

PROCESI JEZIKOVNEGA SEPARATIZMA PRI ČEZMEJNIH JEZIKOVNIH MANJŠINAH: PREVZEMANJE, PRILAGAJANJE IN PREVAJANJE COVIDNE TERMINOLOGIJE MED SLOVENCI IN SLOVENKAMI V ITALIJI

Matejka GRGIČ

Univerza v Ljubljani, Filozofska fakulteta, Aškerčeva 2, 1000, Ljubljana, Slovenija
SLORI – Slovenski raziskovalni inštitut, Ul. Beccaria 6, 34133, Trst, Italija
e-mail: matejka.grgic@ff.uni-lj.si

Damjan POPIČ

Univerza v Ljubljani, Filozofska fakulteta, Aškerčeva 2, 1000, Ljubljana, Slovenija
e-mail: damjan.popic@ff.uni-lj.si

IZVLEČEK

Članek prinaša rezultate primerjalne korpusne analize, izvedene na besedilih, ki so izšla v Primorskem dnevniku (slovenskem dnevniku, ki izhaja v Italiji) v razdobju 2020–2021. Ta besedila primerjamo s podatki iz slovenskih korpusov, pri tem pa se omejujemo na leksiko, povezano z epidemijo covid-19. Namen članka je prikazati stopnjo razlikovanja leksike slovenskega standardnega jezika na obeh straneh meje. Analiza je namreč potrdila hipotezo, da prihaja na tem območju do visoke variantnosti, ki je ne moremo pripisati dejavnikom, s katerimi običajno razlagamo odstopanja od standarda, pojave jezikovnega stikanja in jezike v manjšinskem položaju.

Ključne besede: Slovenci v Italiji, manjšinski jeziki, terminologija, covid-19, sociolingvistika, korpusno jezikoslovje

PROCESSI DI SEPARATISMO LINGUISTICO TRA LE MINORANZE LINGUISTICHE TRANSFRONTALIERE: ADOZIONE, ADATTAMENTO E TRADUZIONE DELLA TERMINOLOGIA COVID-19 TRA GLI SLOVENI IN ITALIA

Il contributo presenta i risultati di un'analisi comparativa condotta su corpora di testi pubblicati dal Primorski dnevnik (quotidiano in lingua slovena edito in Italia) nel periodo 2020-2021. I dati emersi sono stati confrontati con quelli rilevati da altri corpora sloveni, con particolare attenzione al lessico riferito al contesto epidemiologico da covid-19. L'obiettivo è mostrare il grado di differenziazione della lingua slovena standard nelle due aree a ridosso del confine italo-sloveno. L'analisi ha confermato l'ipotesi secondo la quale il territorio in questione sarebbe interessato da un alto grado di variabilità linguistica non associabile a fattori che normalmente incidono su tali discostamenti, su fenomeni di contatto linguistico o altre dinamiche tipiche delle lingue minoritarie.

Parole chiave: Sloveni in Italia, lingue minoritarie, terminologia, Covid-19, sociolinguistica, corpora linguistici

UVOD

Jezikovni separatizem poimenujemo proces, po katerem (manjša) skupnost govorcev in govork nekega jezika uporablja pretežno ali izključno regionalne, lokalne različice tega jezika, ki se zaradi pojavov intenzivnega stikanja z drugim jezikom in/ali marginalizacije oz. izključevanja govorcev in govork od svoje širše referenčne skupnosti vse bolj oddaljujejo od jezikovnega kontinuuma, znotraj katerega so se razvile. V tem smislu je jezikovni separatizem oblika jezikovne variantnosti, se pa od običajne/pričakovane variantnosti razlikuje po stopnji izraženosti pojavov, vrsti pojavov, nekaterih vzrokih in predvsem po (sociolingvističnih) posledicah.

Proces, ki ga sicer spremljamo pri različnih skupnostih govorcev in govork, ima pri čezmejnih manjšinskih skupnostih nekatere specifične, ki jih bova avtorja osvetlila v tem prispevku, in sicer na primeru leksike, ki se je pojavila ob izbruhu koronavirusne bolezni leta 2020, in ukrepov, ki so sledili v letih 2020 in 2021.

Izhajala bova iz podatkov o rabi – ali, natančneje, o prevzemanju, prilagajanju in prevajanju – terminologije, povezane s koronavirusno boleznijo (v nadaljevanju tudi: covidom). Primer te leksike je z več vidikov izjemno reprezentativen za pojav, ki ga želiva osvetliti. Prvič, raba terminologije je znotraj sporazumevalnih praks v nekem jeziku specifična, saj teži (v večji meri od splošne leksike) k enotnosti/doslednosti, enoznačnosti in ustaljenosti. To pomeni, da je raba leksemov, ki jih percipiramo kot termine, v večji meri implicitno ali eksplicitno standardizirana, kar nadalje pomeni, da temelji na konsenzu govorcev in/ali na normiranju ter da se redkeje (če sploh) spreminja v času in prostoru. Drugič, nenadni izbruh covida z vsemi družbenimi in kulturnimi posledicami (Riman, 2022) je zahteval hitro prilagajanje jezika novim okoliščinam in sporazumevalnim potrebam, kar se je, tudi zaradi značilnosti pojava, pokazalo prav na ravni leksike in, specifično, terminologije. Tretjič, ker je šlo za globalen pojav s podobnimi posledicami v Italiji in Sloveniji, ne moremo govoriti o potrebi po specifični terminologiji znotraj skupnosti govorcev

in govork slovenskega jezika v Italiji, češ da gre za poimenovanje specifične predmetnosti, kot na primer na področju (upravnega) pravna, vzgoje in izobraževanja, financ itd.

V nadaljevanju izhajava torej iz predpostavke, da bi zaradi narave samega pojava, tj. pandemije koronavirusne bolezni, ki je skoraj istočasno zajela ves svet, pričakovali enotno, enoznačno in ustaljeno rabo področne terminologije v celotnem slovenskem jezikovnem prostoru. Najina hipoteza pa je, da izkazuje korpus besedil v slovenskem standardnem jeziku, nastalih na območju naselitve slovenske manjšine v Italiji, variantnost, ki je ne moremo pripisati dejavnikom, s katerimi običajno razlagamo odstopanja od standarda, pojave jezikovnega stikanja in jezike v manjšinskem položaju, npr. vplivu večinskega/dominantnega jezika (italijanščine) ali različni predmetnosti, ampak predvsem jezikovni marginalizaciji, tj. omejenim jezikovnim stikom z referenčnim slovenskim prostorom. Hipotezo bova preverjala s primerjalno analizo korpusa, ki zajema dva letnika časopisa *Primorski dnevnik* in ki ga bova primerjala z obstoječim referenčnim korpusom za slovenščino, obenem pa tudi z rezultati novega jezikovnotehnološkega orodja za slovenščino, Sledilnika CJVT.¹

IZHODIŠČA

V skupnost govork in govorcev slovenskega jezika v Italiji vključujemo rojene govorce in govork slovenščine, ki se opredeljujejo za (zamejske) Slovence in Slovenke, nato rojene govorce in govork slovenščine, ki se opredeljujejo drugače (Pertot, 2017), in številčno manj relevantno skupino drugih govorcev in govork, ki so se slovenščine naučili v odrasli dobi.² Če se omejimo zgolj na slovensko skupnost (torej na skupnost tistih, ki se opredeljujejo za Slovence oz. Slovenke), govorimo o avtohtoni čezmejni narodni manjšini, torej o skupini prebivalstva, ki je bila v določenem obmejnem prostoru prisotna že pred vzpostavitvijo trenutne državne meje, ki je na tem območju manj številna od večinskega prebivalstva in ki se identificira v odnosu z referenčnim kulturnim in nacionalnim prostorom onkraj meje (Brezigar, 2017, 22–23; Jagodic, 2017).

1 Raziskava je delno nastala v sklopu sodelovanja med SLORI – Slovenskim raziskovalnim inštitutom in Znanstveno-raziskovalnim središčem Koper. V času zbiranja gradiva in nastajanja prispevka (2021) je bila Matejka Grgič članica programske skupine ZRS Koper P5-0409 (Razsežnosti slovenstva med lokalnim in globalnim v začetku tretjega tisočletja), Damjan Popič pa član programske skupine P6-0215 (Slovenski jezik – bazične, kontrastivne in aplikativne raziskave). Oba programa je sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije iz državnega proračuna. Na dan oddaje prispevka v objavo (15. 3. 2023) sta oba avtorja člana programske skupine P6-0215 (Slovenski jezik – bazične, kontrastivne in aplikativne raziskave).

2 Spomnimo, da slovenščina v Italiji ni obvezen učni predmet na šolah z večinskim (italijanskim) jezikom, zato spadajo v to kategorijo predvsem govorcev in govork, ki sicer izhajajo iz italijansko govorečih družin, a so obiskovali šole s slovenskim učnim jezikom ali dvojezično šolo v Špetru (Bogatec, 2015) oz., redkeje, govorcev in govork, ki so se slovenščine naučili v odrasli dobi kot drugega/tujega/sosedskega jezika (Cavaion, 2014).

Slovenska skupnost v Italiji je – ponekod prej, ponekod šele pred kratkim – razvila specifične jezikovne poteze, ki so bile predvsem v zadnjih dveh desetletjih predmet številnih raziskav (Grgič, 2017; Jagodic et al., 2017; Mezgec, 2016; Melinc Mlekuž, 2019). Vse doslej izvedene raziskave pa so bile količinsko omejene na manjše korpuse besedil oz. na preverjanje domnev o posameznih pojavih jezikovnega stikanja. Take analize so potrdile prisotnost tovrstnih pojavov, niso je pa mogle kvantitativno meriti.

Pri poskusu odgovora na vprašanje, v katero smer se razvija jezik govorcev in govork slovenščine v Italiji, ki se opredeljujejo za Slovence in Slovenke, bova v pričujočem prispevku izhajala iz izhodišča, da je vsak jezikovni kontinuum in vsak njegov (lokalni) segment artikularna, kompleksna struktura sporazumevalnih sredstev, podkodov, ki jih govorcev in govork izbirajo in uporabljajo glede na njihovo pragmatično in simbolično funkcijo (Chambers, 2003; Grgič, 2016a). V tem smislu je torej nemogoče govoriti o nečem, kar bi preprosto poimenovali »značilnost slovenščine v Italiji«. Prav tako je težko določiti, kje je v jeziku nasploh meja med standardnimi in nestandardnimi različicami (Davila, 2016). Za potrebe pričujočega prispevka bova kot besedila, napisana v standardnem jeziku, obravnavala besedila, ki so namenjena javnosti in niso napisana v lokalnem narečju.³

S sodobnimi metodami korpusnega jezikoslovja lahko na širšem segmentu besedil preverimo, katere izbire govorcev in govork slovenščine v Italiji se razlikujejo od izbir govork in govorcev slovenščine v Sloveniji in posledično od besedil, ki nastajajo v primerljivih okoliščinah in s primerljivim namenom. Na podlagi tega lahko spremljamo stopnjo oddaljenosti dveh ali več kodov, variant oz. idiomov znotraj danega jezikovnega kontinuuma. To najbolj opazimo pri besedilih, v katerih je pričakovana variantnost minimalna – to so strukturirana besedila z ustaljenimi sporazumevalnimi vzorci in visoko koncentracijo terminov za poimenovanje iste predmetnosti (eno-/istoreferenčni termini oz., drugače rečeno, termini z eno/isto referenco na obeh straneh meje). Skratka: stopnjo variantnosti najlažje opazujemo, če omejimo vse ostale spremenljivke, razen kraja nastanka besedila. Pri korpusni analizi teh besedil lahko opazujemo procese krepitve in šibitve jezikovnega kontinuuma, torej približevanja in oddaljevanja od standardov, ki so se uveljavili oz. se uveljavljajo

v referenčnem prostoru za razvoj slovenskega jezika. Taki procesi ne zadevajo samo terminologije in terminoloških izbir, ampak tudi druge jezikovne prvine in ravni, na primer splošno leksiko, oblikoslovje, besedotvorje, skladnjo itd. Medtem ko je variantnost pri splošnem, nestrokovnem sporazumevanju višja in sploh prisotna v vsem jezikovnem kontinuumu, saj predstavlja enega njegovih značilnih elementov (Chambers, 2003), je v znanstvenih, strokovnih in poljubnih znanstvenih ali strokovnih besedilih manj pogosta in na splošno »manj zaželeno«.

Redke raziskave, ki so bile doslej opravljene na področju slovenske terminologije in strokovnega jezika v Italiji (Grgič, 2020, 104), se običajno nanašajo na področja, kjer prihaja do večjih razlik pri poimenovanju predmetnosti in, posledično, do potreb po specifičnih terminih. Tako področje je predvsem področje (upravnega) prava, v manjši meri tudi področji (javnih) financ in računovodstva ter šolstva – ker je institucionalna ureditev teh področij v Italiji drugačna kot v Sloveniji, se pojavlja tudi potreba po drugačni, specifični terminologiji, s katero bi v slovenščini poimenovali italijansko oz. lokalno predmetnost. Vprašanje je seveda, kdaj je uvedba novega, drugačnega termina nujna, kako ta termin izbrati in kako zadovoljiti potrebi po njegovi dosledni rabi vsaj znotraj lokalnega okolja (Grgič, 2020, 109).

Pri raziskavah s področja terminologije na naselitvenem področju Slovencev v Italiji (in v drugih primerljivih kontekstih) moramo nazadnje upoštevati še eno specifiko dvojezičnega okolja. Namreč, pomemben segment besedil nastaja kot prevod ali prilagoditev (npr. povzemanje) italijanskih izvirnikov. To velja npr. za novice, ki jih objavljajo mediji (in jih oblikujejo na podlagi novinarskih sporočil ali agencijskih vesti, napisanih v italijanščini), za raznovrstna sporočila, ki jih svojim deležnikom pošiljajo slovenske šole, dvojezične občine, javni zavodi, stanovska združenja in subjekti zasebnega prava (podlaga zanje so pravilniki, smernice in druga besedila nastala v italijanskem jeziku) ter za druga slovenska besedila. Potreba po visoko razvitih in splošno razširjenih prevajalskih veščinah je torej na takem območju izrazita, izobraževalni sistem pa je ne podpira, saj tovrstnih vsebin šolski kurikulumi ne predvidevajo, možnosti študija prevajanja v slovenščino na lokalnih univerzah so močno okrnjene, sploh pa prevajanje ni prisotno kot transverzalna kompetenca v izobraževalnih programih na različnih stopnjah (Melinc, 2019; Bogatec, 2018).

3 Primeri rabe narečij v okoliščinah, kjer bi pričakovali nelokalne oz. standardne variante jezika, sta npr. spletni strani Občine Rezija (<http://www.comune.resia.ud.it/>) in tednika Novi Matajur, ki poleg prispevkov v italijanščini in standardni slovenščini objavlja tudi novice in komentarje v lokalnem narečju (<https://novimatajur.it/>).

METODOLOGIJA

V tem odseku opišemo potek in metodologijo analize ter izgradnjo jezikovnega vira, na katerem temelji naša analiza leksike, povezane s covidom-19. V osnovi je analiza v prispevku dvoravninska:

1. Najprej poskušamo ugotoviti, kakšna je področna leksika v zamejskem okolju, kar lahko najobjektivneje storimo s korpusom. Ker je tematika prispevka vezana na relativno novo pojavnost, je obstoječi jezikovni viri seveda ne zajemajo. Zaradi tega smo morali za izvedbo raziskave zgraditi lasten korpus iz relevantnih besedil s področja jezikovnega stika in ga primerjati z že obstoječimi besedilnimi zbirkami, da lahko tako identificiramo izstopajoče jezikovne elemente ali strukture.⁴
2. Ko lahko izoliramo celotno leksiko, ki se uporablja zgolj v zamejskem okolju, v centralnem standardu pa je ne najdemo, je v tej podmnožici tudi celotno besedje, vezano na pandemijo. To besedje lahko potem primerjamo z rezultati Sledilnika CJVT (<https://viri.cjvt.si/sledilnik/slv/>), ki za primerljivo časovno obdobje (2021) podaja izstopajoče pogosto besedje v slovenskih spletnih medijih.

V naslednjem odseku opišemo novi korpus in njegovo strukturo, obenem pa tudi referenčni korpus, ki smo ga izbrali za primerjavo.

Podatki

Kot gradivo za lastni korpus smo uporabili dva letnika dnevnika *Primorski dnevnik*,⁵ in sicer vse številke iz let 2020 in 2021. Zaradi (potencialnega) razlikovanja med letnikoma smo oba letnika oblikovali kot podkorpusa, tako da je mogoče analizirati zgolj posamezno leto in tako poizvedovati za morebitnimi spremembami v jezikovni rabi (npr. ali se je raba v enem letu značilno spremenila). V Tabeli 1 so podani pregled in osnovne lastnosti korpusa.

To pomeni, da smo v korpus zajeli 618 številčk dnevnika, ki skupaj zajemajo 22.958.558 pojavnic oz. 17.812.352 besed, ki skupaj tvorijo 848.547 povedi, same pojavnice pa zasedajo 1361 oblikoskladenjskih oznak.⁶

Pri tem je nujno poudariti, da smo oba letnika časopisa prejeli v formatu PDF, kar pomeni, da je v ustvarjenem korpusu tudi določena količina šuma, ki nastane zaradi same narave formata PDF. Šum je vezan predvsem na oblikovanje dokumentov, saj je to v formatu PDF fiksno, za razliko od denimo zapisa v formatu XML. Zaradi tega so pojavnice v korpusu včasih nepravilno

Tabela 1: Korpus v številkah.

Kategorija	Število
Besedila (številčk časnika <i>Primorski dnevnik</i>)	618
Pojavnice	21.501.607
Besede	17.812.352
Uporabljene oblikoskladenjske oznake	1.361
Povedi	848.547

združene (npr. *Primorskidnevnik*) ali razpisane (npr. pri deljajih: *razdeli-ti*), prav tako pa so deli besedil nepravilno združeni. Po pregledu ene od številčk *Primorskega dnevnika* ocenjujemo, da je v korpusu zaradi vhodnega formata do 5 % šuma.

Zaradi velikosti korpusa je možnost ročnega popravljanja napačno segmentiranih besed v danih okvirih praktično nemogoča, hkrati pa šum ne onemogoča objektivne korpusne analize, in sicer iz dveh razlogov:

1. Če se pri korpusni analizi fokusiramo na najpogostejše pojavnice, se med njimi šum ne pojavlja, saj so napake pri pretvorbi arbitrarne in na ravni celotnega korpusa posledično marginalne.
2. Če se pri korpusni analizi fokusiramo na redke pojavnice, moramo v vsakem primeru rezultate pregledati ročno, s čimer lahko šum odpravimo.

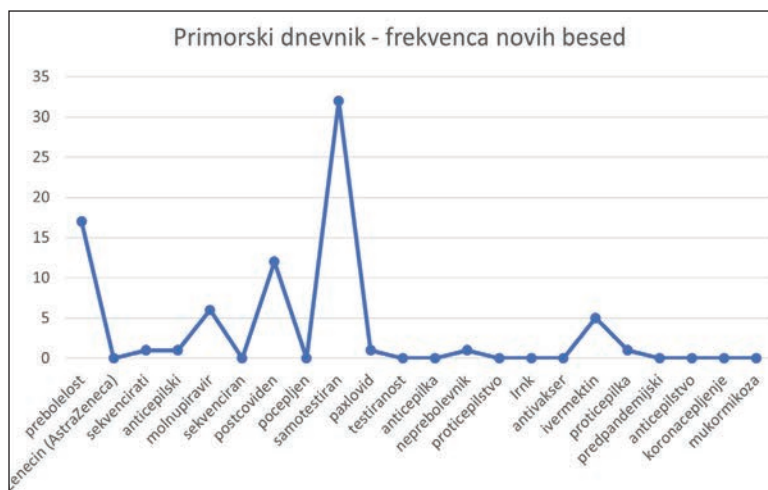
Nastali korpus smo analizirali, obenem pa smo ga primerjali z že obstoječim korpusom slovenščine, in sicer smo za to nalogo izbrali korpus Slovenian Web 2015 (slTenTen15). Ker smo za analizo izbrali orodje SketchEngine, smo morali za primerjavo izbrati čim bolj primerljivo zgrajen in označen korpus, ki je moral hkrati zajemati čim več besedil in biti čim bolj sodoben. Tem pogojem je najbolj zadostil korpus slTenTen15, ki premore skoraj milijardo pojavnic.⁷ Pri analizi posameznih letnikov smo kot izhodišče izbrali enega od podkorpusov, kot referenčni korpus pa preostali del korpusa *Primorskega dnevnika*. Na ta način lahko primerjamo specifično besedje v enem izmed podkorpusov in s tem posledično primerjamo celoten letnik z drugim. V naslednjem razdelku podajamo rezultate analize po posameznih vidikih primerjave.

⁴ Tako za gradnjo korpusa kot tudi poznejšo analizo smo uporabili orodje SketchEngine (<https://www.sketchengine.eu/>).

⁵ Na tem mestu se iskreno zahvaljujemo celotnemu uredništvu in tehničnemu osebju *Primorskega dnevnika*, da so nam omogočili dostop do zbirke besedil.

⁶ Seznam oznak in njihov opis je na voljo na spletni strani <https://www.sketchengine.eu/slovene-tagset-sl/>.

⁷ Korpus vsebuje 988.513.467 pojavnic oziroma 829.544.337 besed.



Slika 1: Distribucija novih besed v korpusu Primorskega dnevnika.

ANALIZA

V tem odseku podamo rezultate korpusnojezikoslovne analize, pri kateri se fokusiramo na različne vidike leksemov, vezanih na epidemijo covid-19.

Nove besede

Sodeč po podatkih Sledilnika CJVT, so bili v letu 2021 v slovenskih spletnih medijih pogosti in statistično relevantni naslednji leksemi, ki jih podajamo v Tabeli 2, zraven pa dodajamo še frekvence teh leksemov v korpusu besedil časnika *Primorski dnevnik*, če se te tam seveda pojavljajo. V skladu s tematiko prispevka smo v tabelo vključili zgolj lekseme, vezane na covid-19.⁹

Kot prikazuje Tabela 2, se nove besede v letu 2021, ki so vezane na covid-19, nanašajo predvsem na cepljenje oziroma nasprotovanju cepljenju (to pomensko polje je še posebej dobro zastopano in variantno), učinkovine/zdravila in cepiva ter na status (preboleznosti ipd.). Večina leksemov, ki so se v letu 2021 množično pojavili v slovenskih spletnih medijih, je v *Primorskem dnevniku* izostala, zgolj peččica je bila normalno zastopana. Distribucijo lem v korpusu ponazarja Slika 1.

Kot prikazuje Slika 1, najdemo v korpusu *Primorskega dnevnika* predvsem izraze, vezane na samotestiranje in prebolelost, drugih izrazov je izjemno malo ali pa jih sploh ni. Zanimivo je, da so v korpusu skoraj povsem umanjale besede za nasprotnike cepljenja (*anticepilski*, *anticepilka*, *proticepilstvo*, *antivakser*, *proticepilka*, *anticepilstvo*), najdemo zgolj eno pojavitev *anticepilke* in eno pojavitev *proticepilke*. Prav tako se v določeni meri pojavljajo (domnevna) zdravila

oziroma cepiva (*AstraZeneca*, *paxlovid*, *ivermektin*). V naslednjem poglavju obravnavamo najbolj značilne besede v slovenskih spletnih medijih v letu 2021 in jih primerjamo z njihovo rabo v korpusu *Primorskega dnevnika*.

Najbolj značilne besede v slovenskih spletnih medijih (2021)

Pri analizi pogostnosti najbolj značilnih leksemov v slovenskih spletnih medijih v letu 2021 smo najprej prenesli seznam s Sledilnika CJVT in lema pripisali frekvence iz korpusa *Primorskega dnevnika*. S seznama 100 leksemov smo izločili vse, ki se ne navezujejo na covid-19, pri tem pa smo bili izjemno permissivni, tako da smo upoštevali tudi šibke/potencialne pomenske povezave, zato smo denimo obdržali tudi lekseme, kot sta denimo *izpolnjevanje* (v smislu pogoja PCT), *različica* (virusa), *temnordec* (regija ipd. na covidnem zemljevidu) ipd. Na podlagi tega merila smo obdržali 53 leksemov, kar pomeni, da je bila v letu 2021 več kot polovica vsega najbolj značilnega besedja v slovenskih spletnih medijih vezana na epidemijo covid-19. Leksemi so podani v Tabeli 3, spremljajo pa jih podatki o pogostnosti pojavljanja v korpusu *Primorskega dnevnika*, zaporedno mesto po frekvenci pojavljanja v korpusu (1.–53.), zaporedno mesto po relevantnosti glede na Sledilnik CJVT (1.–53.) in absolutni razliki med obema uvrstitvama. V tabeli so leme razvrščene po absolutni razliki med relevantnostjo v slovenskih spletnih besedilih leta 2021 in njihovo frekvenco v korpusu *Primorskega dnevnika*.

Kot prikazuje Tabela 3, je relevantnost leksema v obeh jezikovnih virih, korpusu *Primorskega dnevnika* in Sledilniku CJVT, identična zgolj v enem primeru (*precepljenost*). Pri

8 Imenovalniška oblika ni vključena na seznamu Sledilnika CJVT.

9 Tovrstni leksemi so sicer lani močno prevladovali; poleg leksemov, vezanih na covid-19, so se na seznam novih besed na sledilniku prebile zgolj naslednje leme: *metaverzum*, *biopropan*, *neunijski*, *brezupnica*, *astrotrufing*, *kriptoposel*, *paraodbojka*, *protishod*, *NFT*, *zvo-novje*, *politkolesariat*, *monosežigalnica*, *nevrorazličen*, *klančarski*, *kvačkar*, *zaplezanost*, *paralokostrelec*.

Tabela 2: Nove relevantne besede (Sledilnik CJVT) in frekvenca v Primorskem dnevniku.

Sledilnik – lema	Primorski dnevnik – frekvenca
prebolelost	17
AstraZenecin (AstraZeneca) ⁸	0 (270)
sekvencirati	1
anticepilski	1
molnupiravir	6
sekvenciran	0
postcoviden	12
pocepljen	0
samotestiran	32
paxlovid	1
testiranost	0
anticepilka	0
neprebolevnik	1
proticepilstvo	0
iRNK	0
antivakser	0
ivermektin	5
proticepilka	1
predpandemijski	0
anticepilstvo	0
koronacepljenje	0
mukormikoza	0

drugih leksemih je v rabi znatno razhajanje, v največji meri pri leksemu *cepivo*, ki je od vseh najbolj značilnih leksemov leta 2021 v korpusu najbolj zastopan, medtem ko se na Sledilniku CJVT ni znašel med najbolj značilnim besedjem. Po drugi strani pa tudi najbolj značilna beseda sledilnika (*PCT*) v *Primorskem dnevniku* ni tako pogosto zastopana, kot bi pričakovali. Podobno velja tudi za kratico *PCR*, ki je v korpusu sicer pogosta, vendar manj od pričakovanega glede na relevantnost. Po drugi strani pa je v korpusu izjemno pogosta beseda *potrdilo*, ki je sicer po podatkih sledilnika pogosta, a ne v tolikšni meri. Ti podatki kažejo, da tovrstna primerjalna analiza lahko poda uvid v razlike v rabi področnega besedja in nas s tem tudi usmerja pri nadaljevanju analize.

Primerjava splošnih besed

V naslednjem segmentu primerjamo 50 splošnih besed s seznama besed, ki se v slovenskem jeziku

uporabljajo najpogosteje (Pollak idr., 2020), ki so se v letu 2021 uporabljale najpogosteje, zatem pa ta seznam primerjamo s frekvenčnim seznamom iz korpusa *Primorskega dnevnika*. Seznam je podan v Tabeli 4, krepko je pisana leksika, vezana na epidemijo covid-19.

Kot lahko razberemo iz Tabele 4, je od 50 najpogosteje rabljenih besed v slovenščini v letu 2021 kar polovica povezana s covidom-19, obenem pa lahko vidimo tudi, da so visoke tudi frekvence v korpusu *Primorskem dnevniku*. Če na primer primerjamo relativno frekvenco najpogostejših petih splošnih besed v slovenščini (ukrep, maska, širjenje, obvezen, zdravje) v *Primorskem dnevniku* z referenčnim korpusom slovenščine Gigafida 2.0,¹⁰ dobimo distribucijo, kot jo ponazarja Tabela 5.

Oba stolpca pri obeh zbirkah podatkov predstavljata isto enoto (relativno frekvenco, tj. frekvenco na milijon besed), a zaradi ilustrativnosti ohranjamo obe. Na ta način lahko vidimo, v kolikšni meri izrazi, vezani na covid-19, v korpusu *Primorskega dnevnika* po frekvenci izstopajo. Glede na to, da gre pri Gigafidi 2.0 za referenčni korpus, tj. za korpus, ki v največji možni meri odslkava jezik, kakršen je oziroma kot ga je s korpusom (trenutno) sploh mogoče ponazoriti, nam te številke povedo, kako močno se raba jezika oziroma vsaj naštetih izrazov razlikuje od *splošnega jezika*. Kot kažejo številke, je razhajanje precejšnje, in sicer je splošni izraz *ukrep* približno trikrat pogostejši kot v referenčnem korpusu, medtem ko sta *maska* in *širjenje* kar približno desetkrat pogostejša. *Obvezen* je pogostejši približno štirikrat, medtem ko je pri zdravju razmerje nekoliko manj izrazito v korist korpusa *Primorskega dnevnika*, in sicer približno 3 : 4. Izraz *zdravje* je torej tudi v splošnem jeziku dovolj pogost, hkrati pa v korpusu ni dovolj izpostavljen, da bi bila razlika višja.

Kot vidimo, je bila raba specifične splošne leksike (tiste, ki je vezana na covid-19) v korpusu znatno izrazitejša kot v referenčnem korpusu. V nadaljevanju primerjamo specifično leksiko, vezano na epidemijo covid-19, po posameznih tematskih področjih.

Področna analiza covidne leksike

Covidno leksiko, obravnavano v tem poglavju, lahko grobo razdelimo na a) leksiko, vezano na status in potrjevanje, b) leksiko, vezano na testiranje, in c) leksiko, vezano na poučevanje, študij in delo. Izraze, vezane na status in potrjevanje, ki smo jih našli in izpostavili z analizo, podaja Tabela 6.

Tabela 6 navaja izraze za status (prebolevnost) in potrjevanje statusa med epidemijo covid-19, ki smo jih zaznali pri primerjavi pogostega besedja in ročnim pregledovanjem zadetkov (gl. podpoglavje 4.2). S

¹⁰ Največji korpus slovenskega jezika Gigafida 2.0 zajema besedje do leta 2018, tako da ne zajema leksike o novem koronavirusu, je pa občutno večji, zato so absolutne frekvence bistveno večje, tako da primerjamo relativne frekvence.

Tabela 3: Seznam najbolj relevantnih besed Sledilnika CJVT s pripisanimi frekvencami v korpusu.

Lema	▼ Frekvenca v korpusu Primorskega dnevnika ▼	Mesto po relevantnosti (Sledilnik CJVT)	▼ Absolutna razlika (PD: Sledilnik) ▼
cepivo	3683	1	44
venski	841	7	48
izpolnjevanje	220	14	53
različica	996	6	41
potrdilo	3605	2	35
samotestiranje	24	29	2
precepljen	3	43	17
prebolelost	17	33	8
cepljenje	2799	3	26
prebolevnost	11	37	15
cepljenost	6	39	21
sekvenciranje	1	47	29
gama	97	20	38
Janssen	30	28	11
antigenski	106	19	36
strjevanje	18	32	49
odmerek	1694	4	18
neželen	41	27	40
temnordeč	10	38	50
delta	185	15	4
prebolevník	126	17	6
epidemičen	12	35	46
trombocit	11	36	47
proticepilski	6	41	52
PCT	420	11	1
južnoafriški	83	22	32
Modernin	2	44	34
vektorski	21	31	22
omikron	47	26	19
prebolel	62	23	30
HAT	6	40	33
tromboza	21	30	37
Moderna	124	18	13
sedemdneven	59	24	20
cepljen	826	8	5
AstraZeneca	270	12	9
strdek	95	21	24
precepiti	13	34	31
pocepiti	1	45	42
proticepilec	1	46	43
nenaročen	5	42	45
ivermektin	1	48	51
cepiti	1333	5	3
cepilen	610	9	7
PCR	523	10	12
Pfizerjev	50	25	27
necepljen	231	13	14
precepljenost	155	16	16

Tabela 4: Frekvenca splošnih besed v korpusu PD.

ukrep	11238
spleten	10873
maska	4729
širjenje	4568
dijak	4191
obvezen	4175
zdravje	3743
potrdilo	3605
bolnik	3262
test	3033
vabljen	2919
pouk	2915
pozitiven	2879
val	2865
omejitev	2681
izpostaviti	2514
nega	2177
omrežje	2136
faza	1844
pozvati	1833
protest	1702
dovoljen	1652
digitalen	1579
blizu	1554
intenziven	1552
negativen	1543
regija	1495
pozivati	1268
vključno	1254
nahajati	1252
navesti	1241
portal	1101
porast	1049
živo	1025
različica	996
soočati	935
priporočilo	929
oskrba	822
potrjen	755
pameten	660
zdraviti	643
populacija	545
kraljestvo	433
preboleti	362
terasa	329
dobava	299
zadati	237
vesti	194
izboljševati	150
kopalnica	113

krepro smo zapisali izrazje, ki se v okviru spopadanja z epidemijo znotraj meja Slovenije ni uporabljalo oz. je govorcem slovenščine, ki niso seznanjeni s slovenščino v jezikovnem stiku med italijanščino in slovenščino, (povsem natančen) pomen neznan. Podčrtali smo izrazje oz. sintagme, pri katerih je pomen sicer razumljiv oziroma ulovljiv, vendar gre za sintagme, ki v centralnem standardu slovenščine niso (povsem) stalne.

Kot lahko vidimo, je v korpusu *Primorskega dnevnika* precej dobro zastopana leksika, vezana na potrdilo, ki se v italijanskem sistemu imenuje *green pass*. Izraz je dobesedno prenesen iz italijanščine (ne moremo reči, da prihaja iz angleščine, saj v angleščini ta sintagma ne pomeni ničesar specifičnega, sploh pa ne v povezavi s covidom-19), ta zveza pa se v korpusu *Primorskega dnevnika* pojavlja zelo pogosto (471-krat). Le 16-krat se sicer pojavi zveza *super green pass*, so pa razširjeni prevedki obeh zvez, in sicer v korpusu najdemo *zeleno potrdilo* (795-krat), *covidno potrdilo* (1890-krat) in *okrepljeno covidno potrdilo* (30-krat). Medtem ko lahko sklepamo, da bi govorniki znotraj meja Slovenije kljub neustaljenosti razumeli, kaj pomeni *covidno potrdilo* (saj je to tudi termin, ki se uporablja za EU-certifikate), lahko utemeljeno sklepamo, da bi bili brez poznavanja aktualnega italijanskega sistema pri izrazih *zeleno potrdilo*, *green pass*, *super green pass* in tudi *okrepljeno covidno potrdilo* povsem nemočni. Izstopa tudi hibridna raba, kot se izkazuje pri primeru *covidno PC-potrdilo*, ki se v korpusu pojavi 61-krat in pri katerem gre za mešanje standardnega termina (*PC*) z opisnima iz okolja jezikovnega stika med slovenščino in italijanščino (*covidno* in *potrdilo*). Zanimivo je, da kljub pojavnosti covidnega PC-potrdila v korpusu ne najdemo *covidnega PCT-potrdila*.

Standardni izrazi, vezani na status in potrjevanje, ki so (bili) v rabi v centralnem standardu (*PC-pogoj*, *PCT-pogoj*, *pogoj PCT*, *pogoj PC*), so v korpusu *Primorskega dnevnika* sicer prisotni, vendar v bistveno manjši meri kot drugi, zaznamovani izrazi.

V nadaljevanju analiziramo besedje, vezano na testiranje. Frekvence v korpusu podaja Tabela 7.

Kot lahko vidimo, sta daleč najpogostejša izraza *test* in *bris*, najdemo pa tudi tri primere *tampona*, ki je seveda v centralnem standardu v tem pomenu neobstoječ. Pri tem je treba izpostaviti, da smo ročno pregledali vse konkordance (skupno 34) in izločili vse druge pomene izraza *tampon*. V centralnem standardu je bistveno bolj razširjen izraz *test*, medtem ko izraz *bris* označuje sam tehnični vidik testiranja in se torej pri poročanju o covidu-19 skorajda ne uporablja. Glede na podatke korpusa *Primorskega dnevnika* imamo torej na področju jezikovnega stika med slovenščino in italijanščino dva konkurenčna izraza, enega vezanega na centralni standard slovenščine (*test*) in enega, ki je značilen za območno rabo (*bris*). Da sta na območju jezikovnega stika v veliki meri medsebojno zamenljiva, kaže Tabela 8, ki ponazarja najpogostejše

Tabela 5: Relativne frekvence za najpogostejših pet covidnih leksemov v Primorskem dnevniku in Gigafidi 2.0.

Izraz	Relativna frekvenca – PD	Frekvenca na milijon besed – PD	Relativna frekvenca – GF	Frekvenca na milijon besed – GF
ukrep	0,05350 %	535,03	0,01780 %	178,04
maska	0,02289 %	228,87	0,002663 %	26,63
širjenje	0,02290 %	228,96	0,002060 %	20,6
obvezen	0,02236 %	223,61	0,005233 %	52,33
zdravje	0,01791 %	179,06	0,01296 %	129,6

kolokatorje za oba izraza (izbrali smo prilastke in glagole, ki se najpogosteje vežejo z obema izrazoma).

Kot ponazarja Tabela 8, sta oba izraza v veliki meri zamenljiva, razlikujeta pa se predvsem pri glagolih, zlasti z vidika, da lahko *test* pokažemo oziroma potrdimo, medtem ko se *bris* uporablja predvsem v smislu kliničnega odvzema in poznejše analize. Da bi videli, ali je med obema konkurenčnima izrazoma kakšna razlika v času uporabe oziroma ali je prišlo v obdobju 2020–2021 do razlik v rabi, analiziramo distribucijo obeh izrazov v korpusu. Distribucija obeh izrazov je podana na Sliki 2, slika pa ponazarja kronološko zgradbo korpusa. To pomeni, da položaj pri izhodišču koordinat, povsem levo, ponazarja 1. številko *Primorskega dnevnika* v letu 2020, položaj pri 100 %, povsem desno, pa predstavlja zadnjo številko *Primorskega dnevnika* v letu 2021.

Kot prikazujeta zgornja grafa, je raba izraza *bris* (zgornji graf) izrazito narasla pomladi 2020, nato pa je sledil še en izrazit porast proti koncu leta 2020, nato pa se je raba postopoma začela umikati izrazu *test*, ki kaže bolj razpršeno rabo skozi oba letnika (ob upoštevanju dejstva, da je *test* seveda splošnejša beseda).

V nadaljevanju se osredotočamo na področji poučevanja in dela, ki sta (bili) zaradi epidemije močno zaznamovani. Lekseme, vezane na obe področji, podajamo v Tabeli 9.

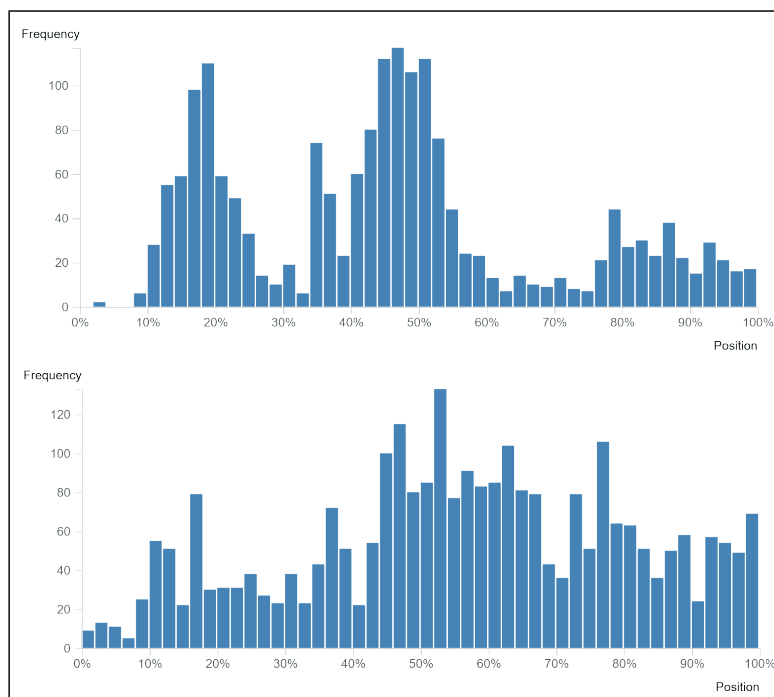
Tabela 9 sopolstavlja izrazje, ki smo ga identificirali kot specifično za korpus *Primorskega dnevnika* (pisano krepko), ter izrazje, ki je v rabi v centralnem standardu slovenščine. Kot še posebno zanimiv se je izkazal izraz *telematski*, ki se v korpusu uporablja v zvezi s številnimi storitvami na daljavo, medtem ko se v centralnem standardu skorajda ne uporablja, razen v navezavi na dejansko telematiko. Zanimiv je tudi izraz *smart working* (sic), ki se uporablja v točno tej obliki, prevzeti iz italijanščine – podobno kot pri izrazu *green pass* ne moremo reči, da je prevzet iz angleščine, saj ta leksem v angleščini ne pomeni ničesar specifičnega oziroma vsaj ne tistega, kar pomeni v italijanščini. Izraz je zanimiv tudi z vidika strukture, saj sta tovrstni zapis in raba (pisano narazen in *working* v obliki gerundija) tudi izrazito nestandardna v

Tabela 6: Področno covidno izrazje v Primorskem dnevniku (status in potrjevanje).

Izraz	Frekvenca (Primorski dnevnik)
covidno potrdilo	1890
zeleno potrdilo	795
green pass	471
pogoj PCT	303
PCT-potrdilo	128
PC-potrdilo	67
PCT-pogoj	64
covidno PC-potrdilo	61
pogoj PC	59
okrepljeno covidno potrdilo	30
potrdilo PCT	24
super green pass	16
PC-pogoj	16
potrdilo PC	2
covidno PCT-potrdilo	0

angleščini, sicer običajnih zapisov, ki bi bili standardni, pa v korpusu ne najdemo (npr. *smartwork*, *smart work* ipd.). Treba pa je poudariti, da sta oba izraza rabljena izrazito redkeje kot povsem standardna *delo od doma* in *delo na daljavo*.

Pri izrazih, pri katerih lahko pomen razberemo, vendar raba v centralnem standardu ni ustaljena, lahko izpostavimo predvsem *didaktiko na daljavo* in *pouk v prisotnosti*, ki oba kažeta močan vpliv italijanščine, sta pa razmeroma redko rabljena glede na prevladujoča in standardna izraza *pouk na daljavo* in *pouk v razredu/v živo*.



Slika 2: Distribucija izrazov bris (zgornji graf) in test (spodnji graf) v korpusu Primorskega dnevnika.

V naslednjem odseku analiziramo covidno leksiko glede na leto objave, tj. primerjamo podkorpusa besedil *Primorskega dnevnika* iz let 2020 in 2021.

Leksemske razlike med 2020 in 2021

Analizo leksemskih razlik najprej izvedemo na primeru enobesednih leksemov, nato pa se osredotočimo na večbesedne. Pri tem uporabljamo podkorpus iz leta 2021 kot fokus, podkorpus iz leta 2020 pa kot referenčno zbirko podatkov. Enobesedne lekseme podaja Tabela 10.

Tabela 10 prikazuje najbolj značilno covidno besedje v *Primorskem dnevniku* v letu 2021 v primerjavi s predhodnim letnikom. Pri tem moramo izpostaviti, da je covid-19 daleč najpogostejša tematika, na katero se navezuje leksika, edini temi, ki ju lahko med najbolj pogostim besedjem iz leta 2021 še razberemo, se nanašata na županske volitve in olimpijske igre.

Kot lahko vidimo, je bilo v letu 2021 v *Primorskem dnevniku* bistveno več napisanega o cepivih in potrdilih, saj sta bili to dva glavna vidika covidne realnosti v letu 2021. V Tabeli 11 podajamo še večbesedne termine, pri katerih pa vključujemo tudi izrazje drugih pomenskih polj, tako da lahko iz distribucije in frekvenc razberemo pomembnost/razširjenost določene teme v letih 2020 in 2021.

Najbolj značilni v 2021 v primerjavi z 2020 sta sintagmi *covidno potrdilo* in *zeleno potrdilo*, ki se v letu 2020 še nista pojavili v *Primorskem dnevniku*, na dnu

seznama 25 najpogostejših najdemo tudi *obvezno covidno potrdilo*. Podobno kot pri enobesednih leksemih lahko vidimo, da se večina primerov nanaša na epidemijo covid-19 in volitve, pri drugih pa gre povečini za sistemske dele časopisa, ki so formulacijski (npr. *indeks sredi dneva*, ki se nanaša na vreme).

RAZPRAVA

Analiza korpusa besedil, ki so izšla v *Primorskem dnevniku* med letoma 2020 in 2021 in ki obravnavajo epidemijo covid-19, potrjuje hipotezo, da prihaja med govorce in govorkami slovenskega jezika v Italiji do nekaterih pojavov, ki jih po vrsti in količini ne moremo pripisati zgolj običajnim dejavnikom, ki vplivajo na jezike v manjšinskem položaju (Meecham & Rees-Miller, 2001, 539; Fought, 2006, 31).

Jezikovno (in specifično leksikalno) variantnost običajno pričakujemo pri manj formalnih sporazumevalnih kodih, pri referiranju na različno predmetnost/pojavnost in, seveda, pri jezikovnih različicah (idiomih), ki jih sami govorce in govorce prepoznavajo kot lokalne/regionalne (Fought, 2006, 37). Korpusna analiza pa je pokazala, da je leksikalna variantnost v besedilih, ki nastajajo na območju poselitve slovenske skupnosti v Italiji, znatna tudi takrat, ko gre za strokovno terminologijo (in bi torej pričakovali večjo doslednost), za ubesedovanje predmetnosti/pojavnosti, ki se na obeh straneh meje ne razlikuje, ter za besedila, ki jih govorce in govorce dojemajo kot standardna.

Tabela 9: Področna covidna leksika v Primorskem dnevniku (poučevanje in delo).

Izraz	Frekvenca (Primorski dnevnik)
didaktika na daljavo	23
poučevanje na daljavo	16
pouk na daljavo	381
pouk v razredu	20
pouk v prisotnosti	12
pouk v živo	69
pouk v učilnici	5
pouk v predavalnici	0
delo od doma	98
delo na daljavo	177
smartwork	0
smart working	25
telematski	41

Vprašani, ki si ju na tej točki moramo zastaviti, sta dve:

1. kaj povzroča tolikšno variantnost?
2. katere so posledice take variantnosti?

V tem prispevku se bova osredotočila na prvo vprašanje.

Da ima slovenščina v Italiji nekatere samosvoje poteze, so raziskovalke in raziskovalci v glavnem pripisovali stiku z romanskimi jeziki (Jagodic et al., 2017; Grgič, 2016b) in drugim dejavnikom, predvsem geografski obrobni, specifični politični zgodovini in prisotnosti državnih meja (Jagodic et al., 2017). Vendar pa se v danem primeru (tj. covidni terminologiji v letih 2020 in 2021) zdi, da nobeden od teh dejavnikov ne more pojasniti tolikšnega odstopanja, kot ga izkazujejo korpusni podatki. Covidna terminologija se je v vse jezike prevajala istočasno in to iz angleščine, kar pomeni, da so bili italijanski in slovenski termini novinarjem na razpolago istočasno. Prav tako ni mogoče reči, da bi na terminologijo lahko vplivala specifična politična zgodovina, saj je epidemija koronavirusa globalen in nov pojav. Tudi geografska obrobni in prisotnost državne meja nimata več vloge, ki sta jo imeli v preteklosti – k temu sta pripomogla padec notranjih meja znotraj schengenskega območja in (kar je za dani primer še pomembnejše) razvoj tehnologije, ki danes omogoča hitro in neposredno spremljanje novic na svetovnem spletu in družbenih omrežjih.

Tabela 10: Primerjava najpogostejših enobesednih leksemov med letnikoma 2020 in 2021.

Izraz	Frekvenca (2021)	Frekvenca (2020)
potrdilo	3709	462
cepljenje	1997	458
odmerek	1349	292
obvezen	2790	2018
cepivo	1935	1063
covidno	920	1
zelen	2510	1966
cepljen	878	26
cepiti	956	177

Analiza, opravljena za slovenščino v Italiji, je klub temu pokazala, da se je terminologija, povezana s koronavirusom, v veliki meri prevajala iz italijanščine in se nato na različne načine in do različnih stopenj prilagajala slovenskemu standardu, kar smo doslej opazovali že na drugih besedilih, ki pa so izkazovala drugačne značilnosti (Grgič, 2016b in 2020). Lahko torej rečemo, da pojav ni odvisen od zunanjih dejavnikov (meje, zgodovinskih okoliščin, vpliva italijanščine), ampak od samih procesov, na katerih temelji nastajanje besedil, in pa od dinamik oblikovanja skupnosti skozi rabo jezika v sporazumevalnih praksah (Eckert, 2000; Meyerhoff, 2008). Vprašanje je seveda, v kolikšni meri ti procesi temeljijo na zavestnih odločitvah govorcev in govork in v kolikšni meri gre za usvojene mehanizme, ki so se že povsem uveljavili v skupnosti in niso več predmet sprotnega odločanja.

Količina in vrsta odstopanj od splošnega standarda ter postopno prilagajanje standardu potrjujejo hipotezo, da gre za ne povsem zavestne odločitve govorcev in govork oz. za spontane mehanizme upovedovanja, ki so na ravni skupnosti (torej tudi bralk in bralcev) sicer običajni. V prid hipotezi, da vendarle prihaja do naknadne zaznave odstopanja (ali na ravni piscev ali na ravni bralcev), govori dejstvo, da se terminologija s časom približa splošnemu standardu. To vidimo na nekaterih primerih, kot so *tampon* → *bris/test* (kjer sta termina *bris* in *test* zelo hitro izpodrinila paronim *tampon* – it. *tampone*), *didaktika na daljavo* → *poučevanje/pouk* na daljavo (*poučevanje* in *pouk* sta nadomestila paronim *didaktika* – it. *didattica*), *pouk v prisotnosti* → *pouk v razredu/v živo* (besedni zvezi *v razredu* in *v živo* se uporabljata namesto kalka *v prisotnosti* – it. *in presenza*). Drugi primeri kažejo na delne prilagoditve (*green pass* → *zeleno potrdilo* → *covidno potrdilo* vs. *PCT pogoj*): taki primeri so po-

Tabela 11: Primerjava 25 najpogostejših večbesednih leksemov med letnikoma 2020 in 2021.

Termin	Frekvenca (2021)	Frekvenca (2020)
covidno potrdilo	1890	0
zeleno potrdilo	795	0
županski kandidat	592	46
covidni bolnik	402	64
odmerek cepiva	404	85
rajonski svet	653	404
sprejemati po telefonu	716	538
cepilni center	283	7
število mest	524	366
starostna skupina	274	55
krajina po deželnem zakonu	230	1
prestonica kulture	500	355
evropska prestolnica kulture	482	333
jezik s podnapisi	250	39
evropska prestolnica	498	365
imate v rokah	230	25
tiskarno sinoči	230	25
kulturni center	1020	1056
izbranka tedna	199	0
zaprt na voljo	230	44
indeks sredi dneva	194	0
desna sredina	316	158
izvirni jezik	224	42
izvirni jezik s podnapisi	209	26
obvezno covidno potrdilo	169	0

gostejši, ko gre za razlike v predmetnosti med obema državama, Italijo in Slovenijo. Analiza kolokacij pa na drugi strani pokaže, da so te tudi po prilagoditvi terminologije neznačilne in da odstopajo od standarda: tak primer je kolokacija *narediti tampon* → *narediti/odvzeti bris* vs. *testirati se*.

Količine in vrste variantnosti, ki jo izkazujejo analizirana besedila, ne moremo pojasniti s sodobnimi družbenimi dinamikami čezmejne narodne manjšine, ki živi ob odprti notranji schengenski meji, razpolaga z digitalnimi tehnologijami in dostopa do spletnih vsebin. Stopnja variantnosti, ki jo opažamo v teh besedilih, je značilna bolj za jezikovni otok (t. i. jezikovno enklavo, prim. Bauer et al., 2005) in nakazuje na dejstvo, da se je v desetletjih pred padcem meje in razvojem novih tehnologij med slovensko skupnostjo v Italiji, ki je res živela dokaj ločeno od referenčne skupnosti v Sloveniji, razvila tesna, nepropustna družbena mreža (Milroy, 2002). Ta je bila hkrati dovolj obsežna in dinamična, da je omogočila srednjeročno ohranjanje jezika, a tudi dovolj nepropustna in tesna, da je spodbudila razvoj vzporednega jezikovnega standarda, ki se je ob pomanjkanju stikov z drugimi kodi slovenskega jezikovnega kontinuuma uveljavil kot edini standard. Tako so se razvile sporazumevalne prakse in jezikovne rabe, ki so izrazito lokalne, a se zaradi pomanjkanja izpostavljenosti drugim rabam znotraj istega jezikovnega kontinuuma dojemajo kot (edine) standardne. V primeru skupnosti govork in govorcev slovenščine v Italiji ne govorimo torej o večstandardnosti (torej o sočasni prisotnosti več standardov istega jezikovnega kontinuuma), ampak o vzporedni enostandardnosti: lokalni standard je v nekaterih primerih nadomestil osrednjeslovenskega oz. se uveljavil kot edini standard, ki ga govorcev in govorcev uporablja v svojih sporazumevalnih praksah oz. do katerega sploh dostopajo.¹¹

Gornji tezi, ki izhajata iz analize izbranega korpusa, potrjujejo tudi nekatere druge doslej opravljene raziskave. Te so že pokazale, da se je po 2. svetovni vojni predvsem v Trstu in Gorici, delno pa tudi na Videmskem, razvila bogata dejavnost, ki je poleg šol in društev vključevala tudi številna podjetja (Guardiancich, 2021) in s tem ustvarila okoliščine za usvajanje in rabo slovenskega jezika. Prav zaradi svojega obsega in strukture se je ta mreža zdela dovolj učinkovita za zagotavljanje jezikovne izpostavljenosti (Brezigar et al., 2021). Veliki premiki, do katerih je ob prehodu med 20. in 21. stoletjem prišlo na več ravneh,¹² pa so kmalu pokazali na pomanjkljivosti tako strukturiranih mrež in njihovih premis. Lokalno in regionalno samozadostnost, na katerih je temeljila družbena mreža slovenske skupnosti v Italiji, sta neposredno podpirali tudi eksplicitna jezikovna politika obeh

11 Čeprav obravnava posledic jezikovnega separatizma presega omejitve tega prispevka, naj na tem mestu le omeniva sovplivanje jezikovnih rab in sporazumevalnih na oblikovanje in percepcijo identitet posameznikov in skupnosti (Pertot, 2017) ter učinke okrnjene izpostavljenosti na sporazumevalno zmožnost ter s tem na opolnomočenje govork in govorcev (Grgič, 2016a).

12 Naj tu navedemo le nekatere: na meddržavni ravni osamosvojitve Slovenije, njen vstop v EU in padec meje; na globalni ravni globalizacijo in posledično regionalizacijo, hiter razvoj tehnologije s 4. industrijsko revolucijo, ublagovljenje jezikov, fluidnost identitet itd.; na lokalni ravni propad t. i. družbenega gospodarstva, sprejem t. i. zaščitne zakonodaje, spremenjen odnos večinskega prebivalstva do manjšinskega jezika itd. (Heller, 2010).

držav, Italije in Slovenije, in implicitna jezikovna politika same skupnosti (Brezigar et al., 2021), ki sta naslavljali predvsem zaščito in ohranjanje jezika kot tradicionalne vrednote manjšinske skupnosti, nista pa spodbujali njegovega razvoja vzporedno z razvojem celotnega slovenskega jezikovnega kontinuuma. Raziskave o stikih med slovenskima skupnostma na obeh straneh meje (Brezigar & Vidau, 2021) danes kažejo, da so ti sicer pogosti, ne pa zadostni in ustrezni, da bi zagotovili učinkovito izpostavljenost govorcev in govork, ki živijo v Italiji, različnim kodom slovenskega jezikovnega kontinuuma. Meja je zlasti nepropustna na področju medijskih vsebin in družbenih omrežij,¹³ kar je glede na značilnosti sodobnih kanalov (splet, telefon, družbena omrežja ...) še posebej presenetljivo. Prav ta podatek nam kaže na dejstvo, da jezikovni separatizem, ki smo mu priča med slovensko skupnostjo v Italiji, ni (več) posledica zunanjih okoliščin (meje, zgodovinskih danosti ...), ampak predvsem (samo)marginalizacije in (samo)izključevanja govorcev in govork iz širšega prostora, v katerem je slovenski jezikovni kontinuum prevalentni sporazumevalni kod, s perpetuiranjem predvsem lokalnih družbenih mrež in identitet (Pertot, 2017).

SKLEP

Korpusna analiza potrjuje hipotezo, da je regionalni jezikovni kod, ki se je razvil med skupnostjo govorcev in govork slovenskega jezika v Italiji, že postal prevladujoči standard in zavzel področja, na katerih bi v jezikovnih rabi pričakovali nadregionalno kontinuiteto. Stopnja variantnosti, torej količina in vrsta odstopanj od splošnega standarda, kaže na dejstvo, da gre v primeru slovenske skupnosti v Italiji za razširjen pojav jezikovnega separatizma. Ta pojav, katerega razsežnosti spominjajo prej na jezikovni otok kot na čezmejno narodno manjšino v Evropi 21. stoletja, lahko razumemo v luči sociolingvističnih dinamik, ki so skupnost zaznamovale v drugi polovici 20. stoletja (oblikovanje močne lokalne družbene mreže, pomanjkanje stikov z referenčnim prostorom, sprejetje konservativnih jezikovnih politik) in ki se ob pomanjkanju ustreznega jezikovnega načrtovanja v 21. stoletju le še nadaljujejo in stopnjujejo ter s tem vplivajo na identitetne opcije posameznikov in posameznic, na sporazumevalno zmožnost in opolnomočenje govork in govorcev ter, srednje-ročno, na jezikovno vitalnost celotne skupnosti.

¹³ Podatek izhaja iz raziskave, ki jo je v sklopu programske skupine Razsežnosti slovenstva med lokalnim in globalnim v začetku tretjega tisočletja (P5-0409) izvedla Matejka Grgič.

PROCESSES OF LINGUISTIC SEPARATISM IN CROSS-BORDER LINGUISTIC MINORITIES:
ADOPTION, ADAPTATION, AND TRANSLATION OF COVID TERMINOLOGY AMONG
SLOVENIANS IN ITALY

Matejka GRGIČ

University of Ljubljana, Faculty of Arts, Aškerčeva 2, 1000, Ljubljana, Slovenia
SLORI – Slovene Research Institute, Ul. Beccaria 6, 34133, Trst, Italy
e-mail: matejka.grgic@ff.uni-lj.si

Damjan POPIČ

University of Ljubljana, Faculty of Arts, Aškerčeva 2, 1000, Ljubljana, Slovenia
e-mail: damjan.popic@ff.uni-lj.si

SUMMARY

The paper reports on the results of a comparative corpus analysis of texts published in Primorski dnevnik, the Slovenian daily published in Italy, in 2020 and 2021. We contrast these texts with information from Slovenian corpora and language resources. The research is focused on the lexis connected with the Covid-19 epidemic as this topic brings about a number of specifics such as rapid genesis and determinologization, single- or self-referentiality, using loanwords (from English), etc. The research is based on corpus-linguistic survey of 618 issues (two volumes) of the Trieste daily newspaper Primorski dnevnik. We used this body of texts to create a corpus that can be used to discern the differences in the use of lexemes in different datasets. The aim of the paper is to demonstrate the level of difference in the lexis of Standard Slovene on both sides of the border. For this reason, we chose to employ corpus methodology as it allows us to compare large bodies of data. The analysis confirmed our initial hypothesis that there is a high level of variation in language use in this region and that this variation cannot be attributed to the factors that normally explain the deviations from standard language, the effects of languages in contact and minority languages.

Keywords: Slovenians in Italy, minority languages, terminology, covid-19, sociolinguistics, corpus linguistics

VIRI IN LITERATURA

- Bauer, Thomas, Epstein, Gil S. & Ira N. Gang (2005):** Enclaves, Language, and the Location Choice of Migrants. *Journal of Population Economics*, 18, 4, 649–662.
- Bogatec, Norina (2015):** Šolanje v slovenskem jeziku v Italiji. *Razprave in gradivo: revija za narodnostna vprašanja*, 74, 5–21.
- Bogatec, Norina (2018):** Kdo v Italiji obiskuje vrtce in šole v slovenskem jeziku. V: Žele, Andreja & Matej Šekli (ur.): *Slovenistika in slavistika v zamejstvu*. Ljubljana, Zveza društev Slavistično društvo Slovenije, 125–129.
- Brezigar, Sara (2017):** Slovenska skupnost v Italiji med preteklostjo in prihodnostjo. V: Bogatec, Norina & Zaira Vidau (ur.): *Skupnost v središču Evrope: Slovenci v Italiji od padca Berlinskega zidu do izzivov tretjega tisočletja*. Trst, ZTT & Slori, 22–30.
- Brezigar, Sara & Zaira Vidau (2021):** Mladi govorci slovenskega jezika v Italiji in njihov odnos do Republike Slovenije. *Razprave in gradivo: revija za narodnostna vprašanja*, 87, 87–106.
- Brezigar, Sara, Grgič, Matejka & Devan Jagodic (2021):** Koncept deželne jezikovne politike za slovenščino: teoretska izhodišča, cilji, področja ukrepanja in institucionalni okvir. Tretja deželna konferenca o varstvu slovenske jezikovne manjšine. V: Jagodic, Devan (ur.): *Tretja deželna konferenca o varstvu slovenske jezikovne manjšine*, Trst, 12. in 19. november 2021. Trst, Regione autonoma Friuli Venezia Giulia, Consiglio regionale, 133–158.
- Cavaion, Irina M. (2014):** Testing Feasibility of Cross Border Contacts within Primary Neighbouring Languages Classroom. *Annales, Series Historia et Sociologia*, 24, 2, 277–292.
- Chambers, Jack K. (2003):** *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford, Blackwell.
- Davila, Bethany (2016):** The Inevitability of 'Standard' English: Discursive Constructions of Standard Language Ideologies. *Written Communication*, 33, 2, 127–148.
- Eckert, Penelope (2000):** *Linguistic Variation as Social Practice*. Oxford, Blackwell Publishers
- Fought, Carmen (2006):** *Language and Ethnicity*. Cambridge, Cambridge University Press.
- Grgič, Matejka (2016a):** Jezik: sistem, sredstvo in simbol. Identiteta in ideologija med Slovenci v Italiji. Trst, Slovenski raziskovalni inštitut SLORI.
- Grgič, Matejka (2016b):** Lo sloveno in Italia: fenomeni di contatto linguistico tra pragmatica, percezione e ideologia. *Ricerche slavistiche*, 14=60, 387–415.
- Grgič, Matejka (2017):** Italijansko-slovenski jezikovni stik med ideologijo in pragmatiko. *Jezik in slovstvo*, 62, 1, 89–98.
- Grgič, Matejka (2020):** Terminology as a Cultural Project: The Consequences of the Double Standard in a Minority Context. V: Mikolič, Vesna (ur.). *Language and Culture in the Intercultural World*. Newcastle upon Tyne, Cambridge Scholars Publishing, 102–124.
- Guardiancich, Igor (2021):** When Political Rights Do Not Translate into Economic Power: The Rise and Fall of the Slovenian Minority's Economy in Italy (1954–2020). *Nationalities Papers*, 1–18.
- Heller, Monica (2010):** The Commodification of Language. *Annual Review of Anthropology*, 39, 1, 101–114.
- Jagodic, Devan (2017):** Slovenci v Italiji: poselitveni prostor in demografsko gibanje. V: Bogatec, Norina & Zaira Vidau (ur.): *Skupnost v središču Evrope: Slovenci v Italiji od padca Berlinskega zidu do izzivov tretjega tisočletja*. Trst, ZTT & Slori, 40–49.
- Jagodic, Devan, Kaučič Baša, Majda & Roberto Dapit (2017):** Jezikovni položaj Slovencev v Italiji. V: Bogatec, Norina & Zaira Vidau (ur.): *Skupnost v središču Evrope: Slovenci v Italiji od padca Berlinskega zidu do izzivov tretjega tisočletja*. Trst, ZTT & Slori, 66–88.
- Meecham, Marjory & Janie Rees-Miller (2001):** Language in Social Contexts. V: O'Grady, William, Archibald, John, Aronoff, Mark & Janie Rees-Miller (ur.): *Contemporary Linguistics*. Boston, Bedford/St. Martin's, 537–590.
- Melinc Mlekuž, Maja (2019):** Sporazumevalna zmožnost v šolah s slovenskim učnim jezikom v Italiji. *Razprave in gradivo: revija za narodnostna vprašanja*, 83, 67–82.
- Meyerhoff, Miriam (2008):** Communities of Practice. V: Chambers, Jack K., Trudgill, Peter & Natalie Schilling-Estes (ur.): *The Handbook of Language Variation and Change*. Oxford, Blackwell Publishing, 526–548.
- Mezgec, Maja (2016):** Linguistic Landscape as a Mirror: The Case of the Slovene Minority in Italy. *Razprave in gradivo: revija za narodnostna vprašanja*, 77, 67–85.
- Milroy, Lesley (2002):** Social Networks. V: Chambers, Jack K., Trudgill, Peter & Natalie Schilling-Estes (ur.): *The Handbook of Language Variation and Change*. Oxford, Blackwell Publishing, 549–572.
- Pertot, Susanna (2017):** Identitetne prigode Slovencev v Italiji. V: Bogatec, Norina & Zaira Vidau (ur.): *Skupnost v središču Evrope: Slovenci v Italiji od padca Berlinskega zidu do izzivov tretjega tisočletja*. Trst, ZTT & Slori, 89–101.
- Pollak, Senja, Arhar Holdt, Špela, Krek, Simon & Marko Robnik Šikonja (2020):** Reference List of Slovene Frequent Common Words. Slovenian language resource repository CLARIN.SI, <http://hdl.handle.net/11356/1346> (zadnji dostop: 2022-12-22).

Riman, Barbara (2022): Slovenska skupnost na Hrvaškem v obmejnem območju in meja v času covid-19. *Annales, Series Historia et Sociologia*, 32, 2, 165–178.

SketchEngine (2022). <https://www.sketchengine.eu/> (zadnji dostop: 2022-12-22).

Sledilnik 1.0 (2022). <https://viri.cjvt.si/sledilnik/slv/> (zadnji dostop: 2022-12-22).