of economic deterioration or sociocultural displacement, were far less frequent. Despite the fact that migrants were the primary target of SUD in this case study, a closer examination of the comments also confirmed the commenters' dissatisfaction and frustration with the state and the way the political leadership has handled the situation. The perception of threat in the context of anti-migrant speech often stems from difficult personal circumstances in which the arrival of migrants and refugees may be seen as an aggravating factor.

As referring to media reports and various opinion makers with the aim of legitimizing anti-migrant comments is a widely used strategy in online SUD, more research should be focused on how influential discursive actors contribute to establishing the framework for intolerant, discriminative and otherwise harmful speech. Accountability for disseminating negative attitudes toward vulnerable minority groups should be highlighted and discussed. An in-depth examination of discursive devices most frequently and effectively deployed in SUD is crucial to understanding, recognizing, and being able to respond to online hate speech. Also, an ongoing and regularly updated overview of online SUD would be useful in following current topical and argumentative trends that should be addressed in public debate when speaking against hate speech. It is particularly worrying that the many forms of socially unacceptable discourse so often remain unrecognized by those who help create it, until they become ingrained in the public conscience as part of the new normal.

### **Acknowledgments**

The author acknowledges the financial support from the Slovenian Research Agency (research core funding No. P6-0215).

### **REFERENCES**

- Baider, F., & Kopytowska, M. (2018). Narrating hostility, challenging hostile narratives. *Lodz Papers in Pragmatics*, 14(1), 1–24.
- Baker, P., & McEnery, T. (2005). A Corpus-based Approach to Discourses of Refugees and Asylum Seekers in UN and Newspaper Texts. *Journal of Language and Politics*, 4(2), 197–226.
- BBC News (2015). Migrant crisis: Hungary closes border with Croatia. Retrieved from <https://www.bbc.com/news/world-europe-34556682>
- Beißwenger, M., Horsmann, T., & Zesch, T. (2018). Part-of-speech tagging for corpora of computer-mediated communication: a case study on finding rare phenomena. In D. Fišer & M. Beißwenger (Eds.), *Investigating Computer-Mediated Communication: Corpus-based Approaches to Language in the Digital World* (pp. 192–219). Ljubljana: ZIFF.
- Buzan, B., Waever, O., & Wilde, de J. (1998). *Security: A new framework for analysis*. Boulder, Colorado: Lynne Rienner.

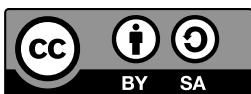
- van Dijk, T. A. (1991). *Racism and the Press*. London and New York: Routledge.
- van Dijk, T. A. (2001). *Critical discourse analysis*. In D. Schiffrin, D. Tannen & H. E. Hamilton (Eds.), *Handbook of discourse analysis*, 349–371.
- Erjavec, K., & Kovačič, M. P. (2012). You Don't Understand, this is a New War. Analysis of Hate Speech in News Websites' Comments. *Mass Communication and Society*, 15(6), 899–920.
- European Council (2016). EU-Turkey Statement, 18 March 2016. Retrieved from <https://www.consilium.europa.eu/en/press/press-releases/2016/03/18/eu-turkey-statement/>
- Fairclough, N. (2001). *Language and Power*. London, New York: Longman.
- Fišer, D., & Beisswenger, M. (Eds.). (2017). *Investigating Computer-Mediated Communication: Corpus-based Approaches to Language in the Digital World*. Ljubljana: ZIFF.
- Gorjanc, V., & Fišer, D. (2018). Twitter in razmerja moči: diskurzna analiza kampanj ob referendumu za izenačitev zakonskih zvez v Sloveniji. *Slavistična revija*, 66(4), 473–495.
- Gorjanc, V., & Fišer, D. (2020). Twitter discourse on LGBTQ+ in Slovenia. In D. Fišer & P. Smith (Eds.), *The Dark Side of Digital Platforms: Linguistic Investigations of socially unacceptable online discourse practices* (pp. 36–55). Ljubljana: ZIFF.
- Huysmans, J. (2006). *The politics of insecurity: Fear, migration and asylum in the EU*. London: Routledge.
- Innes, A. (2010). When the threatened become the threat: The construction of asylum seekers in British media narratives. *International Relations*, 4, 456–477.
- Keren, G. (2011). On the definition and possible underpinnings of framing effects. *Perspectives on framing*, 3–33. New York: Taylor and Francis.
- Kogovšek Šalamon, N. (2017). Mass Migration, Crimmigration and Defiance: The Case of Humanitarian Corridor. *Southeastern Europe*, 41(3), 251–275.
- Ladić, M., & Vučko, K. (2016). Slovenia's response to increased arrivals of refugees: "We don't want them, but we also don't understand why they don't want to stay." In N. Kogovšek Šalamon & V. Bajt (Eds.), *Razor-wired: Reflections on migration movements through Slovenia in 2015* (pp. 7–10). Ljubljana, Slovenia: Peace Institute.

- Ljubešić, N., Fišer, D., & Erjavec, T. (2019). The FRENK Datasets of Socially Unacceptable Discourse in Slovene and English. Retrieved from <https://arxiv.org/pdf/1906.02045.pdf>
- Merrill, S., & Åkerlund, M. (2018). Standing Up for Sweden? The Racist Discourses, Architectures and Affordances of an Anti-Immigration Facebook Group. *Journal of Computer-Mediated Communication* 23(6), 332–353.
- Sindoni, M. G. (2018). DIRECT HATE SPEECH VS. INDIRECT FEAR SPEECH: A multimodal critical discourse analysis of the Sun's editorial "1 in 5 Brit Muslims' sympathy for jihadis". *Lingue e Linguaggi*, 28, 267–292.
- Slovenske novice (2015). VIDEO: Begunci zavračajo pomoč Rdečega križa? Retrieved from <https://www.slovenskenovice.si/novice/svet/video-begunci-zavracajo-pomoc-rdecega-kriza>
- Vehovar, V., Povž, B., Fišer, D., Ljubešić, N., Šulc, A., & Jontes, D. (2020). Družbeno nesprejemljivi diskurz na Facebookovih straneh novičarskih portalov. *Teorija in praksa*, 57(2), 622–645.
- Vezovnik, A. (2017). Securitizing Migration in Slovenia: A Discourse Analysis of the Slovenian Refugee Situation. *Journal of Immigrant and Refugee Studies*, 16(1–2), 39–56.
- Waldron, J. (2012). *The Harm in Hate Speech*. Cambridge: Harvard University Press.
- Wodak, R., & Boukala, S. (2015). European identities and the revival of nationalism in the European Union: A discourse historical approach. *Journal of Language and Politics*, 14(1), 87–109.

## PERCEPCIJA GROŽNJE V PROTIMIGRANTSKEM GOVORU NA SPLETU: ŠTUDIJA PRIMERA

Članek na osnovi kritične analize odzivov komentatorjev na medijske objave v času evropske migracijske krize opisuje percepcijo beguncev kot grožnje v slovenskem spletnem diskurzu. Raziskava izhaja iz stališča, da je dojemanje pojava migracij kot grožnje v središču družbeno nesprejemljivega diskurza, v katerem so begunci in migranti prikazani kot nezaželen in potencialno nevaren element. V okviru celovitega projekta, ki je vključeval obravnavo javnih odzivov na medijsko poročanje o prihodu migrantov v Slovenijo, so bili za analizo izbrani komentarji, usmerjeni proti beguncem in migrantom. Relevantni komentarji so bili nato označeni z namenom opredelitve najpogostejših tematskih sklopov, vezanih na percepcijo grožnje. Besedilna analiza se osredotoča na opis glavnih kategorij domnevne grožnje ter na različne diskurzivne značilnosti in strategije, vsebovane v komentarjih. Čeprav obravnava pojava v večji meri temelji na kvalitativnem opisu, je v članku predstavljen tudi pregled pogostnosti in porazdelitve prepoznanih kategorij.

**Ključne besede:** korpusna analiza, kritična analiza diskurza, protimigrantski govor, sovražni govor, evropske migracije



To delo je ponujeno pod licenco Creative Commons: Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna. / This work is licensed under the Creative Commons Attribution-Share-Alike 4.0 International.

<https://creativecommons.org/licenses/by-sa/4.0/>

## OKROGLA MIZA »(BLIŽNJA) SREČANJA OBLIKOVALCEV JEZIKOVNE POLITIKE«

Ina FERBEŽAR, Igor CETINA, Alojz IHAN, Marko STABEJ, Lana ZDRAVKOVIČ in Tina ZUPANČIČ

*Ferbežar, I. et al. (2020): Okrogla miza »(Bližnja) srečanja oblikovalcev jezikovne politike«. Slovenščina 2.0, 8(1): 92–112.*

DOI: <https://doi.org/10.4312/slo2.0.2020.1.92-112>

V Ljubljani sta med 6. in 8. 11. 2019 potekala 54. srečanje in javni posvet združenja ALTE<sup>1</sup> (Association of Language Testers in Europe). Srečanje na temo Enojezično testiranje v večjezični resničnosti? Jezikovne ideologije in njihov vpliv na jezikovno testiranje sta organizirala Univerza v Ljubljani, Filozofska fakulteta in njen Center za slovenščino kot drugi in tuji jezik pri Oddelku za slovenistiko. V tem okviru je 8. 11. 2019 potekala okrogla miza (Bližnja) srečanja oblikovalcev jezikovne politike, na kateri so sodelovali **Igor Cetina** z Ministrstva za notranje zadeve RS, **prof. dr. Alojz Ihan** z Medicinske fakultete Univerze v Ljubljani, **prof. dr. Marko Stabej** s Filozofske fakultete Univerze v Ljubljani, **dr. Lana Zdravkovič** z ljubljanskega Mirovnega inštituta in **Tina Zupančič** iz podjetja Revoz.

Okroglo mizo je moderirala **dr. Ina Ferbežar** s Centra za slovenščino kot drugi in tuji jezik Filozofske fakultete ljubljanske univerze. Pri zapisu so bile izjave za namene lažjega branja skladenjsko prilagojene značilnostim pisnega jezika, vendar po principu minimalne intervencije.

**Ina Ferbežar:** Najprej bi se rada zahvalila svojim gostom. Lepo pozdravljeni. Zelo redko se srečamo oblikovalci jezikovne politike, takšni in drugačni odločevalci. Danes je naš namen, da se z mojimi sogovorci, ki prihajajo z zelo različnih področij, pogovarjamo o zapletenem razmerju med potrebo po

---

1 <https://www.alte.org/>

priseljavanju (recimo zaradi pomanjkanja delavcev na nekaterih področjih pri nas, zlasti v zdravstvu trenutno in turizmu), med dejanskimi sporazumevalnimi potrebami priseljencev, med pričakovanji jezikovne skupnosti glede njihovega znanja slovenščine in zahtevami države po tem znanju. In seveda na koncu o jezikovnih pravicah in dolžnostih vseh partnerjev v komunikaciji.

Moji gostje so – bom začela na svoji levi strani – doktorica Lana Zdravković, filozofinja, strokovnjakinja za migracije pri Mirovnem inštitutu in aktivistka na področju človekovih pravic. Za nas posebej zanimiva tudi zato, ker ima sama migrantsko preteklost oziroma migrantsko ozadje. Profesor Marko Stabej, nekateri ste ga danes dopoldne že slišali. Redni profesor za slovenski jezik na Filozofski fakulteti v Ljubljani in strokovnjak za jezikovno politiko. Gospod Igor Cetina je višji svetovalec pri Ministrstvu za notranje zadeve. Skrbnik projektov, namenjenih pomoči pri vključevanju migrantov v slovensko družbo. Profesor Alojz Ihan, zdravnik, redni profesor mikrobiologije in imunologije na medicinski fakulteti naše univerze<sup>2</sup> in predvsem tudi – kar je za nas še posebej zanimivo – tudi zelo priznan književnik, pesnik in esejist. Na koncu, na moji skrajni desni, Tina Zupančič, vodja službe za izobraževanje pri družbi Revoz (eno od podjetij v sklopu Renaulta), ki bo predstavila pogled delodajalca.

Hvala lepa, da ste prišli.

Jaz sem si tole zamislila bolj kot neformalni pogovor, čeprav so vsi moji gostje že dobili vsaj približna vprašanja vnaprej. Ampak domnevam, da bomo improvizirali. Odločitev, morda povem še to, zakaj v slovenščini. Promoviramo večjezičnost. Danes in te dni smo slišali v glavnem angleščino. To je en razlog. Drugi razlog, mogoče za nas, ki smo predstavniki univerze, še bolj pomemben, tudi simbolni, je ta, da je stoletnica ljubljanske univerze, univerze, ki ima učni jezik slovenski. Pred sto leti je bilo prvo predavanje v slovenščini o slovenskem jeziku. In tretji razlog: vsi smo govorci slovenščine. Bilo bi mogoče nekoliko neobičajno, da bi med sabo govorili angleško. Tako da, hvala za razumevanje.

Mogoče bi začela s profesorjem Ihanom. Pri zdravnikih gre za odločanje o življenju in smrti. Mogoče malo okoliščin za začetek. Trenutno imamo v Sloveniji zelo veliko pomanjkanje zdravnikov. Verjamem, da je tako marsikje, tudi v

---

2 Univerze v Ljubljani



drugih državah Evrope. To se je že zgodilo pred leti. Takrat pred leti je bil problem slovenščina. Takrat je minister rekel, da je slovenščina ozko grlo in je znižal raven znanja za zdravnike, da o tem pač odločijo delodajalci. Večina zdravnikov je takrat šla delat izpit na ravni A2. Tak je bil zakon. No, leta sedemnajst pa se je zakon spremenil. Od takrat naprej morajo zdravniki, če želijo delati v Sloveniji, opravljati izpit na ravni C1.

Profesor Ihan je strokovnjak, je sam delodajalec, zelo dobro pozna zadevo, dober pregled ima nad situacijo. Zato vas prosim, če mogoče malo predstavite vse skupaj. Kakšni so tokratni ukrepi, od kod bo Slovenija dobila zdravnike, kako bomo z njihovo slovenščino? Če lahko na kratko predstavite malo situacijo na tem področju.

**Alojz Ihan:** Ja. Kot ste omenili, pojav pomanjkanja zdravnikov je evropski, tudi svetovni, ampak vsekakor je to evropski pojav. Primarno je nastal zaradi tega, ker še pred desetimi, petnajstimi leti nismo niti slutili, kako zahtevni bodo postali pacienti glede na vso tehnologijo, ki je omogočena. Se pravi še dvajset let nazaj ni bilo običajno, da si vsak človek zamenja očesne leče, da si vsak človek slej ko prej zamenja kolena, kolke [smeh]. Ko je človek dobil raka, je bilo pet let že kar uspeh. Danes ima pravzaprav večina ljudi, ki dobi raka, pred sabo petnajst, dvajset let zelo kompleksnih posegov. Ne umrejo. Tudi avtoimunske bolezni zahtevajo ogromno tehnologije. Zaradi tega je po celi Evropi veliko pomanjkanje zdravnikov. V Evropi opažamo migracije zdravnikov z vzhoda, migracije zdravnikov od tam, kjer je investicija, kjer je denarja na bolnika manj, v države, kjer je količina denarja na bolnika višja. Obstaja en pas evropskih držav, ki se vleče od Skandinavije preko Nemčije, kjer je potrošnja za zdravstvo na prebivalca nad štiri tisoč evrov. Potem je drugi pas, v katerega sodimo tudi mi, kjer je potrošnja nekako dva do tri tisoč evrov. In je potem tretji pas, kjer je potrošnja nekako pod tisoč petsto evrov letno.

Opažamo migracije: iz tretjega pasu migrirajo zdravniki, medicinske sestre v drugi pas. Zakaj ne v prvega? Zaradi tega, ker je drugi pas kulturno bližji jugovzhodu kot pa prvi pas, se pravi, težje je iti večini sester in zdravnikov npr. iz Srbije ali pa Bolgarije ali pa Bosne neposredno v Skandinavijo. Lažje jim je iti v Slovenijo. Čeprav se tudi to dogaja, ampak ne toliko. Te migracije so dejstvo. Vsi se selijo za eno ali pa dve državi bolj proti severu in pa proti zahodu. V Sloveniji imamo isto: migracije iz Slovenije. Okrog 20 odstotkov

naših diplomantov želi delati preko meje, v Avstriji, čeprav je pri nas ogromno pomanjkanje zdravnikov. Mi bi takoj potrebovali zdravnike za tri ali pa štiri letnike na vseh medicinskih fakultetah. Najboljši način je, da pridejo k nam iz držav nekdanje Jugoslavije zato, ker smo kulturno dosti blizu. Oni nas še nekako razumejo in mi njih. Tudi zainteresirani so. Kadar delamo poizvedovanje npr. v Srbiji, je ogromno interesa, da bi prišli v Slovenijo. Seveda, uvoziti zdravnike je za državo velik plus, ker ni treba investirati dvesto, tristo tisoč evrov za vse izobraževanje. Zaradi tega si prizadevamo, da bi nekako omogočili uvoz zdravnikov, ampak hkrati mora pa to biti varno.

Če mi zahtevamo, da pride sem zdravnik, ki je že jezikovno popolnoma suveren, potem bomo predpostavljali, da bo zdravnik pustil službo v Beogradu in bil potem tukaj eno leto ali pa dve, živel od zraka in se potem tukaj zaposlil. To ni realno. Zaradi tega želimo narediti nek sistem, in ga že delamo, da jezikovni tečaj ne bi bil tako splošen, da bi bil jezikovni tečaj zdravniški. V okviru fakultete imamo programe komunikacije s pacientom. V teh programih, ki jih imajo naši študentje, sodelujejo igralci. Igralci se naučijo določeno bolezen. Potem pride študent k njemu, ga sprašuje, igralec mu pove vse o tisti bolezni, potem dobi dokumentacijo in mora napisati, kaj je s pacientom, kaj kaže dokumentacija, kako ga bo zdravil, v katere diagnostike ga bo poslal. Te module, te programe že imamo za naše študente. Zato smo rekli: zakaj bi se na jezikovnem izpitu moral človek pogovarjati o plezanju po hribih ali pa o igranju hokeja? Sploh če pride iz Makedonije, se bo zelo težko pogovarjal o hokeju, v Sloveniji pa vsak ve vse o hokeju [smeh]. Zaradi tega se nam je zdelo pretirano, da bi zahtevali od tistih, ki se učijo jezik, da so suvereni na vseh kulturnih področjih, ki so nam v Sloveniji samoumevni, medtem ko prišlekom pa niso samoumevni. Zato se nam je zdelo, da je za zdravnika čisto dovolj, da je suveren v tem modulu, ki sem ga opisal. Se pravi, da dobi igralca, da se mora znati z njim pogovarjati o njegovih težavah, da mora znati potem pregledati dokumentacijo v slovenščini, da mora znati napisati, kako bo potekalo zdravljenje, in da to predoči komisiji, ki bo ocenila, če je na medicinskem področju dovolj suveren, da je varen za pacienta. To bi radi uvedli na zdravniški zbornici. Ne bi nižali nivoja izpita, ampak bi za začetek zahtevali ta zakonsko predpisan nivo izpita v območju komunikacije s pacientom. Sicer je iz številnih jezikovnih študij znano, da je nepopolno obvladovanje jezika v nekaterih strokah

lahko nevarno. Manj nevarno je npr. v splošni medicini. Če zdravnik svojega pacienta pozna in ta redno hodi k njemu, ve za njegove bolezni, potem se oba nekako prilagodita v tej komunikaciji. Imeli smo že izkušnje s splošnimi zdravniki, tako kot je kolegica povedala, v tistem obdobju, ko zakon ni zahteval znanja slovenščine. Takrat ljudje niso bilo tako zelo nezadovoljni s splošnimi zdravniki. Rekli so: »Hvala bogu, vsaj zdravnika imamo. Dajmo primerjati s stanjem, ko ga nimamo.« Če nas ta zdravnik pozna in imamo z njim kontinuirana srečanja, nas ta zdravnik toliko razume in mi njega, da ne bo kakšnih velikih težav.

Drugo pa je recimo na urgenci. Na urgenci je vso komunikacijo od pacienta potrebno dobiti v petih, šestih minutah. Potem sledi ukrepanje, ki pa je lahko življenjsko ogrožajoče, če je napačno. Na urgenci pa mora tisti, ki se pogovarja s pacienti, zelo dobro vedeti, kaj pomenijo tudi medvrstični izrazi. Ljudje imamo za to, da smo pijani ali pa da smo zadrogirani ali pa da smo se z nekom stepli, velikokrat zelo različne izraze. Recimo nekdo bo rekel: »Malo je čuden«. To lahko v nekem kraju pomeni, da se je napil, in tisti, ki sprašuje tega pacienta in tega ne ve, lahko naredi hude napake. Lahko nekoga, ki je povedal, da mu je slabo, razume, da se je napil. Pa ne bo vedel, da se je on pravzaprav udaril v glavo in da bi ga moral poslati na slikanje. Na urgenci je 90 odstotkov ljudi, ki so poškodovani, ki morajo samo malo počakati. Problem je pa ujeti tistih 10 odstotkov ljudi, ki jih je treba dati naprej.

**Ina Ferbežar:** Če lahko samo na kratko povzamem. Vi predlagate za različne specializacije malo različne nivoje znanja in dejansko uporabo jezikovnega profila, to, čemur je nenazadnje namenjen Skupni jezikovni okvir<sup>3</sup>. To je ena zadeva. Mogoče pa tole dopolnitev. Zakaj izpit za zdravnike? Ko je izšel zakon, da morajo zdravniki opravljati izpit na C1, so zdravniki prišli in je bil zelo, zelo slab uspeh. Komaj 20 procentov zdravnikov je opravilo ta izpit, ker seveda pridejo in zelo na hitro hočejo na izpit. Kar profesor tudi predlaga, če sem vas prav razumela, je nek specializiran tečaj.

**Alojz Ihan:** Ja. Idealno bi bilo, da bi se delodajalec zavezal, da bo zdravnik imel eno leto takšnega statusa, da bo s pacienti delal samo pod mentorstvom

---

3 *Skupni evropski jezikovni okvir: učenje, poučevanje, ocenjevanje – SEJO*, prevod dokumenta Sveta Evrope o jezikovnem poučevanju naspluh: <https://centerslo.si/wp-content/uploads/2015/10/SEJO-komplet-za-splet.pdf>

in da bi mu v tem času obvezno omogočal in tudi plačeval tečaj, enoletni tečaj slovenskega jezika, v katerem bi bila skupina tečajnikov samo zdravnikov.

**Ina Ferbežar:** Homogena (skupina).

**Alojz Ihan:** Samo tisti kandidati za zdravniški izpit naj bi bili skupaj na tečaju, zato da bi vadili predvsem komunikacijo s pacientom, ki jo bodo potem tudi na izpitu imeli. To je nekako želja.

**Ina Ferbežar:** Hvala lepa. Tukaj smo zdaj pri dostopu do znanja. Seveda, idealno bi bilo, če bi zato poskrbel delodajalec. Kar nekaj dostopa do znanja ponuja država. In bi se tukaj preselili mogoče h gospodu Cetini, ki je predstavnik Ministrstva za notranje zadeve. Moram najprej povedati, da Center za slovenščino<sup>4</sup> zelo dobro sodeluje z notranjim ministrstvom. Jaz mislim, da nam kolegi iz tujine lahko včasih tudi zavidate, da je tako. Dostikrat nas pokličejo in nas vprašajo za mnenje. Včasih pa tudi ne.

**Igor Cetina:** In obratno. [smeh]

**Ina Ferbežar:** In obratno, seveda. Moram seveda povedati okoliščine. Slovenija ponuja dostop do znanja migrantom. Razmeroma pozno smo začeli z brezplačnimi tečaji, in sicer se program imenuje Začetna integracija priseljencev.<sup>5</sup> To je 180-urni tečaj, ki se ga udeležujejo državljani tretjih držav. In gospod Cetina je strokovnjak za to. Mogoče za začetek: Kakšni so trenди, kakšno je trenutno stanje, koliko ljudi se uči, se vključuje v te programe začetne integracije?

**Igor Cetina:** Če dovolite, ne bi se osredotočil samo na program Začetne integracije priseljencev. 18. septembra je minilo deset let, odkar je Ministrstvo za notranje zadeve podpisalo prvo pogodbo za takratni tečaj slovenskega jezika. Začetna integracija priseljencev je program, v katerem se izvaja tečaj od leta 2012 naprej. Mi gledamo na to kot na kontinuiteto izvajanja, ne pa neki hud presek. Jaz imam tukaj par števil in če imate čas in voljo, vam jih lahko malo predstavim. V grobem: številke zajemajo področje od septembra 2009 do 31. oktobra letos<sup>6</sup>. V tem času je z obiskovanjem tečajev, ki jih organiziramo mi,

---

4 <https://centerslo.si/>

5 Začetna integracija priseljencev, krajše ZIP.

6 2019

začelo 12 810 ljudi. Največ jih je v starostni skupini od 26 do 35 let in največ je žensk. Med 12 810 je kar 7885 žensk, medtem ko je slabih 5000 moških. Če pogledamo države: tečajniki so prihajali iz preko stotih držav. Iz nekaterih je samo en tečajnik, tako da se bom osredotočil samo na tistih glavnih pet držav, v katerih imamo trimestna števila. Največ jih je, seveda, iz Bosne in Hercegovine, skupaj 6221. Preostale države so tudi v prvi vrsti s področja bivše Jugoslavije. Črna Gora 116, Hrvaška – čeprav Hrvati od leta 2015 niso več upravičeni do teh tečajev – 179, 101 Kitajec, 1240 oseb s Kosova, 770 iz Ruske federacije, 966 iz Severne Makedonije, 1304 iz Srbije ter 577 državljanov iz Ukrajine. To so samo tisti, kjer imamo trimestne številke, vse ostalo so statistično manjše številke. To je nekako to.

Kakšni so trendi v zadnjem obdobju? Niti ne spremljamo toliko po državljanstvih, gledamo bolj celostno sliko. Za tiste, ki vas zanima bolj splošno področje migracij, sem pustil naših nekaj statističnih poročil zunaj in si lahko pogledate.

**Ina Ferbežar:** Zelo veliko ljudi med temi, ki se udeležujejo teh programov, opravlja izpit iz znanja slovenščine. Ministrstvo za notranje zadeve uporablja izpite iz znanja slovenščine, ki jih izvajamo na Centru za slovenščino. To niso izpiti, ki so narejeni posebej za ta namen. To so splošni izpiti. Tukaj prihaja pri nas malo do kratkega stika, ker recimo ljudje, ki se udeležujejo Začetne integracije priseljencev, 180-urni tečaj, lahko gredo na izpit A2, ki pa je delan za popolnoma druge namene, z drugimi cilji in veliko več ur se predvideva, da prideš do stopnje A2, in sicer od 300 do 450 ur. Moje vprašanje tukaj bi bilo: zakaj ste se odločili za to kombinacijo, če gre za različne cilje, za različno število ur?

**Igor Cetina:** No, pa smo spet pri komunikaciji. Prvič, ker Center za slovenščino takrat ni obvestil, da bo spremenil svoje izpite, ko je leta 2016 šel na višji nivo.

**Ina Ferbežar:** Popolnoma enak izpit je ostal.

**Igor Cetina:** Glavni razlog je sama narava programa Začetna integracija priseljencev. Poleg jezikovnih vsebin ta vsebuje tudi vsebine iz poznavanja slovenske zgodovine, družbe in ustavne ureditve. Ta je bila na začetku mišljena kot ločen program, vendar je bila izredno slabo obiskana, tako da smo ga zdaj združili s tečajem slovenščine. Udeleženci lahko spoznavajo vse vsebine.

Program predvideva tudi spoznavanje lokalnega okolja, tako da se ljudje lažje vpnejo tudi v lokalno okolje. Priznam, da ne poznam toliko programa, o katerem ste govorili. Se mu nisem toliko posvečal, tako da ne vem, ali tudi ta to vsebuje. Ampak ZIP pač to vsebuje in ni osredotočen izključno na slovenščino.

**Ina Ferbežar:** Saj zato sprašujem.

**Igor Cetina:** Če dovolite, nadaljujem. Zdaj bom zelo suhoparno povedal, ker se vsi se zavedamo, da to ni res. Ampak mi gledamo na to, da nekdo dobi državljanstvo, kot na zaključek integracije. To je zelo formalno gledano. In mi pač ponujamo možnost. Brezplačen tečaj kot vstop v to možnost in izpit kot zadnji korak, zato da lahko to oseba izkoristi, zato da lahko pridobi dokument, ki je eden od pogojev za pridobitev državljanstva. Nikjer ne piše, da morajo iti takoj, ko opravijo tečaj, tudi na izpit. Važno je samo, da so 80 procentov prisotni na tečaju.

**Ina Ferbežar:** Zakaj tako vztrajam pri tem vprašanju. Namreč, malo paradoksalna situacija je nastala: ljudje, ki so prosilci za slovensko državljanstvo, so že dolgo časa v Sloveniji.

**Igor Cetina:** Vsaj pet let.

**Ina Ferbežar:** Potem se udeležijo tečaja začetne integracije priseljencev, ki je namenjena res tistim, ki začnejo. Poleg tega pa ta tečaj, ta program, ne predvideva nobenega izpita ravno zato, ker vsebine, kot so ustavna ureditev, poznavanje družbe – tega ne preverjamo, ker o tem lahko govorimo šele na nivoju B2 – to so abstraktne zadeve. To je problem. Zato, malo moram izzivati, saj veste. Bomo nadaljevali to debato. Vsekakor ti ljudje gredo na izpit. Tukaj mogoče nadaljujem pri vas, doktorica Zdravković.

Mimogrede, ko sva se z gospo Zdravković prvič srečali, mi je rekla: »Midve se že pozna. Pri vas sem delala izpit.« Sem rekla: »O, moj bog. A je bilo hudo?« Je rekla: »Ne! Super izkušnja!« Tako da, hvala, Lana, za to »super izkušnja«.

Moje vprašanje bi bilo, v zvezi s priseljenci, ki morajo na izpit. Vam namreč ni bilo ravno treba, za vas je bila to dobra izkušnja, za večino pa ni. Za večino teh kandidatov, o katerih je pripovedoval gospod Cetina, je to argument moči. Argument moči, ki ga naša država izvaja. In me zanima, od česa je odvisno, ali bo izkušnja bolj negativna ali bolj pozitivna, če se udeležiš izpita? In kako vi kot

filozofinja gledate na to: a ima izpit tudi moč argumenta, a je vedno samo nek pokazatelj moči, kontrole in tako naprej? Ti dve vprašanji, če lahko odgovorite.

**Lana Zdravkovič:** Dejstvo je, da se vsaka nacionalna država vzpostavi, med drugim tudi s homogenizacijo, unifikacijo in kodifikacijo nacionalnega jezika. Vstop v ta jezik predstavlja del nekega širšega pogleda integriranja v družbo. Ti procesi učenja jezika, jezikovnih tečajev so potemtakem del nekega širšega sindroma integracije v neko družbo, ki predpostavlja tudi sprejemanje nekatereh pravil, odgovornosti in tudi poznavanje nekega širšega konteksta države, v katero se preseliš in katere del na nek način postaneš.

Ti vstopi so lahko bolj ali manj boleči, bolj ali manj travmatični. Odvisno seveda od tega, iz kakšnih okoliščin določena oseba prihaja v določeno državo in tudi, nenazadnje, v kakšno situacijo prihaja. Seveda je tukaj zelo pomembno razredno vprašanje oziroma ekonomska situacija osebe, in s kakšno situacijo se sooča, ko pride v to državo. Ali jo že čaka neka služba, kot je prej bilo povedano, neki zdravnik pride, tukaj ga čaka vse. Ali pa pride nekdo kot, agambenovsko povedano, *homo sacer*, neko golo človeško bitje. To so zdaj begunci, s katerimi se mi soočamo. Danes je ravno eden umrl na meji. Seveda, ne vem, če veste, Slovenija se ponaša z žico na svoji zunanji meji, tako da je to tudi ena lepa pridobitev, ki jo imamo.

Če nekdo pride v takšnih okoliščinah, ko nima ničesar in ga nič tukaj ne čaka oziroma zelo malo, potem je seveda ta vstop zelo travmatičen. Jaz delam z migranti, z begunci, prosilci za azil. Pri svojem delu se soočam z deljenjem izkušenj z njimi. Tudi iz njihovih pripovedovanj, pa tudi meni se zdi tako zelo logično, da je teh 180 ur, ki je zapovedanih, zelo malo.

**Igor Cetina:** Ja, ampak mislim, da imajo te osebe ...

**Lana Zdravkovič:** Imajo 320 ur potem, ampak ni dovolj, pravijo, tako da je mogoče treba razmisliti o tem. Mi smo delali tudi par projektov na Mirovnem inštitutu,<sup>7</sup> kjer smo to raziskovali. Ugotavljamo, da to ni dovolj za osebo, ki pride iz nekega totalno drugega jezikovnega področja. Ne bi bilo slabo, da je mogoče teh ur več. Je pa fascinantno potem, ko spoznaš na primer enega Afganistanca ali pa Sirca, ki se je v šestih mesecih naučil slovenščine. To je res

---

7 <https://www.mirovni-institut.si/>

fascinantno. Jaz sem iz Srbije prišla, iz Beograda, in vem, kako je. Že meni se je bilo težko slovenščine naučiti, pa si mislim, kako je šele enemu, ki je z nekega čisto drugega jezikovnega področja.

Tudi preverjanja znanja seveda predstavljajo neko moč države po eni strani. Po drugi strani pa tudi neki čisto pragmatični korak k temu, da se oseba vključi v neki dominanten jezik, kar je, zelo pragmatično, tudi potrebno. Jaz mislim, da bi bilo potrebnih več ur in bolj sistematično vključevanje v to. Tudi pri svojem delu in pri delu s prijatelji migranti, s katerimi jaz sodelujem, vidim, kako je tudi njim pomembno, da se slovenščine čim prej naučijo, ker je od tega odvisno tudi njihovo vključevanje na trg dela. Oni so izobraženi za določene poklice, ampak se ne morejo vključiti, dokler slovenščine ne znajo dobro, tako da je to absolutno zelo pomembna stvar.

Verjamem pa, da ti jezikovni izpiti, na splošno neka preverjanja znanja travmatizirani osebi in osebi iz nekega čisto drugačnega kulturnega okolja predstavljajo velik problem. Jaz recimo vem, da ko sva midva z mojim možem delala takrat izpit – to sva midve preverjali zadnjič – je on znal, on zna slovensko bolje kot jaz, tudi govori slovensko bolje kot jaz. Jaz imam ta balkanski naglas, ki se mu [smeh] iz nekaterih svojih osebnih razlogov nočem odreči. On govori zelo dobro, ampak je bil ocenjen slabše kot jaz. On je očitno imel neko tremo, ki je jaz očitno nisem imela. Definitivno ni nujno, da izpit pokaže realno stanje. Mogoče bi tudi preverjanja lahko zasnovali kako drugače, če bi že želeli oceniti neko realno stanje.

Še nekaj sem hotela reči, glede kulturnih razlik. Jaz imam zdaj sodelavca iz Sirije, ki dela z nami na Mirovnem inštitutu, ki smo ga zaposlili kot kulturnega mediatorja. On stalno izpostavlja te kulturne razlike, v smislu, kako imamo neke druge prioritete v zahodnem svetu. Recimo glede zamujanja, glede nekih natančnosti. Jaz bom stalno izpostavljala, da ne sme zamujati, da ko smo zmenjeni, da mora priti pravočasno. In on pravi: »Ne vem, ne razumem sploh tega koncepta. Pomembno je, da se mi imamo dobro [smeh], da se jaz spočijem, in če zamudim, to ni tak problem.« To so zanimive stvari, ko se soočaš s takimi totalno različnimi svetovi.

Bi pa povedala glede kontinuitete, da je Slovenija migracijsko strategijo sprejela šele letos. Ker je integracija samo en del, eden od šestih delov. Tako da ...



**Igor Cetina:** Neki dokumenti so že iz leta 2008.

**Lana Zdravković:** Dokumenti so, nek celovit dokument glede integracije pa je sprejet letos.

**Igor Cetina:** Hotel bi samo eno stvar povedati. Ko smo prej govorili o izobraževanju in o testih in tako naprej. Jaz bi tukaj poudaril, da so naši programi vsi mišljeni za starejše od 15 let. 15 let je meja za obvezno osnovno šolo v Sloveniji. Dosti je vezano tudi na specifiko andragoškega izobraževanja. Pogosto gre za slabše izobražene ljudi. Slabše izobražene – zdaj ne govorim o neizobraženih. Oni imajo srednjo šolo, poklicno šolo ali pa kaj takega. Zadnjič so se šolali leta ali pa desetletja, preden so prišli sem, tako da se morajo v okviru teh 180 ur, ki jih imajo na voljo, še enkrat začeti naučiti učiti. Pogosto je, glede na to, da so tako učitelji kot tudi mi, osredotočeni na vsebino, zelo velik izziv. Po mojem se to tudi pozna pri preverjanju, ker imajo tudi malo strahospoštovanja do avtoritete.

**Ina Ferbežar:** Še nekaj, če vas lahko dopolnim. Ja, to je res, mi opažamo, sploh pri malo nižje izobraženih, pri zdravnikih to ni problem, ker točno vedo, da gre za dva popolnoma drugačna jezika. Zaradi zgodovinskih okoliščin, o katerih je Marko<sup>8</sup> že prej pripovedoval na svojem plenarnem predavanju in Petra<sup>9</sup> je malo omenjala. Vajeni smo bili, da se nekako razumemo v isti državi, zdaj pa ni več tako. Dostikrat se tudi zgodi, da se ljudje, ko pridejo na izpit, niti ne zavedajo, da ne govorijo zares slovensko, ker v resnici oni ne govorijo več niti svojega jezika. Ampak se pač ne zavedajo, niti dostikrat nimajo dostopa do informacij. Ne vejo, da bi se pravzaprav lahko učili. Tukaj se mi zdi, da je potem tudi vprašanje motivacije. Dostikrat je, seveda, izgovor: »Tukaj sem zaradi dela.« Osebno velikokrat slišim: »Tukaj sem zaradi dela, ne zato, da bom hodil v šolo.« Zato tudi več hodijo ženske. To je pač dejstvo, s katerim se mi dnevno soočamo tudi na Centru za slovenščino. Samo to sem hotela dopolniti. Kar izvolite.

**Igor Cetina:** Še eno stvar bi dodal. Če govorimo o južnoslovanskih govoricah, predvsem s področja bivše Jugoslavije. To so edine skupine v Sloveniji,

---

8 Marko Stabej, ki je na posvetu sodeloval s prispevkom *Sojezičje: utopija ali realnost?*

9 Petra Likar Stanovnik, ki je na posvetu sodelovala s prispevkom *Kako predsodki vplivajo na ocenjevanje (na primeru slovenščine)*.

za katere lahko trdimo, da imajo tukaj neko skupnost, bolj ali manj trajno. Pogosto je tukaj že druga, tretja generacija. Številni med njimi ne vedo, da ne govorijo slovensko, tudi zaradi tega, ker se jih preprosto razume. Njihovi otroci, njihovi vnuki, ljudje, ki so doma odraščali ob južnoslovanskih jezikih, zdaj delajo v trgovini, delajo na pošti, delajo v uradih. Pogosto lahko s svojo mešanico opravijo praktično večino zadev. Ne morejo pa preiti preko formalnega preverjanja znanja.

**Lana Zdravković:** Jaz sem se hotela navezati. Hotela sem izpostaviti, kakšna je percepcija osebe, ki ne govori dobro slovensko glede na poreklo te osebe. Če imamo osebo z zahoda, visoko izobraženo, belo, ki govori recimo angleško, francosko, to ni tak problem. Jaz poznam tudi ljudi, umetnike ali pa visoko izobražene ljudi, znanstvenike, ki tukaj delajo in sploh ne govorijo slovensko, ki so tukaj po deset let in funkcionirajo na ravni angleščine. To ni tak problem, ker to so kulturni, fini, izobraženi ljudje. Potem pa imamo te, ki prihajajo s tako imenovanega globalnega juga, ki so pa »umazani«, to so begunci, to so tujci, ki so rangirani drugače. Jih nočemo, ker prihajajo sem »vzet naše ženske, naše službe«. Percepcija je čisto drugačna. Jaz temu rečem sindrom »Kako dobro govorite slovensko«. Ko sem se jaz preselila sem in smo se pogovarjali s prijatelji in znanci, sem jaz vedno to doživljala kot eno kulturno klofuto: »O, kako dobro govorite slovensko, glede na to, da niste od tukaj.« To se na ta način začne. Če mi slišimo nekoga govoriti albansko ali pa makedonsko ali pa arabsko ... Jaz sem se enkrat z enim Albancem pogovarjala. Je rekel: »Mi govorimo naš jezik, pa vidim, da je arabski jezik podoben. Mi, ko govorimo, vedno izgleda, da se kregamo med sabo [smeh]. Malo agresivno vse to skupaj izgleda.« Ko mi potem to slišimo na ulici, je to čisto neka druga percepcija te osebe, tega tujca, kot pa percepcija tujca, ki prihaja z zahoda in govori lepo francosko ali pa italijansko in nam je to vse tako malo romantično. Tega se je tudi treba zavedati: kako percipiramo ljudi s tako imenovanega globalnega juga in z »razvitega zahoda«. Tudi osebe, ki prihajajo iz višjih razredov, ki so izobražene, in tiste, ki so pač gola človeška bitja in nimajo nič drugega razen tega, da so pač ljudje.

**Ina Ferbežar:** Hvala lepa. Težko temo smo odprli. Se bomo še vrnili. Tukaj smo imeli tri govorce, kjer smo se koncentrirali na enojezičnost, kjer je v središču slovenščina, izpiti so iz slovenščine, so enojezični. Zdaj bi pa prešli

malo drugam. Mi živimo v večjezični realnosti. To je bilo danes že velikokrat izrečeno. Smo raznojezični, smo plurilingvalni. To je pač realnost. Nekateri delodajalci živijo tako resnično. Na primer gospa Zupančič, ki prihaja iz podjetja Revoz, ki je eden večjih zaposlovalcev v Sloveniji. Nekih, ste rekli, tri, štiri tisoč ljudi je zaposlenih pri vas. To podjetje živi večjezičnost in raznojezičnost. Ali lahko predstavite, kako to izgleda v praksi v vašem podjetju, koliko jezikov imate, a ste jih kdaj prešteli v vašem podjetju? Prosim.

**Tina Zupančič:** Seveda imamo veliko, veliko jezikov, ker smo mednarodno podjetje. Sodelujemo tudi z ostalimi tovarnami, ker smo le ena izmed štiridesetih Renaultovih tovarn po svetu. V tovarni to izgleda tako, da večinoma seveda govorimo slovensko. Je pa zelo prisotna tudi francoščina, včasih je bila veliko bolj. Zadnje čase opažamo, da postaja čedalje bolj prisotna tudi angleščina, predvsem na sestankih. Ogromno sestankov imamo z drugimi tovarnami po Skypu in tukaj je francoščino že zamenjala angleščina. Je pa tako: ko imamo projekt, pridejo na srečanje inženirjev sodelavci iz Romunije, naš direktor je Turek, ki govori tekoče angleško, francosko, slovensko se trudi s posameznimi besedami, si zapisuje zanimive izraze. Mi smo veseli, seveda. Tujcem, ki pridejo v tovarno, ponudimo možnost učenja slovenščine. To se mi zdi zelo pomembno. Vsi ti sodelavci pravijo, da mogoče ne toliko zaradi življenja v tovarni in komuniciranja s sodelavci, predvsem da se lažje vključujejo v družbo, v okolje. To bi bilo mogoče za začetek.

**Ina Ferbežar:** Ko sva se pogovarjali, ste omenili, da imate vi neko prakso, da imate različne napise v različnih jezikih v menzi, povsod.

**Tina Zupančič:** Ja.

**Ina Ferbežar:** Kako to funkcionira?

**Tina Zupančič:** Menije imamo recimo v slovenščini, angleščini, francoščini. Imamo tudi službo za prevajanje, ki pomaga pri jezikovnih zaprekah. Ko imamo recimo izobraževanja v angleščini, se s službo za prevajanje dogovorimo, da pridejo prevajat v slovenščino. Potem napisi, če pogledamo računalnik, imamo Renaultovo računalniško okolje, kjer imamo možnost izbire vseh jezikov, ki jih ponuja Renault Renaultovim tovarnam. Zdaj so nam povezali katalog izobraževanj, tako da si lahko nastaviš jezik in izbereš katalog izobraževanj v svojem jeziku. To se mi zdi tudi v redu. Sámó izobraževanje je

potem v slovenščini. Ravno zaradi tega tudi spodbujamo učenje slovenščine. Varstvo pri delu na primer ponudimo v jeziku tega sodelavca, ampak naprej pa vidijo, da se morajo tudi oni naučiti slovensko, da se je za lažje vključevanje v skupnost in okolje dobro naučiti slovensko.

**Ina Ferbežar:** Se pravi, vi jim organizirate tudi tečaje slovenščine?

**Tina Zupančič:** Tudi.

**Ina Ferbežar:** Kako motivirani pa so? Koliko časa vztrajajo? Ker dostikrat je pri teh tečajih tako, da pridejo, potem pa če ni ravno obvezno ... Ne tako kot recimo pri ZIP-ih, kjer imajo vendarle neko boniteto in vsaj na izpit grejo lahko zastonj. Koliko časa potem ...?

**Tina Zupančič:** To je pa čisto odvisno. Kar se tiče tečajev slovenščine, smo nekako razdelili tečaje slovenščine glede na materni jezik sodelavcev. Pred dvema letoma smo se soočili s situacijo pomanjkanja delovne sile in je k nam prišlo veliko sodelavcev iz Bosne, iz Hrvaške. Takrat smo bili v nekem trenutku soočeni s situacijo, da je bilo potrebno organizirati tečaje slovenščine. Ponuditi smo jim želeli možnost teh tečajev in so jih opravljali pred službo, nismo jih mogli ponuditi znotraj delovnega časa. Jaz se trudim, da bi bilo med delovnim časom, če se le da. Tako imamo francoščino, nekaj slovenščine je tudi med delovnim časom, ampak za spremljevalne službe. To so bili pa delavci večinoma iz proizvodnje in je bilo moteče za proizvodni proces. Ne bi bilo možno izpeljati, da bi bilo med delovnim časom, in smo jih povabili izven delovnega časa. Je pa res, da je bilo na začetku navdušenje veliko, na koncu pa so večinoma ostale ženske. [smeh] Tako kot je bilo povedano. Je pa glavna motivacija za učenje, se mi zdi, za osnovni nivo, namenjen za nižja delovna mesta, kjer ni obvezno znanje slovenščine. Je pa obvezno, čim grejo na višja delovna mesta, kjer imajo več komunikacije z ostalimi sodelavci, tudi strokovne, tem delavcem smo pa kar v program poleg vseh izobraževanj dali tudi slovenski jezik.

**Ina Ferbežar:** Na izpit jim pa ni treba?

**Tina Zupančič:** Imajo ob koncu tečaja izpit. Ni pa, da bi bilo to na uradni ravni. Ker tudi pri drugih jezikih – angleščini, francoščini – ne zahtevamo tega.

**Ina Ferbežar:** Vi vidite pravzaprav dodano vrednost v tej svoji večjezičnosti. Ali bi imeli kakšne smernice za kakšna podjetja, recimo v turizmu? Tam zlepa ne vidimo, da bi se učili slovenščine. Mi dobivamo informacije, da morajo turistični delavci delati izpit na A1, ki pa trenutno še ne eksistira v Sloveniji. Ali imate kakšen nasvet?

**Tina Zupančič:** Mi, kar smo naredili, smo se povezali v okolju in izkoristili razne programe, ki jih ponuja ministrstvo, tudi z jezikovnimi šolami smo sodelovali. Predvsem se mi zdi, da mora biti vsebinsko usmerjeno, ne pa v celoti. To potem lahko delaš na višjem nivoju, osnove pa morajo biti. Se mi zdi, da neko osnovno, splošno sporazumevanje mora biti, potem pa lahko greš na strokovni nivo.

Meni je bila zelo všeč naslednja ideja: ko je pred nekaj leti prišel novi vodja, je rekel, da so njemu ponudili učenje slovenščine še pred nastopom funkcije v Sloveniji. Z ženo sta šla na tečaj zato, da so ju pripravili. Mi smo potem samo nadgradili njegovo znanje. Tudi nekateri Francozi, ki pridejo v Slovenijo, imajo že v pogodbi napisano: tečaj slovenščine v Sloveniji. Imajo 60-urni tečaj, jaz ga organiziram in so zelo zadovoljni. Je pa učenje jezika dolgotrajen proces. Ne moremo pričakovati, da se bojo v enem letu, oziroma če pridejo na misijo za tri leta, da se bojo naučili v treh letih tekoče govoriti slovensko. Je pa res – so mi sami povedali – da jim zelo pomaga, ko gre vodja med delavce, da lahko tudi on od njih iz prve roke izve, da ne rabijo imeti vedno prevajalke, da se ne rabi angleščine posluževati, ker tudi vsi francosko ne govorimo.

**Ina Ferbežar:** Se pravi, ta čisto človeška komunikacija.

**Tina Zupančič:** Ja, čisto zaradi tega. Tudi da mogoče pride do realnih informacij. Tudi to je mogoče en faktor. Tako da zelo spodbujamo tudi učenje slovenščine.

**Ina Ferbežar:** Hvala lepa. Zdaj pa pričakujem mogoče en tak bolj splošen pogled. Pa bom profesorja Stabeja prosila. Danes smo že dvakrat slišali, da je za Slovence slovenski jezik že celo zgodovino zelo močno identifikacijsko sredstvo, nekako temelj narodnega obstoja, da se to odraža v zelo osebnem, zelo čustvenem razmerju do slovenščine. To je tudi že gospa Zdravković omenjala. Glede na vse tole povedano, kaj bi rekli, kakšne jezikovne ideologije usmerjajo trenutno jezikovno politiko v Sloveniji? Kako krmariti med to čustvenostjo na

eni strani in med takole realnostjo, ki jo imamo na drugi, kjer je samoumevno, da je več jezikov, da govorimo različne jezike? Kako loviti, kako držati to ravnovesje? Da ne bojo, ne bomo prizadeti tisti, ki smo zelo čustveni, pa tudi tisti ne, ki racionalno grejo v to?

**Marko Stabej:** Hvala. Jaz sem zelo vesel, da smem tole poslušati, in sem vesel tega, kar je bilo tukaj izrečeno. Na tako zgoščen način tako rekoč tako različnih glasov na to temo že zlepa nismo slišali. Jaz niti ne bi o čustvenosti sami po sebi govoril. Sem pa že prej omenil: čustvena reakcija je pogosto povezana z iskrenimi procesi, pogosto pa z nevednostjo. Meni se zdi, da so tukaj sami primeri, kako je edina rešitev po eni strani humanističen odnos do vsega skupaj, po drugi strani pa pragmatično reševanje težav, ne bom rekel težav, ampak ovir za skupni cilj. Pri tem je treba čustva po mojem seveda pustiti na neki način, naj živijo, ampak ne kot glavni argument česarkoli. Moja izkušnja skupaj s tvojo, saj sodelujeva že dolgo časa in tudi z nekaterimi med vami, je, da stvari gredo postopno. Izkaže se, da je za nekoga 180 ur premalo, to že vemo, da je 320 ur premalo. Ampak še pred desetimi leti se je zdelo 50 ur za vsakogar preveč oziroma odveč. Pričakovanje, da se ozaveščenost vseh lahko spremeni tako »on/off, pak, pa smo ozaveščeni«, je irealno. Važno pa je, zato se mi tudi današnji dogodek zdi razmeroma važen (ne vem, če je, ampak je dragocen), da slišiš različne glasove. In ob neki dobronamernosti lahko najdeš rešitve, ki jih prej ni bilo.

Če smem kolega Ihana rahlo parafrazirati. Če smo za migracijo zdravnikov in večjo potrebo med drugim odgovorni tudi vse zahtevnejši pacienti ali pa vse bolj razvjeni pacienti, na nek način enako velja za udeležence izobraževalnih procesov, pa tudi migracij. In je prav tako. Hkrati je pa sistem marsikdaj narejen – spet se opravičujem – po analogiji, tako kot je bila medicina včasih predvsem, da zdravi in ne preprečuje. Če imaš vse sisteme narejene predvsem za zdravljenje, najbrž stežka kakšen preventivni program zaženeš. V Sloveniji sta vsaj dva taka, ki sta zelo uspešna. Eden je Dora za raka na prsah, presejalni, drugo je Svit za ...

**Alojz Ihan:** Za raka črevesja.

**Marko Stabej:** ... za raka črevesja. Jaz si predstavljam, da ti sistemi zelo veliko stanejo, ampak najbrž konec koncev mnogo manj, kot bi bilo dolgotrajno zdravljenje marsikoga, ki na teh testih ne bi pravočasno odregiral.

Pri jeziku smo daleč od takih presejalnih testov zato, ker se nam je dolgo časa zdelo samoumevno, da kdor pride, če hoče, naj se že nauči. Če se noče naučiti, je sam kriv. Najbolje pa, da ne hodimo nikamor in da smo tukaj mi, drugi pa naj bodo drugje. Vsi se tega šele navajamo in redko sedimo skupaj. Tak primer, kot je bil predlog kolega Ihana, neko vmesno leto zdravnika, ki pride, da že dela in se lahko sproti uči. Analogija je tudi na univerzi, vsaj v Ljubljani. Po dolgih letih spraševanja, kako poservisirati ljudi, ki pridejo sem študirat in morajo študirati vsaj večino programa v slovenščini, na kakšen način jih naučiti slovensko. Nerealno je, da jih boš v enem mesecu naučil, čeprav so to inteligentni mladi ljudje, da jih boš naučil na ta način, da ga potem vržeš v seminar ali kamorkoli. Prišlo je spoznanje, da je treba organizirati neko posebno obliko, ki je do takrat ni bilo. Imenuje se Leto plus, kjer lahko deloma študirajo, deloma pa so obvezani, da se učijo slovenščino in se jim ne šteje v siceršnji rok študija, se pravi neka dodatna pravica. In so se začele stvari v zadnjih dveh letih bliskovito urejati. Seveda pa imamo vsi zaradi tega organizacijske, kadrovske, konceptualne težave. Ampak mi smo tukaj zato, da se s temi težavami soočimo in jih rešujemo. Marsikdo, zlasti ki je zunaj realnih okolij, pa misli, da se ukvarja z usmerjanjem jezika in jezikovne situacije, pa strelja kot najslabši kavboj v najslabšem filmu. On bi vse takoj. Ali bom jaz zmagal ali bo kdo drug zmagal, če lahko tako. V tem smislu je teh refleksov še preveč, hkrati so pa ta vprašanja marginalna, čeprav gre za vprašanja življenja in smrti, sprejetosti, nesprejetosti. Tukaj se mi zdi, da smo zelo napredovali, zlasti ker ljudje konkretno začnejo razmišljati. Dober primer, se mi zdi, je Revoz. Na koncu koncev vaša firma mora profitirati. To ne pomeni, kot marsikdo, ki ima stereotipne predstave, ki je bolj domač ljubitelj slovenščine in njenega statusa, misli in je prepričan, da vsi, ki hočejo zaslužiti, itak vse samo v angleščini počnejo. Kdor resno služi, ve, da je to bistveno kompleksnejša zadeva. Ne vem, če sem odgovoril, kako s tem ravnati. Treba je imeti forme, da si poveš. Treba je imeti *self-assessment*, da sam pri sebi znaš razmišljati, kaj bi pa lahko naredil, da bi bilo boljše, kaj je zdaj realno, kaj ni realno, kaj bo realno. Kdaj bom zaprosil za nek dodatni denar, pa ga bom čez tri ali pa pet let dobil. To šele nastaja.

**Ina Ferbežar:** Ja, hvala lepa. To se mi zdi zelo pomembno: kaj je realno. S tem smo se tudi zelo veliko s profesorjem Ihanom ubadali v naši delovni

skupini, kjer pripravljamo ta izpit, o katerem je na začetku pripovedoval. Izpit, ki je zdaj na C1, za zdravnike sploh ni realen. Je pa tudi tako, da marsikateri zdravnik ima sam nerealna pričakovanja najprej do sebe, češ, saj se jaz lahko to hitro naučim, ker prihajajo večinoma iz sosednjih, bližnjih dežel, kjer se govori slovenščini sorodni jezik. Malo podcenijo slovenščino pa malo precenijo sebe. Se mi zdi ta rešitev res dobra. A se vam zdi, profesor Ihan, da bodo delodajalci pripravljani dati čas in denar za to, kot deloma, vsaj deloma v Revozu počnejo? Se vam zdi to realno? Jaz imam malo zlo slutnjo, da se bo pri denarju zelo hitro ustavilo.

**Alojz Ihan:** Trenutno se mi ne zdi realno. [smeh]

**Alojz Ihan:** Realno je, da se ta proces začne v državnih bolnicah. Država ni tako zahteven lastnik, kar se tiče denarne racionalnosti, medtem ko pa ni realno, da bi to financirali delodajalci v privatnem sektorju. Ti upori so že bili in se je potrebno tem uporom izogniti in vzpostaviti sistem, na primer v zdravstvenih domovih, v državni mreži zdravstvenih domov.

Zdaj imamo že številne kraje, ki so brez zdravnika. Tam pa se morajo lokalne skupnosti odločiti. Primarno zdravstvo financirajo lokalne skupnosti. V Sloveniji država financira bolnice, medtem ko te splošne ambulante financirajo občine. Že zdaj se tam, kjer ni zdravnika, morajo občine odločiti, da če hočejo dobiti zdravnika, mu morajo ponuditi stanovanje. Če mu ne ponudijo stanovanja, ne bodo dobile zdravnika. Nek tak tečaj je malenkost v primeri s stanovanjem. Vzorec se da narediti, ampak da bi pa to postalo splošno obvezno, je pa vprašanje.

**Ina Ferbežar:** Bomo poskusili. Imamo še kar nekaj časa, ker smo zamudili 15 minut. Jaz bi predlagala, če bi imel kdorkoli iz publike kakšno vprašanje za moje goste. Kot ste videli, odpade naša popoldanska delavnica, ki naj bi nadaljevala tale pogovor. Če bi bila kakršnokoli vprašanja, jaz mislim, da bodo moji gosti čisto veseli. Če ne, imam pa jaz kar nekaj vprašanj še za njih za rezervo.

Ne? No, dobro.

Potem pa imam eno, pravzaprav za vse. Veliko se govori o tem, da je Slovenija prehodna država, da vse tendira proti severu. To je tudi profesor Ihan čisto na začetku rekel, da se vse seli proti severu. Mi dejansko v Izpitnem centru



slišimo od naših kandidatov za izpite: »Ah, jaz to rabim samo zato, da bom dobil potni list, potem grem pa naprej.« Tudi od zdravnikov smo to že slišali. Se mi pa zdi, da zelo veliko ljudi ostaja, da nimajo kakšnih intenc hoditi daleč. Kakšna je vaša izkušnja, kakšno je vaše mnenje o tem? Smo zares prehodna država, ali smo vendarle ...? Lana, boš ti?

**Lana Zdravković:** Kakor za koga. Meni je zanimiva ena anekdota. Ko se je vzpostavila balkanska ruta in so začeli ljudje iz Sirije, Iraka, Afganistana hoditi čez Slovenijo, je bila glavna skrb nekaterih politikov, da Slovenija ne bi postala žep za migrante. To je pomenilo, da ne bi Avstrija zaprla meje in bi potem oni ostali tukaj. Potem je neka begunka iz enega izmed vlakov izjavila točno to: da nobeden izmed njih ne želi ostati v Sloveniji, da je njim vseeno cilj Nemčija, skandinavske države. Potem je slovenska javnost dobila napad: »Zakaj nočejo ostati? A mi pa nismo dobri?!« [smeh] To je bilo fenomenalno.

Meni se zdi, da so večini tem novim migrantom, ki zdaj prihajajo, cilj druge države. Razlogi so zelo enostavni. Eno je, da tam že imajo vzpostavljene neke socialne kontakte, prijatelje, družino ali kogarkoli že poznajo. To so države, ki imajo migrantsko preteklost že dolgo. Način vključevanja in integracije je vzpostavljen, tudi neformalni predvsem, in te povezave funkcionirajo. Seveda večina ljudi želi iti nekam, kjer je že slišala, da nekaj obstaja, da je nekdo že tam. To so seveda Nemčija, skandinavske države, Italija, Francija. Po drugi strani pa ljudje iz bivše Jugoslavije predvsem zaradi kulturnih povezav nekdanjega skupnega prostora Slovenijo doživljajo kot ciljno državo, velikokrat tudi zaradi skupne preteklosti. Imamo cel val preseljevanja iz republik nekdanje Jugoslavije v sedemdesetih letih prejšnjega stoletja, ekonomske migracije, ko se je začelo. Tudi iz teh razlogov tukaj že obstajajo neke skupnosti, ki so vzpostavljene. Če je za najnovejši val migrantov Slovenija predvsem prehodna država, je za tiste iz našega kulturnega prostora tudi ciljna država.

Bi pa povedala tudi to, da moj sodelavec iz Sirije, ki zdaj dela z nami na inštitutu, on je tudi bil vsepovsod. To so ta »razturena« življenja, ko si šest mesecev tukaj, pa te deportirajo, pa si malo v zaporu in greš nazaj, pa si malo v Italiji, malo v Franciji, malo ne vem kje. On je dobil status v Sloveniji, živi v Sloveniji. Povedal je, da se edino v Sloveniji počuti kot človeško bitje, da se nikjer, niti doma v Siriji, ni počutil tako. Zato je od tam tudi bežal, ker mu ni bilo dobro. Potem je bil vsepovsod v različnih postopkih. Rekel je: »Lana, ti imaš ful

srečo. Dobro si se odločila, da si migrirala v Slovenijo. [smeh] Ker jaz se edino v Sloveniji počutim kot človek.« Meni je bilo to tako: »Mhm.« Ker se tudi jaz v Sloveniji počutim boljše kot pa drugje.

**Ina Ferbežar:** Take stvari si zelo želimo slišati. Še kdo mogoče? Publika?

**Alojz Ihan:** No, mogoče še o zdravnikih. Slovenija je za zdravnike enako prehodna država, kot je za slovenske zdravnike. Približno enak odstotek slovenskih zdravnikov si želi iti drugam, kot je odstotek tistih, ki pridejo v Slovenijo in si želijo iti drugam. To velja tudi za mlade. Moramo pa vedeti, da je tipična pot mladih zdravnikov na primer iz Srbije, Bosne v Slovenijo, da pridejo po dokončani medicinski fakulteti in v Sloveniji zaprosijo že za specializacijo. V Sloveniji je razpisanih veliko več specializacij, kot pa je mest na medicinskih fakultetah. Kar pomeni, da tisti, ki konča medicinsko fakulteto v Beogradu, potem naredi te nostrifikacijske dokumente, ki jih ni težko narediti, ker so fakultete primerljive. Potem naredi izpit. V Sloveniji potem začne s specializacijo, ki jo plača, seveda, in s katero lahko gre enako kot vsak Slovenec karmorkoli v evropski skupnosti. Ampak izšolati se za zdravnika v enem jeziku je tako velika investicija – ker v Sloveniji dobi zdravnik službo takoj, da si večina strok ne želi potem ponovno iti v to jezikovno investicijo v neko drugo okolje. Na koncu je pa treba s pacienti komunicirati. Če se naučiš komunicirati v slovenščini, potem ni enostavno iti v Nemčijo in se začeti učiti spet naslednjih tri, štiri, pet let in jih porabiti za to, da komuniciraš v nemščini. So pa določene stroke, kjer je to vseeno, recimo kirurgom je popolnoma vseeno, ali režejo v ...

**Lana Zdravković:** V slovenščini ali v arabščini. [smeh]

**Alojz Ihan:** ... ali režejo v arabščini ali v slovenščini. Tako da so nekatere stroke, kjer je migracija ogromna. Ampak to je isto pri nogometaših. [smeh]

**Ina Ferbežar:** Potem pa predlagam, da mogoče zaključimo pri nogometaših.

Jaz bi se vsem res najlepše zahvalila, da ste prišli. Mislim, da je res dobro, to, kar je Marko rekel, še enkrat, da se takole srečujemo. Rada bi citirala našo drago profesorico, ki je bila moja mentorica pa tudi Markova mentorica, naša draga mentorica profesorica doktorica Breda Pogorelec, žal pokojna. Vesela bi bila tega dogodka, rekla bi, da je praznik. Njen citat je bil vedno: »O tem se bomo še pogovarjali.« Tako da najlepša hvala mojim gostom, najlepša

hvala vam, čeprav niste nič spraševali. Mogoče boste sedaj pri kosilu. Najlepša hvala mojima dragima tolmačema tamle zadaj, Jana Zidar Forte in Boštjan Zupančič. [aplavz]

Vas toplo vabim na prigrizek za kosilo. Izvolite.

## **KONFERENCA JEZIKOVNE TEHNOLOGIJE IN DIGITALNA HUMANISTIKA 2020**

**Katja MEDEN**

Institut Jožef Stefan

*Meden, K. (2020): Konferenca Jezikovne tehnologije in digitalna humanistika 2020. Slovenščina 2.0, 8(1): 113–119.*

DOI: <https://doi.org/10.4312/slo2.0.2020.1.113-119>

Konferenca Jezikovne tehnologije in digitalna humanistika 2020, ki jo skupaj z Inštitutom za novejšo zgodovino,<sup>1</sup> Centrom za jezikovne vire in tehnologije Univerze v Ljubljani (CJVT UL)<sup>2</sup> ter raziskovalnima infrastrukturama CLARIN-SI<sup>3</sup> in DARIAH-SI<sup>4</sup> organizira Slovensko društvo za jezikovne tehnologije (SDJT),<sup>5</sup> je letos potekala 24. in 25. septembra 2020, že tretjo multidisciplinarno izvedbo konference pa je podprl CLARIN ERIC.<sup>6</sup> Konferenca, ki se lahko pohvali z več kot 20-letno tradicijo delovanja, je leta 2016 v svoj program vključila tudi področje digitalne humanistike in s tem postala pomemben povezovalni člen med omenjenima disciplinama.

Letošnja konferenca bi morala potekati na Inštitutu za novejšo zgodovino v Ljubljani, vendar je bila zaradi pandemije Covid-19 preseljena v virtualno okolje. Digitalna izvedba je predstavljala odmik od tradicionalne organizacije konference, hkrati pa je prinesla tudi nove izzive. Konferenca je tradicionalno potekala dvodnevno, vendar pa se v nasprotju s predlansko ni delila po jezikih, v katerih so bili prispevki predstavljeni, temveč izključno po tematskih sklopih. Prispevki v posameznih sekcijah so bili kot vedno do sedaj razdeljeni na polne in razširjene. Posamezni študentski prispevki so bili v letošnji izvedbi

---

1 <https://www.inz.si/>

2 <https://www.cjvt.si/>

3 <https://www.clarin.si/info/about/>

4 <http://www.dariah.si/>

5 <http://www.sdjt.si/wp/>

6 <https://www.clarin.eu/>

umeščeni znotraj tematsko ustreznih sekcij in ne v ločenem študentskem panelu. Ker se je v letošnjem letu zaradi globalne epidemije konkretno spremenil način dela in znanstvenega komuniciranja, so avtorji posameznih prispevkov svoje predstavitve predhodno posneli, te pa so bile na spletni strani na voljo že pred začetkom konference. To je omogočilo dinamično izvedbo konferenčnih debat, saj so si udeleženci posnetke in zbornik prispevkov ogledali že pred samo konferenco, v okviru panela pa so nato avtorji predstavili le kratek povzetek svojih predavanj, ki jim je sledila razprava.

Konferenco je v sklopu prvega vabljenega predavanja, ki ga je moderiral Filip Dobranić, odprla Sara Tonelli, vodja raziskovalne skupine Digitalna humanistika pri Fondaciji Bruna Kesslerja v Trentu in izredna profesorica na Oddelku za psihologijo in kognitivne znanosti na Univerzi v Trentu. V svojem predavanju z naslovom *Abusive Language Detection: Too much Digital, not enough Humanities?* je predavateljica ustvarila pregled sodobnih pristopov za prepoznavanje nasilnega govora, poudarila vlogo in razumevanje jezika, ki ga uporabljajo različne spletne skupnosti, ter predstavila trenutne raziskave za boljše razumevanje tega fenomena.

Vabljenemu predavanju je sledila prva sekcija na temo jezikovnih in govornih tehnologij, ki jo je povezovala Tanja Samardžić. Med polnimi prispevki lahko izpostavimo predstavitev prispevka Anke Supej, Mateja Ulčarja, Marka Robnika Šikonje in Senje Pollak, v katerem so primerjali spolno pristranost modelov (oziroma njihovih vložitev) z različnimi konfiguracijami in pristopi k računanju analogij. Med predstavitvami razširjenih povzетkov je Darinka Verdonik v imenu soavtorjev predstavila razvoj in delovanje raziskovalne infrastrukture projekta RI-SI CLARIN, v študentskem prispevku pa je Andraž Pelicon predstavil zaznavanje sentimenta v novicah s pomočjo globokih nevronske mreže.

Naslednjo sekcijo, ki je obravnavala jezikovne vire, je moderirala Darinka Verdonik. Vsebovala je predstavitve štirih polnih prispevkov, treh razširjenih povzетkov in študentski prispevek. Simon Krek (s soavtorji) je predstavil najnovejše razvojne dosežke in vsebino učnega korpusa ssj500k,<sup>7</sup> največje in najpogosteje uporabljene odprtokodne zbirke podatkov za obdelavo slovenskega

---

7 <https://www.clarin.si/repository/xmlui/handle/11356/1210>

jezika. Dolores Lemmenmeier-Batinić je izpostavila problem pomanjkanja javno dostopnih virov za srbski jezik, opisala do sedaj ustvarjene korpuse in metode za pretvorbo korpusnih podatkov v standardizirani XML format. Nazadnje je sekcijo s predstavitvijo študentskega prispevka o metodoloških izhodiščih, razvoju in vodilih pri označevanju metaforičnih besed korpusa metafor *KOMET 1.0* sklenila Špela Antloga. Prvi dan konference je zaključil poseben panel projekta *RSDO – Razvoj slovenščine v digitalnem okolju*, v okviru katerega so vodje posameznih delovnih sklopov predstavili glavne in vmesne cilje ter naloge, ki jih bodo uresničevali v okviru projekta.

Uvod v drugi dan konference, ki je bil obarvan s prepletanjem jezikovnih tehnologij, predvsem korpusno analizo in digitalno humanistiko, je ponudil digitalni zgodovinar Kaspar Beelen z Inštituta Alan Turing, ki raziskuje uporabo strojnega učenja v humanističnih raziskavah. V vabljenem predavanju z naslovom *Speaking on behalf of others: Why the digital humanities should care about parliamentary data* je izpostavil pomen in vlogo parlamentarnih podatkov, ki omogočajo vpogled v jezik ter svetovni nazor poslancev in volivcev, ki naj bi jih ti zastopali, s tem pa zagotavljajo podrobno pričevanje o skoraj vsaki temi, ki je bila kdaj koli predmet javne razprave.

V tretji sekciji, ki jo je moderirala Kristina Štrkalj Despot, so bili predstavljeni prispevki s področja korpusne analize. Kristina Pahor de Maiti je predstavila študijo morfosintaktičnih značilnosti komentarjev na Facebooku, da bi prepoznali tiste, ki so vidni v družbeno nesprejemljivem diskurzu (SUD). Jakob Lenardič in Darja Fišer sta analizirala epistemske modalne prislove v 100-milijonskem korpusu slovenskih doktorskih disertacij (korpus KAS). Študentska prispevka sta predstavljala Zoran Fijavž in Eva Trivunović. Zoran Fijavž je v svojem prispevku raziskoval vpliv videovsebin na prisotnost družbeno neperimerne diskurza, kar je detektiral iz nabora komentarjev na Facebooku, povezanih z LGBT, ki so izvirali iz glavnih virov novic v hrvaščini. Eva Trivunović pa je predstavila (iz)biblijske frazeme, njihove variante, prenovitvene in neprenovitvene modifikacije v korpusih *Gigafida 2.0*, *Janes* in *slWaC*.

Zadnjo sekcijo, v kateri sta se združili digitalna humanistika in pedagogika, je moderiral Miran Hladnik, vsebovala pa je predstavitve treh polnih prispevkov, dveh razširjenih povzetkov ter eno predstavitev študentskega prispevka. Katja Meden in Ana Cvek sta predstavili tehnično in vsebinsko nadgradnjo

Zgodovinskega indeksa citiranosti, citatnega indeksa za sistematično popisovanje citiranosti del na področju zgodovinopisja. Andrej Pančur je predstavil najnovejšo pridobitev med demografsko viktimološkimi raziskavami, digitalno bazo podatkov (vojaških) žrtev prve svetovne vojne z ozemlja današnje Republike Slovenije, ki je nastala na podlagi sodelovanja med različnimi raziskovalnimi in kulturnimi ustanovami ter nekaterimi posamezniki. Nazadnje je Magdalena Schlintl v študentskem prispevku predstavila delo z digitalnimi orodji pri izobraževanju učiteljev na primeru poučevanje-učenje-laboratorij.

Konferenca se je zaključila s podelitvijo nagrade za najboljši študentski prispevek, ki jo je prejel Zoran Fijavž za prispevek *Ambivalence of Queer Visibility in Video-Based Social Media Content*. Ker je bilo med posnetimi predstavitvami kar nekaj izstopajočih, se je programski odbor konference odločil v zaključnem delu podeliti tudi nagrado za najboljšo posneto predstavitev po izboru publike. Nagrado je prejela Špela Antloga za predstavitev prispevka *Korpus metafor KOMET 1.0*.

Ne glede na tehnične in organizacijske izzive, ki sta jih prinesla pandemija in virtualno okolje, je bila Konferenca jezikovne tehnologije in digitalna humanistika 2020 uspešno izpeljana, tako z vsebinskega kot tudi z organizacijskega vidika. Ravno zaradi digitalne izvedbe so na spletni strani konference<sup>8</sup> na voljo ne le zbornik prispevkov, temveč tudi vse predstavitve in posnetki diskusij. Prispevki, predstavljeni na konferenci, so omogočili vpogled v nove metode, aplikacije, nadgradnjo in razvoj raziskovalnih polj. Razprave, ki so potekale med predstavitvami, pa so udeležence spodbudile k nadaljevanju raziskav in jim ponudile vpogled v interdisciplinarnost različnih tematskih podpolj jezikovnih tehnologij in digitalne humanistike. Preprosto povedano, konferenca je prinesla priložnost za tesnejše sodelovanje s sorodnimi področji ter postavila nove gradnike pri premoščanju vrzeli med jezikovnimi tehnologijami in digitalno humanistiko. Izbrani prispevki bodo v obliki izvirnih znanstvenih člankov na voljo v naslednji tematski številki revije *Slovenščina 2.0*.

---

8 Konferenca Jezikovne tehnologije in digitalna humanistika: <http://www.sdjtsi.si/wp/dogodki/konference/jtdh-2020/>.

## LITERATURA

- Antloga, Š. (2020). Korpus metafor KOMET 1.0. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 167–170). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Student\\_Antloga\\_Korpus-metafor-KOMET-1.0.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Student_Antloga_Korpus-metafor-KOMET-1.0.pdf)
- Beelen, K. (2020). Speaking on behalf of others: Why the digital humanities should care about parliamentary data. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika: 2*. Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Invited-Lecture\\_Beelen\\_Speaking-on-behalf-of-others-Why-the-digital-humanities-should-care-about-parliamentary-data.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Invited-Lecture_Beelen_Speaking-on-behalf-of-others-Why-the-digital-humanities-should-care-about-parliamentary-data.pdf)
- Fijavž, Z. (2020). Ambivalence of LGBT Visibility in Croatian Video-Based Social Media Content. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 144–148). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Student\\_Fijavz\\_Ambivalence-of-LGBT-Visibility-in-Croatian-Video-Based-Social-Media-Content.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Student_Fijavz_Ambivalence-of-LGBT-Visibility-in-Croatian-Video-Based-Social-Media-Content.pdf)
- Konferenca Jezikovne tehnologije in digitalna humanistika. Dostopno prek <http://www.sdjt.si/wp/dogodki/konferenca/jtdh-2020/>
- Krek, S., Erjavec, T., Dobrovoljc, K., Gantar, P., Arhar Holdt, Š., Čibej, J. in Brank, J. (2020). The ssj500k Training Corpus for Slovene Language Processing. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 24–33). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Krek-et-al\\_The-ssj500k-Training-Corpus-for-Slovene-Language-Processing.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Krek-et-al_The-ssj500k-Training-Corpus-for-Slovene-Language-Processing.pdf)
- Lemmenmeier-Batinić, D., Ljubešić N. in Samardžić, T. (2020). XML-Encoding of a spoken Serbian corpus targeting forms of address. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 127–130). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Lemmenmeier-Batinic-et-al\\_XML-Encoding-of-a-spoken-Serbian-corpus-targeting-forms-of-address.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Lemmenmeier-Batinic-et-al_XML-Encoding-of-a-spoken-Serbian-corpus-targeting-forms-of-address.pdf)
- Lenardič, J. in Fišer, D. (2020). Epistemic modal adverbs in Slovenian academic discourse. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 34–41).



- Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Lenardic-et-al\\_Epistemic-modal-adverbs-in-Slovenian-academic-discourse.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Lenardic-et-al_Epistemic-modal-adverbs-in-Slovenian-academic-discourse.pdf)
- Meden, K. in Cvek, A. (2020). Nadgradnja Zgodovinarskega indeksa citiranosti. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 42–47). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Meden-et-al\\_Nadgradnja-Zgodovinarskega-indeksa-citiranosti.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Meden-et-al_Nadgradnja-Zgodovinarskega-indeksa-citiranosti.pdf)
- Pahor de Maiti, K., Fišer, D., Ljubešić, N. in Erjavec, T. (2020). Grammatical Footprint of Socially Unacceptable Facebook Comments. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 48–57). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_PahordeMaiti-et-al\\_Grammatical-Footprint-of-Socially-Unacceptable-Facebook-Comments.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_PahordeMaiti-et-al_Grammatical-Footprint-of-Socially-Unacceptable-Facebook-Comments.pdf)
- Pančur, A., Blaj Hribar, N., Ojsteršek, M. in Šorn, M. (2020). Projekt Vojaške žrtve prve svetovne vojne na Slovenskem. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 136–140). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Pancur-et-al\\_Projekt-Vojaske-zrtve-prve-svetovne-vojne-na-Slovenskem.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Pancur-et-al_Projekt-Vojaske-zrtve-prve-svetovne-vojne-na-Slovenskem.pdf)
- Schlintl, M., Pawluch, K., Rader, M. in Novak-Geiger, V. (2020). Working with Digital Devices in Teacher Training Using the Example of the Teaching-Learning-Lab. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 171–174). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Student\\_Schlintl-et-al\\_Working-with-Digital-Devices-in-Teacher-Training-Using-the-Example-of-the-Teaching-Learning-Lab.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Student_Schlintl-et-al_Working-with-Digital-Devices-in-Teacher-Training-Using-the-Example-of-the-Teaching-Learning-Lab.pdf)
- Supej, A., Ulčar, M., Robnik-Šikonja, M. in Pollak, S. (2020): Primerjava slovenskih besednih vektorskih vložitev z vidika spola na analogijah poklicev. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 93–100). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Supej-et-al\\_Primerjava-slovenskih-besednih-vektorskih-vlozitev-z-vidika-spola-na-analogijah-poklicev.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Supej-et-al_Primerjava-slovenskih-besednih-vektorskih-vlozitev-z-vidika-spola-na-analogijah-poklicev.pdf)

- Tonelli, S. (2020). Abusive Language Detection: Too much Digital, not enough Humanities?. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika: 1*. Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Invited-Lecture\\_Tonelli\\_Abusive-Language-Detection-Too-much-Digital-not-enough-Humanities.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Invited-Lecture_Tonelli_Abusive-Language-Detection-Too-much-Digital-not-enough-Humanities.pdf)
- Trivunović, E. (2020). Variante in modifikacije (iz)biblijskih frazemov. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 158–166). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Student-Trivunovic\\_Variante-in-modifikacije-\(iz\)biblijskih-frazemov.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Student-Trivunovic_Variante-in-modifikacije-(iz)biblijskih-frazemov.pdf)
- Verdonik, D., Rojc, M., Mlakar, I., Zimšek, D., Kačič, Z., Ojsteršek, M. in Erjavec, T. (2020). Raziskovalna infrastruktura projekta RI-SI CLARIN. V D. Fišer in T. Erjavec (ur.), *Zbornik konference Jezikovne tehnologije in digitalna humanistika* (str. 141–143). Ljubljana: Inštitut za novejšo zgodovino. Dostopno prek [http://nl.ijs.si/jtdh20/pdf/JT-DH\\_2020\\_Verdonik-et-al\\_Raziskovalna-infrastruktura-projekta-RI-SI-CLARIN.pdf](http://nl.ijs.si/jtdh20/pdf/JT-DH_2020_Verdonik-et-al_Raziskovalna-infrastruktura-projekta-RI-SI-CLARIN.pdf)