

TRANSKRIBIRANJE V SOCIOLINGVISTIČNIH RAZISKAVAH IN KORPUSIH GOVORJENE SLOVENŠČINE

MAJA BITENC

Univerza v Ljubljani, Filozofska fakulteta, Ljubljana, Slovenija
Maja.Bitenc@ff.uni-lj.si

Prispevek obravnava transkribiranje kot kompleksen, interpretativen in selektiven proces, pri katerem pretvarjanje iz govorenega v zapisano besedilo odpira številna temeljna in praktična vprašanja. V sociolingvističnih raziskavah ni standardnih transkripcijskih načel, temveč so ta vedno odvisna od teoretske usmeritve raziskovalca, predmeta in namena raziskave. Uvodni del predstavlja splošne teoretske pomisleke in načelne smernice, osrednji del pa se osredotoča na različne prakse transkribiranja v korpusih in raziskavah govorne slovenščine. Posebna pozornost je posvečena problematiki zapisovanja variant fonema /v/, polglasnika, specifičnih narečnih glasov in premen po zvonečnosti. Med drugim so opisane različne prakse transkribiranja v avtoričinih sociolingvističnih raziskavah variantnosti govorne slovenščine in relevantnih socialnopsiholoških tem, od začetnih natančnih fonetičnih transkripcij do različnih stopenj poenostavljanja in prehajanja k ortografskemu zapisu s posebnimi znaki za posamezne glasove. Prispevek izpostavlja tudi izzive glede razločevanja posameznih glasov in nedoslednosti pri njihovem zapisovanju.

DOI
[https://doi.org/
10.18690/um.ff.4.2024.8](https://doi.org/10.18690/um.ff.4.2024.8)

ISBN
978-961-286-882-6

Ključne besede:
transkribiranje,
fonetična in ortografska
načela,
sociolingvistika,
jezikovna variantnost,
fonem /v/,
polglasnik



Univerzitetna založba
Univerze v Mariboru

DOI
[https://doi.org/
10.18690/um.ff.4.2024.8](https://doi.org/10.18690/um.ff.4.2024.8)

ISBN
978-961-286-882-6

Keywords:
transcription,
phonetic and orthographic
conventions,
sociolinguistics,
language variation,
phoneme /v/,
semi-vowel

TRANSCRIPTION IN SOCIOLINGUISTIC RESEARCH AND CORPORA OF SPOKEN SLOVENE

MAJA BITENC

University of Ljubljana, Faculty of Arts, Ljubljana, Slovenia
Maja.Bitenc@ff.uni-lj.si

The article discusses transcription as a complex, interpretive, and selective process in which the transformation of spoken into written text raises numerous fundamental and practical questions. There are no standard transcription principles in sociolinguistic research; they always depend on the researcher's theoretical orientation, subject matter, and research purpose. The introductory part of the article presents theoretical considerations and general guidelines, while the central part focuses on various transcription practices in corpora and studies of spoken Slovene. Particular attention is paid to the notation of the variants of the /v/ phoneme, the semi-vowel, specific dialectal sounds and changes in sonority. Among other things, various transcription practices in the author's sociolinguistic studies of variation in spoken Slovene and relevant socio-psychological topics are described, ranging from initial precise phonetic transcriptions through various levels of simplification to the transition to orthographic representation with special symbols for individual sounds. The article also highlights challenges in distinguishing individual sounds and inconsistencies in their representation.



1 Uvod¹

Transkribiranje oz. pretvarjanje iz govornega v zapisano besedilo je ključen del obdelave gradiva pri vsaki raziskavi govornega jezika, glede katerega je treba odgovoriti na številna temeljna in praktična vprašanja. Pomembno je, da so odločitve, ki jih sprejme raziskovalec, konsistentne, zapisovanje pa natančno, saj transkripcije predstavljajo osnovo tako za nadaljnjo analizo kot interpretacijo podatkov (Tagliamonte 2007; Kvale in Brinkmann 2009: 177–187). Analiza je lahko osnovana tudi na prvotnih virih, torej avdio posnetkih, h katerim se ob transkripcijah vračamo med celotnim analitičnim procesom; v tem primeru da so transkripcije predvsem način predstavitve podatkov bralcem (Lampropoulou 2012: 76–77).



Slika 1 Shema o izgubi informacij pri pretvorbi žive pripovedi v zapis
(graf Barbara Ivančič Kutin, ilustracije Ciril Horjak)

Vir: Ivančič Kutin 2011: 64

¹ Prispevek je nastal v okviru projekta Sociolingvistična variantnost govorne slovenščine na primeru mobilnih govorcev: Viri, metode in analiza (Z6-9371). Avtorica se Javni agenciji za raziskovalno dejavnost Republike Slovenije zahvaljuje za financiranje, Darinki Verdonik, Heleni Dobrovoljc, Karmen Kenda-Jež, Luku Horjaku in Marku Stabeju pa za razprave glede transkripcijskih načel.

Zavedati se je treba, da gre pri transkripciji vedno za interpretacijo, saj nikoli ne more biti povsem zanesljiva predstavitev originalne interakcije. Že z avdio snemanjem se izgubi vrsta informacij, kot so telesna govorica, mimika in gestikulacija, ob zapisu govora pa je nemogoče v celoti opisati ton glasu, intonacijo, dihanje ipd. Obenem je vsaka transkripcija vzeta iz konteksta, saj ne vključuje časovnih in prostorskih dimenzij žive komunikacije (Kvale in Brinkmann 2009: 177–178). Slika 1 nazorno prikazuje izgubo informacij pri pretvorbi žive pripovedi v zapis.

Dejstvo je torej, da pri pretvarjanju besedila iz posnetka v zapis ni mogoče vsega istočasno, izčrpno in neprekinjeno zabeležiti in da se mora raziskovalec vedno odločiti, koliko in katere vrste informacij zapisovati (Brinker in Sager 1996: 35). Ne obstajajo standardna transkripcijska načela, pač pa so ta odvisna od teoretske usmeritve raziskovalca, predmeta in namena raziskave, torej kaj in zakaj se proučuje. Selektivnost je tako tesno povezana s subjektivnostjo. Pomembna dejavnika pri odločanju glede natančnosti zapisa sta tudi čas in finančna sredstva, ki jih imamo na razpolago, saj je transkribiranje zelo zamudno. Okvirna ocena je, da so za prvo transkripcijo enournega posnetka potrebne štiri ure, vendar je to zelo odvisno od kakovosti zvoka, števila govorcev in transkriptorjevega poznavanja varietet(e), ki jih (jo) zapisuje; tako lahko enourni intervju terja delo od enega dneva do enega tedna ali celo več.²

V sociolingvističnih raziskavah se večinoma uporabljajo standardna ortografija in ločila, saj je to ključno za karseda enostavno transkribiranje in berljivost transkripcije. Načeloma se torej besede zapisuje tako, kot so običajno zapisane, razen če obstaja upravičen razlog za drugačno odločitev. Navadno se uporablja tudi standardna ločila, torej pike, vprašaje, vejice itd., kot običajno pri pisanju. Obenem je pomembno, da se transkribira, kar je oseba dejansko rekla, čeprav to ni v skladu s predpisanimi slovničnimi pravili standardnega jezika (Preston 2000; Tagliamonte 2006: 55).

Osnovni vodili pri transkripcijah za potrebe variantnostne analize sta, da naj bodo transkripcije dovolj natančne, da učinkovito ohranijo dovolj podatkov za jezikoslovno analizo, in da naj bodo dovolj preproste, da jih je mogoče sorazmerno lahko brati in zapisovati. Ob upoštevanju časovnih in finančnih omejitev je

² Na zamudnost transkribiranja opozarja veliko avtorjev; Jana Zemljarič Miklavčič (2008: 180–181) npr. piše, da je za transkripcijo dveh minut posnetka povprečno porabila eno uro, pri posnetkih spontanega multiloga pa pogosto ni dosegla niti transkripcijske hitrosti ene minute posnetka v eni uri, pa čeprav je šlo za ortografsko transkripcijo. V tuji raziskavi staranja je izkušena tajnica za transkribiranje ene ure posnetka intervjuja potrebovala približno 5 ur (Kvale in Brinkmann 2009: 180).

ohranjanje ravnotežja med tema dvema ciljema ključno pri oblikovanju korpusa (Tagliamonte 2006: 54). Zapisujemo torej tisto, kar potrebujemo in kar je relevantno za namen sedanje oz. morebitne kasnejše raziskave (Wray in Bloomer 2006: 185).

Glavni izziv pri izdelavi korpusa je zagotoviti karseda avtentično, natančno in dosledno predstavitev posnetkov govornega jezika. Za enotno zapisovanje je potreben transkripcijski protokol, ki kot referenčni dokument za transkripcijsko prakso zagotavlja doslednost znotraj korpusa (Poplack 1989: 434; Tagliamonte 2006: 55).

Za urejanje posnetkov, transkribiranje in obdelavo transkripcij so na voljo različni računalniški programi in orodja. Pogosto uporabljana in prosto dostopna orodja za urejanje posnetkov so Transcriber, WinPitch in Exmaralda, od katerih ima vsak svoje posebnosti. Poslušanje in sorazmerno enostavno urejanje posnetkov omogočata npr. programa Adobe Audition ali brezplačen Audacity. Za poslušanje in natančnejšo analizo krajših izsekov in posameznih glasov je na voljo program Praat. Gradnjo lastnega korpusa s številnimi funkcijami omogoča program Sketch Engine, za korpusno obdelavo transkripcij sta uporabna Notepad++ in PSpad.

V slovenski praksi, tako v dialektologiji kot drugih raziskavah, je še vedno pogosto zapisovanje v programu Word, ki omogoča zapis narečnih glasov z znaki slovenske narečne fonetične oz. Logarjeve transkripcije v vnašalnem sistemu ZRCola, posebej če ni potrebno označevanje prekrivnega govora, obenem pa prav pridejo različne možnosti za oblikovanje besedila in zapisovanje komentarjev. Pri korpusu GOS³ in projektu RSDO⁴ so za transkribiranje uporabili program Transcriber.

2 Transkribiranje v korpusih in raziskavah govornje slovenščine

V okviru dialektologije v slovenskem prostoru je za zapis narečnih govornih besedil od 70. let prejšnjega stoletja v rabi t. i. nova nacionalna oz. Logarjeva transkripcija (Logar 1993: 101–103; Kenda-Jež 1996: VII–XIV in 2011: 27–30). Font z naborom transkripcijskih znakov različnih fonetičnih transkripcij iz slovenske in slovanske zapisovalske tradicije ter mednarodne fonetične abecede IPA (ang.

³ <http://www.korpus-gos.net/Support/About>, dostop 23. 6. 2023.

⁴ <https://rsdo.slovenscina.eu/>, dostop 23. 6. 2023.

International Phonetic Alphabet) je skupaj z vnašalnim sistemom ZRCola⁵ na ZRC SAZU v Ljubljani razvil Peter Weiss. Glede načina zaznamovanja prozodičnih in paralingvističnih prvin ter glasovnih refleksov v dialektoloških študijah prim. Weiss 1990: 181; Kenda-Jež 2002: 166–167.

Pri korpusu GOS so se odločili za t. i. pogovorni zapis, katerega cilj je zvesta predstavitev glasovne podobe govora v karseda berljivi obliki. Uporabljali so veljavni slovenski črkopis in upoštevali odstopa od pravopisne norme, kjer izgovor odstopa od standardne izreke. Transkripcijska načela sta določili Darinka Verdonik in Ana Zwitter Vitez (2011: 58–68).

Tudi pri projektu Razvoj slovenščine v digitalnem okolju (RSDO) govor zapisujejo v skladu s knjižno normo, če izgovor ne odstopa bistveno od zbornega, le polglasnik vedno, tudi v primeru zborne izreke, pišejo s posebnim znakom – @. Pri zapisu besed, ki v izreki odstopajo od zborne ali niso del knjižne leksike, sledijo smernicam, ki sta jih opredelili Darinka Verdonik in Andrejka Bizjak (2022: 25–26).

Za namen proučevanja oblikovanja neknjižnega pogovornega oz. pogovarjalnega jezika na primeru govork iz Mengša in Beltincev, ki obiskujejo srednjo šolo v Ljubljani, je Jožica Škofic posnetke šolskega spraševanja, obnov besedil, medsebojnih pogovorov in intervjujev o jezikovni zavesti uporabljala fonetično transkripcijo po zgledu zapisov v *Slovenskih narečjih* Tineta Logarja (1975 oz. 1993), ni pa označevala tonemov v gorenjskem narečju, saj bi, kot piše, glede na njeno poznavanje lahko prišlo do napak (Škofic 1991: 225–227, transkripcije na str. 227–377).

Pri zapisovanju spontano govornih besedil ljubljanskih govorcev za analizo besedilnih vrst in zgradbe ter sociolingvističnih intervjujev za raziskavo variantnosti govorne slovenščine na primeru šestih mobilnih mladih Korošcev sta se Mojca Smolej in Katja Kotnik odločili za knjižno oz. ortografsko transkripcijo z opredeljenimi izjemami, ki niso skladne s standardnim slovarjem (Smolej 2012: 14–15, zapis besedil na str. 145–236; Kotnik 2022: 48).

⁵ <http://zrcola.zrc-sazu.si/>, dostop 23. 6. 2023.

V raziskavi variantnosti govorne slovenščine pri petih mobilnih govornicah z Idrijskega, ki se šolajo ali delajo v Ljubljani, sem Maja Bitenc (2016a, 2016b) oblikovala več različnih transkripcijskih protokolov. V začetni fazi sem se predvsem zaradi posebnosti in neraziskanosti slovenskih medzvrstnih razlik in razmerij odločila za natančno slovensko narečno fonetično transkripcijo. Čeprav sem že takrat okvirno določila nabor variabel, sem želela preveriti, ali so morda za variantnostno analizo relevantne še druge. V pilotnem poskusu se je izkazalo, da so izbrane variable ustrezne, da je na različnih jezikovnih ravneh sicer še nekaj drugih, vrednih pozornosti, da pa bi bilo transkripcijo vendarle smiselno poenostaviti in pospešiti, saj je za zanesljivost rezultatov potrebna zadostna količina podatkov. V doktorski raziskavi sem tako izbrane izseke iz celodnevnega samosnemanja prvih dveh informantk zapisala v ozki fonetični transkripciji s pisavo ZRCola v skladu z določenimi načeli, pri preostalih treh informantih pa sem se odločila za ortografsko transkripcijo z nekaterimi posebnimi znaki. Pri transkribiranju sociolingvističnih intervjujev so bila transkripcijska načela še nekoliko poenostavljena (Bitenc 2016b: 182–183, transkripcije izsekov iz celodnevnega samosnemanja na 72 straneh so v prilogi 8.4.1, transkripcije relevantnih delov intervjujev pa na 74 straneh v prilogi 8.4.4 na CD-ju v Bitenc 2016a).

Več je torej primerov uporabe ortografske transkripcije z različnimi načeli zapisovanja posameznih jezikovnih elementov, večinoma so tudi ločila zapisana v skladu s standardom. Pri J. Škofic (1991) in M. Bitenc (2016a, 2016b) ni vejic in pik, le vprašaji pri vprašalni intonaciji, pri J. Škofic (1991.) tudi klicaji pri vzklični intonaciji, premori pa so označeni s pikami, od ene do treh pri M. Bitenc (n. d.) in tudi več pri J. Škofic (n. d.), odvisno od trajanja. M. Bitenc (n. d.) sem velike začetnice uporabljala le pri lastnih imenih. Pri K. Kotnik (2022) in M. Bitenc (n. d.) je označeno mesto naglasa, če je nenavadno ali razlikovalno.

3 Transkripcijska načela za specifične pojave govornega jezika

V nadaljevanju izpostavljam primere, glede katerih se pri odločanju za transkripcijska načela znotraj ortografske transkripcije pojavlja največ dilem, predvsem so to načela zapisovanja redukcije, variant fonema /v/, polglasnika, specifičnih narečnih glasov in premen po zvonečnosti.

3.1 Redukcije

V vseh omenjenih korpusih in raziskavah so upošteevane redukcije – glasov, ki niso izgovorjeni, torej raziskovalci oz. transkriptorji ne zapisujejo. Pri RSDO oz. D. Verdonik in A. Bizjak (2022: 25) so navedeni primeri *tud, neki, tko, mam, čevli*; pri M. Smolej (2012: 14) po alinejah glede na obliko in reduciran glas *mlek, okn, mel, igral, spal, dogajal, spal, pripravt, prijatu, drugač, men, clo, zlo, tko*.

3.2 Variante fonema /v/

Slovenski pravopis v poglavju o glasoslovju (Toporišič idr. 2003: 73) glede variant fonema /v/ navaja, da ustničnoustnični šumni *w* (zveneči) govorimo pred zvenečim soglasnikom in na koncu za zvočnikom, *m* (nezveneči) pa pred nezvenečim soglasnikom: *vnuke, barv, odvzeti – vsake, v tebi, predvsem*. Med posebnostmi je omenjeno, da namesto *w* in *m* zlasti pri počasnem govorjenju govorimo tudi *u*, da se pri besedah, ki se začenjajo s pisanim predponskim *u*-, ta izgovarja kot *u*, *w* oz. *m*; pred samoglasniki pa se predlog *v*, če se ne veže na končni samoglasnik prejšnje besede, izgovarja kot *w* ali *u*, odvisno od hitrosti govora.

V pravopisnih pravilih *Pravopis 8.0* in slovarju *ePravopis* je navedeno, da fonem /v/ izgovarjamo zobnoustnično ali dvoustnično: **zobnoustnično** pred samoglasniki iste besede in v nekaterih zvočniških sklopih; **dvoustnično** v dvofonemskih zvezah oz. fonetičnih dvoglasnikih (za samoglasnikom pred soglasniki in v izglasju) ter v vzglasju pred soglasniki in na morfemski meji (Glasoslovni oris, odstavek 39, ostala določila v zvezi s fonemom /v/ so v odstavkih do 45⁶). Po novem predlogu vse tri dvoustnične variante zapisujemo v oglatem oklepaju s črko *u*, po sistemu IPA pa bi bile zapisane s črko *w*. Če je dvoustnična varianta v izglasju, jo zapisujemo tudi s črko *v* in vezajem: očetov [očetov-]. Poenostavitev je pojasnjena s sledečim: »Razlik med tradicionalno **zapisanimi dvoustničnimi variantami fonema /v/** ([w], [ɱ], [u]) večina govorcev slušno ne zaznava, zato so jih jezikoslovci v kodifikacijskih in slovničnih opisih razvrščali in poimenovali glede na glasovno bližino kot **zvenečo** [w], **nezvenečo** [ɱ] in **dvoglasniško** varianto [u] (za samoglasniki). Dvoustnični izgovor se po slušnem vtisu približuje ujevskemu izgovoru, razlikujemo pa ju po zložnosti« (odstavek 40 v navedenem poglavju).

⁶ <https://www.fran.si/pravopis8/Podpoglavje/7-2>, dostop 23. 6. 2023.

V korpusu GOS zvočnik dvoustnični *v*, ki ni nosilec zloga, zapisujejo kot »v« (*prov*, *nav*, *navm*, *odprav*, *davn*, *pršov* – deležnik stanja na -l z glasovno premeno) oz. z »k«, če tako izhaja iz knjižne norme (*kosil*, *mel*). Če je u samoglasniški, tj. nosilec zloga, ga pišejo kot »u« (*pršu*, *vidu*, *u tem delu*) (Verdonik in Zwitter Vitez 2011: 59). Tudi v projektu RSDO zvočnik dvoustnični *v*, ki ni nosilec zloga, zapisujejo s črko »v«, če se pojavi v besednih oblikah, ki niso knjižne: *prov*, *nav*, *navm*, *odprav*, *davn*, *gledavc*, *pov@n*. Posebej opozarjajo na primere *lavfati*, *šlavf*, *genav*, *mav* (*malo*), *šov* (*šel*), *dov* (*dol*), *prov* (*prav*), *dav* (*da bo*), *nov* (*ne bo*), tudi medmet *av*. V poglavju o redukcijah navajajo tudi premene oblik za prihodnjik, izpostavljen je zapis primerov *čev* (*če bo*), *navm* (*ne bom*), *nav* (*ne bo*). Če dvoustnični *v* nastopa v besedni obliki, ki je knjižna in tudi izgovorjena skladno s standardom, ohranijo knjižni zapis: *bil*, *gledal*, *siv*. Če je glas u samoglasniški, tj. je nosilec zloga, pa ga pišejo s črko »u«: *pršu*, *vidu*, *u tem delu*. Tudi predlog *v*, izgovorjen kot samoglasniški *u*, pišejo kot *u* (Verdonik in Bizjak 2022: 26).

M. Smolej (2012: 14) med izjemami od standardnega zapisa navaja, da deležnik na -l za moško obliko ednine zapisuje po izgovorjavi z *u*, po drugi strani pa med nestandardnimi oblikami, ki jih zapisuje skladno s standardnim zapisom, navaja predlog *v*, ki ga zapisuje dosledno s črko *v* tudi v primerih, ko se izgovarja kot *u*. Pri K. Kotnik (2022) so deležniki na -l zapisani z -*u*, izgovor zvočnika *v* je zaradi poenostavitve zapisa in nerelevantnosti razlikovanja za namen raziskave ohranjen kot *v*, le pri (narečnem) dvoustničnem izgovoru je zapisala *w*. Pri J. Škofic (1991) v transkripcijah najdemo zapis *u* in predvsem v vlogi predloga *v*, čeprav je verjetno izgovorjen dvoustnično, npr. *V šo:li pa 'use govo'rim*.

M. Bitenc (2016a, 2016b) v fonetični transkripciji uporabljam *v*, *u* in *w* za narečni dvoustnični glas v položaju pred samoglasniki, ki je tudi varianta proučevane variable, v ortografski transkripciji z nekaterimi posebnimi znaki za variantnostno analizo treh informantov in v dodatno poenostavljeni transkripciji za sociolingvistične intervjuje pa *u* tako za *u* kot dvoustnični *u*, ki se izgovarja pred soglasnikom in na koncu besede, ter *w* za narečni dvoustnični glas pred samoglasniki. Zapisovanje dvoustničnih glasov z *u*-jem je bilo deležno kritike, saj gre za soglasniški glas, zapisan s samoglasniškim. Tako sem se pri podoktorskem projektu pri zapisu dvoustničnih variant fonema /v/ odločala med dosedanjim *u*, ki se kljub kritičnim pomislekom zdi izgovoru in splošnemu slušnemu vtisu najbližji ter se tudi v zapisu pogovornega jezika v spletnih in telefonskih sporočilih kaže kot najpogostejša

izbira,⁷ *v* in *l* v skladu s standardnim zapisom, *u*, kot je v predlogu novega Pravopisa 8.0, in *m*, kot je v sistemu IPA.

Po eni stani bi bilo najbolj korektno za poseben glas uporabljati poseben znak, vendar pa se zdi, da je, če ta glas ni predmet analize in mu ne posvečamo posebne pozornosti, lažje zapisovati in brati *v* in *l*. Tudi pri pretvorbi v poknjžžen zapis za vsebinsko analizo bi zapis s posebnim znakom terjal več sprememb. Posebej delikatni so primeri, ki se lahko izgovorijo z *l* ali *u*, kot npr. v besedi *mal* (npr. *guvurim maŭ dargac̣*). Da bi ohranili razlikovanje med obema načinoma, je v primeru izgovorjave *maŭ* vsekakor na mestu zapis z *v*, kar je skladno tudi s smernicami pri projektu RSDO, saj gre za nekñjižno pozicijo dvoustničnega *v*. Dilema se pojavlja tudi pri zapisovanju nekaterih deležnikov na *-l*. Pri pogovornih oblikah pride do primerov, kot sta npr. *nadzorovou* ali *šov*. Čeprav se po navodilih deležniki načeloma pišejo z *-l*, je pri pogovornih oblikah predviden zapis z *v*, ki se zdi manj nenavaden, torej *nadzorovov* ali *šov*.

3.3 Polglasnik

V pravilih *Pravopis 8.0* in slovarju *ePravopis* pri zapisu izgovora v oglatih oklepajih za polglasnik uporabljajo znak iz IPE, in sicer obrnjeno črko e tj. ⟨ə⟩, kot je navedeno v odstavku {4} poglavja Slovnični oris za pravopis. Zapis polglasnika je sicer opredeljen v odstavkih 76–96 poglavja Glasoslovni oris,⁸ v Slovenskem pravopisu pa na kratko v odstavku 620 s povezavami na druge relevantne odstavke (Toporišič 2003: 70).

V korpusu govornjene slovenščine GOS polglasnik ni zapisan pri zvočnikih l, r, m, n (*złoml*, *pr*, *misl*, *sn*); enoglasovnih predlogih, členkih, čeprav so izgovorjeni s polglasnikom (*s*, *z*, *d*), enozložnih besedah (*nč*, *jž*, *jsi*); lahko je zapisan z e v dvo- ali večzložnih besedah (*kešm*), razen pred l, r, m, n (Verdonik in Zwitter Vitez 2011: 58). V projektu RSDO polglasnik vedno zapisujejo z znakom @: pri zvočnikih l, r, m, n: *s@n*, *p@r*, *mis@l@m*, *bit@r*, *złom@l*, *prjat@lci*; pri enoglasovnih predlogih, členkih ipd., izgovorjenih vokalizirano, s polglasnikom: *s@*, *z@*, *d@*; pri enozložnih

⁷ Pomenljiv je tudi komentar sina, ki je ob prvih vajah branja negodoval nad zapisanim *l*-jem, ki se ne izgovarja kot *l*: »To je mal zoprno, ta glupi *l*. Zakaj se ne piše *u*?«

⁸ https://fran.si/pravopis8/Poglavje/7/slovnici_oris in <https://www.fran.si/pravopis8/Podpoglavje/7-2>, dostop 23. 6. 2023.

besedah: *j@ɤ, n@ɛ*; v dvo- ali večzložnih besedah: *k@šni (kakešni)*; v zborni izreki: *b@z@g, p@s* (Verdonik in Bizjak 2022: 25)

Pri J. Škofic (1991), K. Kotnik (2022) in M. Bitenc (2016a, 2016b) v vseh treh oblikah transkripcije je polglasnik zapisan kot *ə*. M. Smolej (2012) polglasnika ne zapisuje, med nestandardnimi oblikami, ki jih zapisuje skladno s standardnim zapisom, pa omenja kratke samoglasnike, ki so v izgovoru reducirani v polglasnik in pri katerih ohranja zapis npr. s črko *a* oz. *i*, npr. *jaɤ, brət, kear, nič* (Smolej 2012: 14).

3.4 Specifični narečni glasovi

V korpusu GOS so pokrajinsko specifični glasovi zapisani z najbližjimi ustreznimi črkami – *ej* (rejs – res), *aj* (lajtus – letos), *uj* (tujdi – tudi), *ov* (prov – prav), *av* (tav – to), *uo* (puole – pol), *je* (rjekla – rekla), *ju* (tjuk – toliko), *ue* (zmuetlu – zmotilo), *ea* (nea – ne), *ua* (uaknu – okno), *uo* (duobru – dobro); *u* ali *i* tudi za *u* s preglasom, *b* ali *g* za zveneči primorski *b*, *r* za mehkonobni koroški *r* (Verdonik in Zwitter Vitez 2011: 59). Tudi v projektu RSDO diftonge in druge pokrajinsko specifične foneme, ki jih ni v knjižnem jeziku, pišejo z najbližjimi ustreznimi črkami, odvisno tudi od izgovorjave v konkretnih primerih: v navodilih navajajo *ej, ov, je*; »u« za *u* s preglasom; »h« ali »\$g« za zveneči primorski *h*; »r« ali »\$r« za mehkonobni koroški *r* (Verdonik in Bizjak 2022: 26).

M. Smolej (2012: 14) med izjemami, ki niso skladne s standardnim slovarjem, navaja zapisovanje ukanja v prednaglasnem zlogu (npr. *utrok, pupadki, purudnišnica*). K. Kotnik (2022) za zapis narečnih širokih samoglasnikov uporablja strešico (*ožeroma, kēda, dōma, Primōrske*), za oznako narečnih naglasov tudi krativec in ostrivec (*zɤɤčɤer, òtrok, d̂ɤɤač, dōma, babīca, v M̂aribōrn, óbadv̂a*).

Tudi v *Pravopisu 8.0* in slovarju *ePravopis* so pri zapisu v oglatih oklepajih naglasna znamenja na črkah *e* in *o* uporabljena tudi za označevanje kakovosti, tj. širine oziroma ožine; za široka *e* in *o* torej ⟨ê⟩ in ⟨è⟩ oz. ⟨ô⟩ in ⟨ò⟩ (v IPI ⟨ɛ⟩ oz. ⟨ɔ⟩) ter ⟨é⟩ in ⟨ó⟩ za ozka *e* in *o* (v IPI ⟨e⟩ oz. ⟨o⟩) (odstavek {3} v poglavju Slovnični oris za pravopis).

M. Bitenc (2016a, 2016b) v ortografski transkripciji uporablja znake za narečne glasove, ki predstavljajo variante proučevanih variabel, in sicer γ za narečni pripornik γ ; γ , že omenjen w za narečni dvoustnični glas pred samoglasniki, i : za dolgi i (iz izhodiščno-slovenskih $\acute{e}(:)$, $\acute{e}(:)$ in $e(:)$; z izjemo kratkih v zadnjih zlogih); u : za dolgi u (iz izhodiščno-slovenskih $o(:)$ in $o(:)$; z izjemo kratkih v zadnjih zlogih); e : za dolgi ozki e (kjer bi bil v narečju navadno i : ali ie v neidrijskih varietetah; npr. *be'se:d*); o : za dolgi ozki o (kjer bi bila v narečju navadno u : ali uo v neidrijskih varietetah; npr. *'o:sem*); wo za idrijski dvoglasnik uo in je za idrijski dvoglasnik je .

3.5 Premene po zvonečnosti

V GOS in RSDO premen po zvonečnosti v pisavi ne upoštevajo – zapišejo *tud dobr*, *tud tak*, čeprav se izgovori *tut dobr*, *tut tak*, kot navajata D. Verdonik in A. Bizjak (2022: 26). Pri RSDO kot izjemo izpostavljajo predloge s/z in k/h , ki jih pišejo tako, kot so izgovorjeni (Verdonik in Bizjak 2022: 26).

Tudi M. Smolej (2012) v skladu s standardnim zapisom v vseh primerih ohranja zapis zvonečega nezvočnika, K. Kotnik (2022: 48) glede nezapisovanja utemeljuje, da v obravnavanih narečjih gre za takšne premene, kot jih pozna tudi knjižni jezik.

M. Bitenc (2016a, 2016b) v transkripcijah za variantnostno analizo premene po zvonečnosti upoštevam in zapisujem, pri transkribiranju intervjujev, kjer je poudarek na vsebini, pa ne, razen pri zapisu variant proučevane variable (g): te pišem kot $/\gamma/$ oz. $/g/$ skupaj z alofoni.

Zapisovanje in proučevanje premen po zvonečnosti in njihove variantnosti bi bilo smiselno pri proučevanju škofjeloškega, horjulskega in poljanskega narečja rovtarske narečne skupine, v katerih zvoneči nezvočniki pred pavzo in nezvonečimi soglasniki ohranjajo zvonečnost.

4 Dileme pri razločevanju in nedoslednosti pri zapisovanju glasov

V tem poglavju izpostavljam nekatere izkušnje in izzive, s katerimi sem se srečevala v sociolingvističnih raziskavah. Pri poslušanju in zapisovanju se je namreč večkrat pojavil dvom o tem, za kateri glas gre, tudi pri določenih pojavitvah variant izbranih variabel, predvsem pri hitrem govornem tempu, tišjem ali manj artikuliranem govoru

in pri posnetkih z več udeleženci oziroma drugimi spremljajočimi šumi. Izseke s takimi primeri sem večkrat poslušala in nekatere dodatno analizirala. Tako sem npr. za določitev dolžine *a*-ja pogosto porabila precej časa: po večkratnem poslušanju je sledila analiza v programu Praat, pa nazadnje večkrat še ni bilo povsem jasno, ali gre za kratek ali dolg glas. V drugih primerih sem isti glas lahko zaznala kot *ə*, *a* ali *o* (npr. pri *kət/kot/kaɪ*) ali pa je bilo glas v sredinski legi težko določiti. Če se tudi po večkratnem poslušanju, tudi pri upočasnjem predvajanju, nisem mogla odločiti, sem za presojo prosila še koga. Pogosto se je zgodilo, da so drugi – vsi jezikoslovci slovenisti – glas opredelili drugače, kot sem ga opredelila sama.

Izkazalo se je, da poslušalčev oz. transkriptorjev lastni jezikovni sistem oz. sistemi, ki jih ima uzaveščene, znatno vplivajo na zaznavo. Pri jezikoslovcih je ta pogojena tudi s tem, da so zaradi poznavanja sistemskih lastnosti govora nagnjeni k temu, da slišijo, kar želijo slišati, oz. kar mislijo, da bi morali slišati. Naj kot primer navedem akanje. Narečno bi govorec rekel *ɣawa'rim 'ku:kər ɣawa'rim*, vendar se na posnetku namesto /a/-jev kot vmesna varianta pojavlja težko opredeljiv glas, ki smo ga različni poslušalci različno slišali in zapisali. Tako sem se kot narečna govorka pri vmesnem glaslu na tem mestu odločila za zapis *a*-jev, slovenistka iz Vipavske, ki ima v svojem narečju na mestu kratkih nenaglašanih samoglasnikov polglasnik, polglasnike, dialektologinja iz Dolenjske, ki že dolgo živi v Ljubljani, *o*-je, jezikoslovec Ljubljančan pa *a*-je in polglasnike. Vmesni glas bi se dalo bolj natančno zapisati tudi z ožjo transkripcijo kot *a*-jevski polglasnik *ɔ*, vendar se za to nisem odločila, saj sem presodila, da bi vnašanje dodatnih kategorij presojo lahko še otežilo in jo naredilo manj zanesljivo. Tako sem se nazadnje v vseh primerih odločila po svoji presoji ob upoštevanju zaznav drugih jezikoslovcev, kadar so ti podali svoje mnenje. Podobne dileme so bile pogoste pri variablah (v) in (e:) ter pri zlogotvornih zvočnikih, ki predstavljajo potencialne variable. Najbolj sporne primere, ki so jih različni poslušalci, včasih celo en sam, različno zaznavali, sem posebej označila in so v eni izmed različic zapisa obarvani.

J. Škofic (1991) je v primerih, ko se natančnega izgovora besed s posnetka ni dalo razbrati, obliko, ki se ji je zdela verjetnejša, napisala najprej, drugo možno obliko pa v poševnem oklepaju, npr. *do'ma: /du'ma:/* (Škofic 1991: 226). Tudi izkušnje s transkribiranjem pri projektu RSDO kažejo na težave v razločevanju glasov: »v določenih kontekstih se je res težko odločiti med *e* in *@*, pa v nekaterih tudi med kratkim *a* in *@*« (Verdonik v osebni korespondenci jeseni 2022).

Različne vmesne oz. medvarietetne⁹ oblike pri jezikovnem prilagajanju imamo lahko za rezultat številnih strategij kodnega preklapljanja in kodnega mešanja, pri čemer lahko govorimo o t. i. dimenzijah nevtralnosti: govorci kombinirajo elemente različnih varietet in s tem kažejo na družbeno bližino ali identifikacijo z več kodi in njihovimi govorci hkrati. Pri rabi značilnosti različnih kodov se družbena opozicija med njimi nevtralizira. Ta sinteza različnih kodov ali menjavanje med njimi govorcu omogoča izražanje dvojne identitete – tako lahko za različne poslušalce zveni bodisi kot narečni govorec in pripadnik lokalne skupnosti bodisi kot standardni govorec (Mæhlum 2000: 104–105; Bitenc 2016b: 115).

Nedoslednosti pri zapisovanju posameznih glasov glede na določene smernice so opazne pri vseh omenjenih raziskavah, čeprav so transkripcije delo enega transkriptorja. Zagotovo je problematika bolj izrazita v primeru več zapisovalcev v večjih kampanjih. D. Verdonik v osebni korespondenci (jeseni 2022) omenja gradivo, ki so ga transkribirali na Fakulteti za računalništvo in informatiko, kjer so uvedli črko u za u-jevski alofon /v/-ja, pa so to zapisovalci pogosto spregledali in zapisali v ali l. Podobno izkušnjo imam sama s transkriptorko, ki je pomagala transkribirati intervjuje – kljub navodilom za zapisovanje u-ja je v več primerih zapisan v ali l. Pri K. Kotnik (2022) in M. Smolej (2012) se, obratno, kljub načelu zapisovanja z v ali l pojavi tudi zapis z u. D. Verdonik (osebna korespondenca jeseni 2022) obenem izraža nezadovoljstvo z odločitvijo za dosledno zapisovanje polglasnika (z znakom @) v projektu RSDO, ker v tako velikih kampanjah ni mogoče pričakovati, da bo dosledno zapisan.

5 Zaključek

Tudi v korpusih in raziskavah govorne slovenščine so transkripcijski protokoli, načela in prakse različni glede na teoretsko usmeritev raziskovalcev, predmet in namen raziskav, pa tudi obseg projekta, razpoložljivi čas in finančna sredstva. Iskanje kar najbolj optimalnega ravnotežja med čim večjo natančnostjo in doslednostjo ter obenem preprostostjo zapisovanja in branja je vsakič znova izziv. Natančna fonetična transkripcija se uporablja v dialektologiji, sicer pa večinoma ortografske transkripcije z različnimi načini in stopnjami poenostavljanja, navadno s standardno ortografijo in po potrebi posebnimi oznakami za specifične narečne glasove ter ločili. V skladu s splošnimi smernicami se je v izogib (pre)dolgim in (pre)zapletenim

⁹ Gre za oblike, ki jih ne moremo pripisati ne narečju ne standardu (ang. *interdialectal*).

procesom smiselno odredi transkribiranju podrobnosti, za katere kaže, da pri načrtovani študiji ne bodo potrebne, in se odločiti zapisovati za analizo relevantne podatke. Ob morebitni želji po proučevanju drugih variabel se je vedno mogoče vrniti k posnetkom, jih z osredotočenostjo na njihove variante spet poslušati in prilagoditi transkripcije. Zdi se tudi smiselno, da so za morebitne nadaljnje primerjave ali vključitve v podatkovne zbirke (npr. CLARIN.SI) načela zapisovanja čim bližje tem, ki so v veljavi v dosedanjih projektih in vključenih gradivih.

Izkušnje pričajo, da je pri določenih vmesnih glasovih, tudi pri pojavitvah posameznih izbranih variabel, težko presoditi, za katero varianto gre. Pri jezikovnem prilagajanju namreč prihaja do različnih oblik kodnega mešanja in preklapljanja, s katerimi se družbena opozicija med različnimi varietetami nevtralizira, vmesne oblike pa govorcem omogočajo izražanje dvojne identitete. Pri zaznavanju in zapisovanju glasovja se kot pomemben kaže vpliv jezikovnega sistema in vnaprejšnjih pričakovanj transkriptorjev. Za namen transkribiranja predvsem v večjih projektih bi bilo tako smiselno in potrebno usposabljanje in z njim povezana praksa, ki bi te vplive postopoma zmanjšala. Dobrodošle bi bile tudi natančnejše meritve in zato sodelovanje fonetika. Predvsem v primeru večjega števila transkriptorjev pri zapisovanju neizogibno prihaja do določenih nedoslednosti, predvsem pri bolj delikatnih primerih, kot so npr. variante fonema /v/ ali polglasnik.

V okviru projekta *Mezzanine oz. Temeljne raziskave za razvoj govornih virov in tehnologij za slovenščino*¹⁰ so v sklopu Narečna variabilnost med drugim v načrtu naloge, ki bodo lahko pomembno prispevale tudi k enostavnejšemu odločanju in razreševanju dilem glede zapisovanja glasov v nadaljnjih obravnavah jezikovne variantnosti, npr. zasnova sintetične jezikovnogeografske predstavitve fonetičnega/fonemskega inventarja slovenskih narečij, razvoj robustnega modela za prostorski prikaz fonetične variantnosti na podlagi starejših jezikovnogeografskih podatkov, optimizacija široke fonetične transkripcije za govorne korpusne slovenskega jezika in opredelitev razširjenega nabora slovenskih fonemov, ki vključuje na novo opredeljene narečne različice.

¹⁰ <https://mezzanine.um.si/>, dostop 23. 6. 2023.

Literatura

- Maja BITENC, 2016a: *Varietnost govorne slovenščine pri geografsko mobilnih osebah: Doktorska disertacija*. Ljubljana: Filozofska fakulteta.
- Maja BITENC, 2016b: *Z jezikom na poti med Idrijskim in Ljubljano*. Ljubljana: Znanstvena založba Filozofske fakultete.
- Klaus BRINKER, Sven F. SAGER, 1996: *Linguistische Gesprächsanalyse: Eine Einführung*. Berlin: Erich Schmidt.
- Barbara IVANČIČ KUTIN, 2011: *Živa pripoved v zapisu: kontekst, tekstura in prekodiranje pripovedi Tine Kravanja iz Baršice*. Ljubljana: Založba ZRC, ZRC SAZU. (Ethnologica – Dissertationes, 3).
- Karmen KENDA-JEŽ, 1996: Uvodna pojasnila. *Dialektološke in zgodovinske razprave*. Tine LOGAR, ur. Karmen Kenda-Jež. Ljubljana: ZRC SAZU, Inštitut za slovenski jezik Frana Ramovša. VII–XVIII.
- Karmen KENDA-JEŽ, 2002: *Cerkljansko narečje: Teoretični model dialektološkega raziskovanja na zgledu besedišča in glasoslovja: Doktorska disertacija*. Ljubljana: Filozofska fakulteta.
- Karmen KENDA-JEŽ, 2011: Fonetična transkripcija. *Slovenski lingvistični atlas 1: Človek (telo, bolezni, družina)*. Ur. Jožica Škofic. Ljubljana: Založba ZRC, ZRC SAZU.
- Katja KOTNIK, 2022: *Varietnost govorne slovenščine pri geografsko mobilnih mladih govornikih, govornikah izbranega korškega narečja: Magistrsko delo*. Maribor: Filozofska fakulteta Univerze v Mariboru.
- Steinar KVALE, Svend BRINKMANN, 2009: *InterViews: Learning the Craft of Qualitative Research Interviewing*. Los Angeles: SAGE Publications.
- Sofia LAMPROPOULOU, 2012: *Direct speech, Self-Presentation and Communities of Practice*. London: Continuum International Publishing Group.
- Tine LOGAR, 1993: *Slovenska narečja*. Ljubljana: Založba Mladinska knjiga.
- Brit MÆHLUM, 2000: Strategies of Neutrality in the Arctic. *Dialect and Migration in a Changing Europe*. Ur. Klaus Mattheier. Frankfurt am Main, Berlin, Bern, Bruxelles, New York, Oxford, Wien: Peter Lang. 99–108.
- Shana POPLACK, 1989: The Care and Handling of a Megacorpus: The Ottawa-Hull French Project. *Language Change and Variation*. Ur. Ralph Fasold, Deborah Schiffrin. Amsterdam, Philadelphia: John Benjamins. 411–444.
- Pravopis 8.0: Pravila novega slovenskega pravopisa za javno razpravo*. Dostop 23. 6. 2023 na www.fran.si/pravopis8.
- Dennis PRESTON, 2000: Mowr and Mowr Bayud Spellin': Confessions of a Sociolinguist. *Journal of Sociolinguistics* 4/4, 614–621.
- Mojca SMOLEJ, 2012: *Besedilne vrste v spontanem govoru*. Ljubljana: Znanstvena založba Filozofske fakultete.
- Jožica ŠKOFIC, 1991: *Problemi slovenskega pogovornega jezika: Magistrsko delo*. Ljubljana: Filozofska fakulteta.
- Sali A. TAGLIAMONTE, 2006: *Analyzing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Sali A. TAGLIAMONTE, 2007: Representing Real Language: Consistency, Trade-Offs and Thinking Ahead! *Using unconventional digital language corpora 1: Synchronic corpora*. Ur. Joan Beal, Karen Corrigan, Hermann Moisl. Basingstoke, Hampshire: Palgrave Macmillan. 205–240.
- Jože TOPORIŠIČ, Franc JAKOPIN, Janko MODER, Janez DULAR, Stane SUHADOLNIK, Janez MENART, Breda POGORELEC, Kajetan GANTAR, Martin AHLIN (ur.), 2003: *Slovenski pravopis*. Ljubljana: Založba ZRC, ZRC SAZU.
- Darinka VERDONIK, Ana ZWITTER VITEZ, 2011: *Slovenski govorni korpus GOS*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Darinka VERDONIK, Andreja BIZJAK, 2022: *Navodila za zapisanje in označevanje govora v projektu RSDO: Verzija 15*. Maribor: Fakulteta za elektrotehniko, računalništvo in informatiko Univerze v Mariboru.
- Peter WEISS, 1990: *Govori Zadrebke doline med Gornjim Gradom in Nazarjami: Glasoslovje, oblikoslovje in skladnja: Magistrsko delo*. Ljubljana: Filozofska fakulteta.

- Alison WRAY, Aileen BLOOMER, 2006: *Projects in Linguistics: A Practical Guide to Researching Language*. London: Hodder Arnold.
- Jana ZEMljARIČ MIKLAVČIČ, 2008: *Govorni korpusi*. Ljubljana: Znanstvena založba Filozofske fakultete, Oddelek za prevajalstvo. (Zbirka Prevodoslovje in uporabno jezikoslovje).

