

# MODEL ZA KORPUS NAREČNEGA GRADIVA TERSKEGA NAREČJA JANA BAUDOUIA DE COURTENAYJA

Janoš Ježovnik

Inštitut za slovenski jezik Frana Ramovša, ZRC SAZU, Ljubljana

UDK 811.163.6'282.4(450.365)

V prispevku je predstavljen model za izgradnjo nastajajočega korpusa narečnega gradiva terskega narečja Jana Baudouina de Courtenayja. Izpostavljene so nekatere specifične korpusa in gradiva ter opisana načela lematizacije in oblikoskladenjskega označevanja.

narečni korpus, Jan Baudouin de Courtenay, slovenščina, tersko narečje

The article discusses the model for setting up a corpus of dialect materials from the Tersko dialect of Slovene as compiled by Jan Baudouin de Courtenay. The specifics of the material as well as the corpus itself are introduced, and the principles of lemmatisation and POS-tagging are described.

dialect corpus, Jan Baudouin de Courtenay, Slovene, Tersko (Torre) dialect

## 1 Uvod

Delo Jana Baudouina de Courtenayja (1845–1929) o »govorih terskih Slovanov v severovzhodni Italiji«, kot se glasi njegov podnaslov, je izšlo leta 1904 in je druga iz serije knjig gradiv za južnoslovansko dialektologijo in etnografijo, *Materialen zur südslavischen Dialektologie und Ethnographie/Материалы для южнославянской диалектологии и этнографии* (v nadaljevanju: Mat. II). Serija naj bi po načrtih avtorja obsegala več knjig; kot je zapisal v pismu Karlu Štreklju, naj bi knjigam z rezijskimi, terskimi in nadiškimi govori sledilo »še kacic 7–8 tomov s tvarino, nabrano po Goriškem in Kranjskem« (Lenček 1985: 95). Ti načrti so ostali nerealizirani, obsežno rokopisno gradivo za omenjene publikacije pa je Baudouin de Courtenay zapustil arhivu ruske akademije znanosti.

Namen pričujočega prispevka je predstaviti model za vzpostavitev korpusa narečnega gradiva terskega narečja Jana Baudouina de Courtenayja, zbranega v Mat. II. Predstavljeni bodo značilnosti in omejitve obravnavanega gradiva ter način inkorporacije gradiva, torej značilnosti korpusa in načela njegove izgradnje, vključno z načeli lematizacije in označevanja.

## 2 Narava korpusa

Korpus narečnega gradiva terskega narečja Jana Baudouina de Courtenayja lahko razumemo kot **zgodovinski** in kot **dialektološki** korpus.

Zgodovinski je zato, ker vsebuje besedilno gradivo, zbrano v letih 1873 in 1901, in prikazuje stanje v jeziku tistega obdobja. Tega se je treba zavedati tudi pri uporabi korpusa za dialektološke raziskave sodobnega narečja: korpus je lahko v pomoč, izsledke pa je treba preveriti na sodobnem gradivu.

Ker so v korpusu vsebovana besedila narečna, ga lahko štejemo tudi med dialektološke korpuse. Od modela dialektološkega korpusa, ki ga je za slovenščino s korpusom GOKO prva vzpostavila Klara Šumenjak (2013: 32–35), se loči zlasti po naslednjih značilnostih:

1. medtem ko je GOKO v osnovi govorni korpus, kar je za raziskave narečja smiselno, bo korpus terskega narečnega gradiva Jana Baudouina de Courtenayja zaradi svoje pisne predloge le pisni korpus;
2. korpus GOKO poleg fonološkega zapisa narečnega besedila vsebuje poenostavljeni narečni zapis besedila in njegovo poknjženo različico (Šumenjak 2013: 45), ki sta ob poizvedbi prikazana vzporedno; korpus terskega narečnega gradiva Jana Baudouina de Courtenayja bo prikazoval le prepis narečnega zapisa, čeprav dopolnitev z ruskim in morda knjižnoslovenskim prevodom ni izključena;
3. ker gre za inkorporacijo zaključene gradivne celote, bo korpus ob vključitvi gradiva iz Mat II. zaprt za nadaljnje gradivske posodobitve; model korpusa GOKO ostaja za tovrstne dopolnitve odprt;
4. pri korpusu GOKO ni označen izvorni narečni zapis, temveč njegova poknjžena različica, kar omogoča avtomatsko označevanje z označevalnikom za slovenski knjižni jezik; zaradi uporabe prilagojenega nabora oblikoskladenjskih oznak avtomatsko označevanje v primeru korpusa gradiva iz Mat II. ni mogoče, zato bo korpus označen ročno, pri čemer bo zaradi ohranitve enostavnosti iskanja lematiziran v poknjženi obliki.

### **3 O gradivu**

Delo Mat. II je sestavljeno iz dveh delov. Prvi, najobsežnejši del obsega avtorjeve terenske zapise, v drugem pa je zbrano tuje gradivo v manjšem obsegu. Na začetku je obsežen uvod, ki obsega predvsem pojasnila o načinu zapisa, tj. načela označevanja kvalitete in kvantitete različnih glasov ter naglasa, opombe o točnosti zapisa ipd. Zaključni del vsebuje podrobnosti o zapisih in seznam informatorjev (Mat. II 1904: 211–217) ter stvarno kazalo (prav tam: 220).

Narečno gradivo iz prvega dela je Baudouin de Courtenay zapisal v krajih Terskih dolin v letih 1873 in 1901; v času drugega obiska je preveril tudi pravilnost zapisa nekaterih besedil, zapisanih v času prvega obiska; če je bilo mogoče, pri istih informatorjih. Kljub temu je »med ponovno preučitvijo terskih govorov l. 1901 zdaleč privilegiral nabiranje novega gradiva pred izčrpnim preverjanjem starega« (Spinozzi Monai 1988: 48).

Objavljeno je gradivo, zapisano v osemnajstih krajih Terskih dolin. Gradivo posameznega kraja predstavlja zaokroženo poglavje. Na začetku vsakega poglavja so navedeni naslov in vsebina besedil v ruskem in nemškem jeziku, informatorji in morebitne opombe v zvezi z zapisom. Vse narečno gradivo je opremljeno s številkami vrstic in z vzporednim ruskim prevodom.

## VII. Заверх (Zavàrx) (Villanova).

Нѣсколько образцов языка, записанных в 1873 г.

Einige im J. 1873 г. niedergeschriebenen Sprachproben.

Они записаны со слов Анджела Негро по прозв. Саботича (Négro Anželo sorakojoŋ Sabótič), 40 лѣт.

- 418 ¿Káko oblačilo su -mjéle ženè táza - Wàrxax?  
 419 Katére o - mà ragadiŋ, kišni meželànu. Dàm - bòt su -mjéle  
 ženè te - stàre čemòsu bjèlo wòz - wòwne; faculète su -mjéle bjèle  
 420 wuz - nítš še - tè alpiŋr wùs - prèje (alí wus - prèje). Su -mjéle  
 še pàs čàr wùs - korjàna; su -mjéle šè kamžòlu po - zìme wùnz  
 - wòwne, lš bjèlu, lš čòrnu, á po - ljàte w - srákьs w - mònajètax;  
 su -mjéle jamizòt čàr po - ljàte.

- 418 ¿Какое платье было у женщины За-Верхом?  
 419 У одного полосатая матерія, у другого полушерстяная. Нѣ-  
 когда старья женщины имѣли (носили) бѣлый шерстяной пояс;  
 420 платки у них были бѣлые нитяные или же из пряжи. [Мужчины]  
 носили тоже черныи кожаный пояс (кушак); носили тоже зимоу  
 жакет (полукафтан) шерстяной, либо бѣлый, либо черныи, лѣтом  
 же [ходили] в рубашкѣ с рукавами; [кромѣ того] носили лѣтом  
 черныи камзол.

Slika 1: Primer zapisa v Mat II. (1904: 87)

### 3.1 Način zapisa

Način zapisa narečnega gradiva je podrobno opredeljen v uvodu Mat. II (1904: IV–XVIII). Baudouin de Courtenay je pri zapisu uporabljal fonetičnim lastnostim govora prilagojeno latinično pisavo.

Večina grafemov »ne potrebuje posebne razlage« (Mat. II 1904: VII), v nadaljevanju pa so razloženi nekateri posebni znaki. Po potrebi je dodan še približek izgovora, kot ga lahko razberemo iz opisa, v slovenski fonetični transkripciji (FT; prim. Kenda - Jež 2011: 27–30) in zapisu IPA.

Grafemi *š*, *ž*, *č*, *ž* označujejo ustrezne šumniške pripornike in zlitnike (FT: *š*, *ž*, *č*, *ž*; IPA: *f*, *ʒ*, *tʃ* in *dʒ*), pri čemer »se pri izgovoru teh soglasnikov v terskih govorih sliši izrazita srednjejezičnost (palatalnost, 'mehkost')« (Mat. II 1904: VIII). Grafem *x* označuje nezveneči mehkonebni pripornik, zveneči mehkonebni pripornik pa je označen z grafemom *γ* (FT: *x*, *γ*; IPA: *x*, *ɣ*). Grafema *f* in *ɟ* označujeta trdonebna zlitnika (FT: *ć*, *ǰ* oz. *l'*, *d'*; IPA: *tɕ*, *dʒ* oz. *c*, *ʝ*): »Izgovor teh fonemov včasih variira. Zaradi svoje neopredeljenosti spominjata na srednjejezična, 'močno mehka' *l'* in *d'* brez

primesi sičniškega elementa, spet drugič dajeta vtis, ki bi ga lahko boljše odražala grafema *ć* [in] *ž*.« (Mat. II 1904: IX).

V gradivu se variante nosnega zobno-jezičnega zapornika označujejo s štirimi grafemi: *n*, *ñ*, *ŋ*, *ɲ*. Prvi označuje navadni zobno-jezični izgovor, ostali pa:

*ŋ* – trdnebnosni nosni zapornik (FT: *ń*; IPA: *ɲ*);

*ɲ* – mehkonebnosni nosni zapornik, alofon *n* pred mehkonebniki (FT, IPA: *ŋ*);

*ñ* – nazaliziranost predhodnega samoglasnika; znak iz tipografskih razlogov nado-mešča označevanje nosnosti na grafemih, ki označujejo samoglasnike: »*añ* = *a*, *eñ* = *e*, *oñ* = *o* itd.« (Mat. II 1904: X).

Grafemi *ε*, *ə* in *e* označujejo ozki, široki oz. srednji *e*, grafemi *ω*, *ω* in *o* pa ozki, široki oz. srednji *o*. Boudouin de Courtenay navaja, da pri *e*-jevskih in *o*-jevskih glasovih zapis fonetične vrednosti ni vedno točen: tako lahko pisna *e* in *o* ponekod označujeta tudi *ε* ali *ə* oz. *ω* ali *ω*, nekateri zapisani *ε* in *ω* pa *i* oz. *u* (Mat. II 1904: XII–XIII).

Grafema *ɸ* in *ɸ* označujeta izgovor reduciranih glasov, ki imata po de Courtenayjevem opisu fonetično vrednost kratkih *ε* in *ω* (Mat. II 1904: XIII–XIV). V gradivu se prvi znak pojavlja občutno pogosteje od drugega.

Ostrivec nad soglasniki (npr. *́*, *́*, *́*, *́*, *́*, *́*, *́* ...) označuje njihov palatalizirani izgovor; ponekod je takšen izgovor označen z nadpisanim *ʼ* za soglasnikom (npr. *sʼ*, *tʼ*), sicer pa nadpisani znaki, tudi *i*, pomenijo bežen, reduciran izgovor ustreznega fonema (Mat. II 1904: XIV). Označena je le izrazita palataliziranost (prav tam).

Polkrožec nad znakom (*ǎ*, *ě*, *ǎ*, *ě*, *ǎ*, *ě*, *ǎ*, *ě*, *ǎ*, *ě*) označuje izrazito kratkost nenaglašeneega zložnika, črtica nad znakom (*ā*, *ē*, *ē*, *ā*, *ō*, *ō*, *ā*, *ī*, *ū*, *ī*) pa dolžino naglašeneega ali nenaglašeneega zložnika, izjemoma tudi nezložnika (npr. *ñ*). Črtica lahko nastopa v kombinaciji s krativcem ali ostrivcem, ki označujeta tip (tonemskega) naglasa. Prvi načeloma označuje padajoči (brez znaka za dolžino pa lahko tudi (kratki) dinamični naglas), drugi pa rastoči naglas (brez znaka dolžine pa tudi katerikoli naglas na dolgem zložniku ne glede na tonemsko vrednost) (Mat. II 1904: XVI–XVIII).

Boudouin de Courtenay je za zapis kasnejših dodatkov k besedilu uporabljal različne vrste oklepajev (Mat. II 1904: IV, 17).

Okrogli oklepaj vsebuje zapis variantnega dela besedila:<sup>1</sup>

*Tì mь màš povjédàtь, kěj visokò jè tas tlù yòr do tэ přvě zvjédzě (yòr do tьx přvix zvjédzě) [...].* (Mat. II: 4)

Ti mi imaš povedati, kaj visoko je ta s tal gor do te prve zvezde (gor do teh prvih zvezd) [...].

‘Ti mi moraš povedati, kako visoko je od tal gor do prve zvezde (gor do prvih zvezd) [...].’

<sup>1</sup> Daljše navedbe narečnega besedila so urejene tako, da je v prvi vrstici najprej navedeno narečno besedilo v izvornem zapisu skupaj z navedbo strani v Mat. II. V drugi vrstici je isto besedilo oblikovno-pravopisno poknjženo, v oklepajih so pojasnjene morebitne oblikovne posebnosti. V tretji vrstici je zapisan knjižno-slovenski prevod.

Oglati oklepaj vsebuje dodatke k prvotnemu besedilu, nove besede ali pojasnila – v gradivu uvaja v glavnem popravke in dopolnitve prvotnih besedil ob pregledu leta 1901. Oglati oklepaj lahko označuje tudi fakultativnost glasov ali delov besed:

[...] *an ššs[t] sòwdə jə kóštalo po lítre.* (Mat. II 1904: 8)

[...] anu šest soldov je koštalo po litru.

‘[...] in šest soldov na liter je stalo.’

*Ojà jə ujèw ta[ɣa] veličɣa méɣa bràtra [...].* (Mat. II 1904: 12)

Oča je ujel tega velikega mojega brata [...].

‘Oče je ujel mojega starejšega brata [...].’

Zaviti oklepaj vsebuje odvečne ali nerazumljive dele besedila; v glavnem se pojavlja v vzporednem ruskem prevodu.

Zvezdici na začetku in koncu okroglega in oglatega oklepaja (tj. (\* \*) in [\* \*]) označujeta, da so bili dodatki ali popravki dopisani leta 1901 med ponovnim pregledom besedila, zapisanega leta 1873:

*Anzət ɔŋ ə šəw dów ɣ wówkán, ɳəɣá kɔmpáŋan, ánu wəs wštrášən (\*wštrášən\*) ə jən jàw (\*dàw\*): [...]* (Mat. II 1904: 30)

Anzat on je šel dol k volkom, njega kompanjem, anu ves ustrašen (\*ustrášen\*) je jim jal (\*đal\*): [...]

‘In zatem je šel dol k volkovom, svojim tovarišem, in jim ves prestrašen dejal: [...].’

*Tə ɣrjèx tù nedèju plēsàtə [\*Bwéx ɔ nə naɣà\*] nə parpustì bwòžja jušticia [...]* (Mat. II 1904: 35)

Te (= to je) greh tu v nedeljo plesati [Bog on (*os. zaim., 3. os. ed. m. sp., nasl.*) ne naga] ne pripusti božja jušticia [...]

‘Plesati v nedeljo je greh, [Bog ne dovoli,] ne pripusti božja pravičnost [...].’

### 3.2 Omejitve gradiva

Zaradi časovne oddaljenosti je treba gradivo jemati kot zapis zgodovinskega stanja v narečju. Iz njega ne moremo neposredno sklepati o današnjem stanju v narečju, posredno pa lahko sklepamo o nekaterih pojavih, ki se odražajo v sedanji podobi narečja, ter o njihovih zametkih. Tako lahko npr. iz opombe v uvodu, da se pri izgovoru šumnških fonemov »sliši izrazita srednjejezičnost (palatalnost, ‘mehkost’)« (Mat. II 1904: VIII), razberemo dejavnik, ki je pomembno prispeval k današnji izgubi fonološkega nasprotja med sičniki in šumevci zlasti v zahodnem delu narečja.<sup>2</sup> Ker se posebnosti pri kvaliteti sičniških fonemov neposredno ne omenja niti te v gradivu niso dosledno označene, ne moremo samodejno sklepati na prisotnost tega pojava že v tistem času.

Naglasne oznake v gradivu so dokaj nepovedne in nezanesljive, kar v uvodu priznava tudi avtor (Mat. II 1904: XVII–XVIII). Ali gre za nezanesljivost zapisa ali pa je nihanje v realizaciji tonemov prisotno že pri Baudouinovih informatorjih in gre torej za nevtralizacijo tonemskih nasprotij, bo treba še preveriti. Glede na podobno nepoveden zapis v gradivu nadiškega narečja (Spinuzzi Monai 1988: 56–57), ki tonemska nasprotja ohranja še danes (prim. npr. Šekli 2008), bi lahko sklepali, da

<sup>2</sup> S pojavom se celo vzpostavlja tonemski minimalni par [nùɔs] ‘nos’ : [nuɔs] ‘nož’ (Ježovnik 2014: 39).

gre za nezanesljivost zapisa. Gradivo zato ni primerno za natančne naglasoslovne raziskave.

#### 4 Lematizacija

Zaradi kompleksnosti zapisa je potencialno problematičen način podajanja gradiva v obliki, v kakršni je izvorno predstavljeno. Poenostavitev zapisa v skladu s smernicami, predstavljenimi v uvodu, se ne zdi smiselna, ker današnji prikazovalniki omogočajo prikaz naprednega nabora pisav. Poleg tega, da bi bilo tovrstno početje zamudno, obstaja tudi nevarnost nehotene izgube podatkov, saj razmerje med zapisom in okvirnimi fonetičnimi vrednostmi, ki ga vzpostavljajo smernice, ni vedno povsem jasno.

Leme so zapisane v knjižni slovenščini oziroma so poknjžene. V tem se naše delo razlikuje od označevanja terskemu narečju tipološko podobnega rezijanskega narečja (prim. Steenwijk 2010). Ta projekt predvideva lematizacijo v standardni rezijanščini po načelih, predstavljenih v standardizacijskih delih Hana Steenwijka (1994, 1999, 2005). Za knjižno slovenščino smo se odločili, ker za tersko narečje še ni bil zasnovan model sistematične standardizacije različnih krajevnih govorov narečja in njihovega enotnega zapisa, snovanje obojega pa presega cilje in obseg teme. Če bi se odločili za zapis lem v narečju, bi bilo nujno najprej vzpostaviti celovit model za standardizacijo, ki bi zagotovil konsistentno in sistemsko medsebojno ujemanje leksemov in kategorij istega tipa. Zaradi neenakomerne gradivne zastopanosti različnih krajevnih govorov in njihove nezadostne opisanosti to zaenkrat ni mogoče. Prav tako iz podobnih razlogov še ni mogoče vzpostaviti zadovoljivega enotnega sistema zapisovanja. Zaradi kompleksnosti znakov in neenotnosti zapisovanja zgolj prepis z znaki, ki jih uporablja Baudouin de Courtenay, ne pride v poštev, saj bi bilo tako oteženo iskanje po lemah.

Narečnim oblikam so pripisane njihove knjižnoslovenske ustreznice ali njihovi oblikovni približki (npr. *kle, tle* → *tule* ipd.). V primeru večjih odstopanj od knjižnoslovenske ustreznice, zlasti če beseda v knjižnem jeziku ni prisotna, smo uporabili obliko, navedeno v Spinozzi Monai 2009a in jo ustrezno glasovno-pravopisno poknjžili<sup>3</sup> (npr. *djati* → *dejati, ejnje(le)* '(prav) zdaj'). Prav tako smo se v primeru pisnih dvojnic v gradivu odločili za obliko, ki je kot prednostna navedena v istem delu (npr. *žudih, žudi, žudih* 'sodnik' → *žudih* → *džudih*).

Narečne foneme /*ń*/, /*ɲ*/, /*ć*/ in /*ǰ*/ knjižimo po etimološko-zgodovinskem načelu (npr. *svjeća* → *sveća, ğati* → *dejati* [dǰati]). Težavo predstavljajo pozno prevzete besede, skoraj brez izjeme romanizmi. V tem primeru je knjiženje le glasovno-pravopisno, pri tem pa v zapisu ohranjamo fonema /*ć*/ in /*ǰ*/). Slednji je zapisan z grafemom *đ* (npr. *guđa* 'vrsta oblačila' < furl. *gugje*).<sup>4</sup>

3 Izrazje je povzeto po metodologiji knjiženja narečnih hišnih in ledinskih imen, predstavljeni v Klinar idr. 2012: 47–51.

4 Za zapis z grafemom *đ* smo se odločili zaradi enostavnejšega vnosa pri iskanju in zaradi vzporednosti z grafemom *ć* v knjižnih variantah štokavščine (podobna rešitev se pojavlja tudi v Steenwijk 1992). Zapis z dvočrkjema *dj* ali (kot v furlanščini) *gj* bi sicer ustrezal izraziti palatalni naravi glasu, vendar bi lahko prihajalo do zamenjav pri besedah z zaporedji *d + j* in *g + j*. Zapis z dvočrkjem *dž* ne pride v poštev, saj je tako že označen fonem /*ǰ*/; vzporedno z zapisom tega bi lahko po zgledu poljskega pravopisa uvedli dvočrkje

Tujejezični citatni zapisi, ki jih najdemo v gradivu, so lematizirani z ustreznimi tujejezičnimi knjižnimi oblikami, označeni pa so z oblikoskladenjsko oznako X (Ostalo).

Pri lematizaciji lastnih imen upoštevamo rešitve v Spinozzi Monai 2009a, v pomoč pa so nam različna onomastična in toponomastična dela (Merkù 1997, 2006; Dapit 2001).

## 5 Oblikoskladenjsko označevanje

### 5.1 Nabor oblikoskladenjskih oznak in načela označevanja

Za osnovo nabora oblikoskladenjskih oznak smo vzeli specifikacije MULTEXT-East (v. 5) za rezijansko narečje, ki jih je ob sodelovanju pri navedenem projektu izdelal Han Steenwijk (2010) in so označene z licenco Creative Commons.<sup>5</sup> Za to smo se odločili, ker je rezijanščina oblikovno in skladenjsko tipološko zelo podobna terskemu narečju. Omenjeni nabor predstavlja ustrezno osnovo tudi v primerjavi z uveljavljenim naborom oblikoskladenjskih oznak za označevanje slovenskega (knjižnega) jezika JOS (Erjavc idr. 2010), saj ta ni prilagojen narečnim značilnostim, ki jih knjižni jezik ne pozna, npr. določnemu in nedoločnemu členu, v narečju pogosteje rabljenim glagolskim oblikam (deležnik na -ć, vezni naklon ...), s prostorskim prislovom ojačanim prostorskim predlogom (*ta na, tle par, ta v ...*) itn.

Pri označevanju oblik v primeru nejasnosti označimo tiste, ki so skladenjsko pričakovane. S tem se izognemo napačnemu označevanju narečnih oblik, ki so homonimne nekaterim knjižnim:

*Ān mi mǎš povjédǎtǎ, kěj zvjézdě jě tǔ firmaměntu.* (Mat. II 1904: 4).

Anu mi imaš povedati, kaj zvezde je tu v firmamentu.

'In povedati mi moraš, koliko zvezd je na nebu.'

V navedenem primeru obliko *zvjézdě* zaradi skladenjskega okolja označimo kot rodilnik množine, ne pa morebiti zaradi končnice kot tožilnik množine.<sup>6</sup>

Podobno velja za označevanje veznega naklona, ki je v nekaterih oblikah homonimen velelnemu, zato se tudi v tem primeru oznaka določi glede na skladenjsko okolje (npr. obvezno nastopa v odvisnem stavku z ozkim naborom podrednih veznikov).<sup>7</sup>

Na nejasnih mestih si, če je to mogoče, pomagamo z ruskim prevodom. V spodnji povedi bi lahko bila poudarjena oblika označen bodisi kot rodilniška bodisi kot (brezpredložna) mestniška:

*Xódimo májše tu cjérkwo světs Marijǎ Mardalénǎ.* (Mat. II 1904: 106)

Hodimo [k] maši tu v cerkvo svete Marije Magdalene/[k] sveti Mariji Magdaleni.

'K maši hodimo v cerkev svete Marije Magdalene/k sveti Mariji Magdaleni.'

*dž*, vendar bi to precej otežilo vnos besede pri iskanju. Ta omejitev velja tudi za grafem *ǎ*, ki je za zapis podobnega glasu v rezijanščini predviden v Steenwijk 1994 in rabljen v nadaljnjih delih.

<sup>5</sup> Dostopno prek: <http://creativecommons.org/licenses/by-sa/3.0>.

<sup>6</sup> Ostale pojavitve v obravnavanem gradivu, pa tudi oblike, ki se sinhrono pojavljajo v narečju, nam potrjujejo, da se pri nekaterih samostalnikih *o-* in *ǎ-*sklanjatve v rodilniku množine namesto pričakovanih končnih res pojavlja končnica *-e*.

<sup>7</sup> V zvezi z veznim naklonom (subjunktivom oziroma konjunktivom) v terščini prim. Spinozzi Monai 2009b, v rezijanščini pa Ježovnik 2015.

Ker ruski prevod navaja roditeljsko (*святой Марию Магдалины*), in ne mestniške oblike (*\*святой Марию Магдалинѣ*), pri označevanju izberemo prvo.

## 5.2 Nekaj dopolnitev in sprememb izhodiščnega nabora

V nadaljevanju je navedenih nekaj večjih dopolnitev oziroma sprememb izhodiščnega nabora oblikoskladenjskih oznak in sistema označevanja (Steenwijk 2010).

### 5.2.1 Glagol

Iz osnovnega nabora sta bili odstranjeni oznaki *imperfekt* in *aorist*, saj se ti dve obliki v gradivu ne pojavljata. Izjema so (prvotno verjetno) aoristne oblike glagola *biti*, na katere so se naplastile spolske končnice; te so označene z oznako *preteklik*:

*Дá съ бѣ шѹвъ ѿчѣра вѹнѣ о́ру по сѣно[.]* (Mat. II 1904: 92)

Jaz si (1. os. ed. sed.) *bī* (m. sp. ed. pret./aor.) šel včeraj ven goro po seno.

'Včeraj sem (bil) šel na goro po seno.'

*Дí гра́чје је не бѣ̀ полѣнта, ма тѣ̀ нѣ̀бо[.]* (Mat. II 1904: 26)

Di gračje če ne (= na (os. zaim., 3. os. ed. ž. sp., nasl.) je) *ba* (ž. sp. pret. ed. pret./aor.) polenta, ma te (= to je) ne *bo* (sr. sp. ed. pret./aor.).

'Hvala bogu, če je bila polenta, a je ni bilo.'

Deležniki (oziroma deležja) so označeni bodisi z oznako za glagol bodisi z oznaki za pridevnik (deležniki) oziroma prislov (deležja), odvisno od skladenjske vloge v stavku. *Način*, atribut, ki v nabor JOS ni vključen, se označuje le pri deležnikih, in sicer tako, da se skupaj z oznako časa (*sedanjik* ali *preteklik*) tvorijo naslednja razmerja:

- tvornosedanji deležnik (na -ć): *sedanjik, tvorni*;
- tvornopretekli deležnik (na -l): *preteklik, tvorni*;
- trpnopretekli deležnik (na -n/-t): *preteklik, trpni*.

### 5.2.2 Zaimek

Poseben problem predstavlja označevanje sklopljenih oblik naslonske oblike osebnega zaimka za 3. osebo ednine ženskega spola (*na* 'ona') oziroma kazalnega zaimka za srednji spol (*to*) in glagola *biti* (*na* (*j*)e > *ne*; *to* (*j*)e > *te*). Te oblike so označene kot navezne oblike ustreznih zaimkov.

### 5.2.3 Prislov

Ker označevanje korpusa poteka ročno, je mogoča natančnejša pomenska delitev prislovov. Pri atributu *vrsta* je vrednost *splošni* nadomeščena z vrednostmi *kakovostni*, *merni*, *vzročni*, *prostorski* in *časovni* prislov (po Toporišič 2000: 406).

### 5.2.4 Predlog

Atribut *vrsta* v izhodiščnem naboru vsebuje le vrednost *predlog*. Ker so v terščini vse adpozicije predlogi, je ta atribut odvečen in je bil zato odstranjen iz nabora. Z vrednostjo *sestavljen* (atribut *tvorba*) so poleg (sinhrono razvidno) tvorjenih predlogov označeni tudi s prostorskim prislovom ojačani predlogi (npr. *ta za, ta na, ta par* ...).



## Literatura

- DAPIT, Roberto, 2001: *Cognomi e nomi di famiglia dell'Alta Val Torre*. Bardo: Občina Bardo.
- ERJAVEC, Tomaž, KREK, Simon, ARHAR, Špela, FIŠER, Darja, LEDINEK, Nina, SAKSIDA, Amanda, SIVEC, Breda, TREBAR, Blaž, 2010: *Oblikoskladenske specifikacije JOS*. Različica 1.1. <http://nl.ijs.si/jos/index-sl.html>
- JEŽOVNIK, Janoš, 2014: Izvor poimenovanj za orodja in hišne pripomočke v terskem narečju. *Jezikoslovni zapiski* 20/2. 37–52.
- JEŽOVNIK, Janoš, 2015: Vezni naklon v rezijanščini. Danila Zuljan Kumar, Helena Dobrovoljc (ur.): *Škrabčev zbornik 8. Zbornik prispevkov s simpozija 2013*. Nova Gorica: Založba Univerze v Novi Gorici.
- KENDA - JEŽ, Karmen, 2011: Fonetična transkripcija. Jožica Škofic (ur.): *Slovenski lingvistični atlas 1: Človek (telo, bolezni, družina). Atlas (SLA 1.1)*. Ljubljana: Založba ZRC, ZRC SAZU. 27–30.
- KLINAR, Klemen, ŠKOFIC, Jožica, ŠEKLI, Matej, PIKO - RUSTIA, Martina, 2012: *Metode zbiranja hišnih in ledinskih imen*. Jesenice, Celovec: Razvojna agencija Zgornje Gorenjske.
- LENČEK, Rado, 1985: Štrekljeva pisma Janu Baudouinu de Courtenayju in Baudouinova Štreklju. *Slavistična revija* 33/1. 72–100.
- MERKŪ, Pavle, 1997: *La toponomastica dell'Alta Val Torre*. Bardo: Občina Bardo.
- MERKŪ, Pavle, 2006: *Krajevno imenoslovje na slovenskem zahodu*. Uredila Silvo Torkar in Metka Furlan. Ljubljana: Založba ZRC, ZRC SAZU.
- SPINOZZI MONAI, Liliana, 1988: Introduzione/Uvod. Jan Baudouin de Courtenay, 1988: *Materiali per la dialettologia e l'etnografia slava meridionale IV. Testi popolari in prosa e in versi raccolti in Val Natisone nel 1873/Materiali za južnoslovansko dialektologijo in etnografijo. Ljudska besedila v prozi in verzih, zbrana v Nadiških dolinah leta 1873*. Trieste/Trst: Editorialie Stampa Triestina/Založništvo tržaškega tiska, Centro studi/Študijski center Nedizja. 11–57.
- SPINOZZI MONAI, Liliana, 2009a: *Il Glossario del dialetto del Torre di Jan Baudouin de Courtenay*. Videm: Consorzio Universitario del Friuli; Ljubljana: Inštitut za slovenski jezik Frana Ramovša ZRC SAZU, Sankt Petersburg: St. Petersburg branch of the Archive of the Russian Academy of Sciences.
- SPINOZZI MONAI, Liliana, 2009b: Ipotesi di un calco paradigmatico slavo-romanzo (l'imperativo-congiuntivo, uno studio fondato sul Glossario del dialetto del Torre di Jan Baudouin de Courtenay). *Demetrio Skubic octogenario. Linguistica* 49/2. Ljubljana: Znanstvena založba Filozofske fakultete. 295–308.
- STEENWIJK, Han, 1992: *The Slovene dialect of Resia: San Giorgio*. Amsterdam, Atlanta: Rodopi.
- STEENWIJK, Han, 1994: *Ortografia resiana. Tö jošť rozajanskě pisanjě*. Padova: CLEUP.
- STEENWIJK, Han, 1999: *Grammatica pratica Resiana. Il sostantivo*. Padova: CLEUP.
- STEENWIJK, Han, 2005: *Piccolo dizionario ortografico resiano. Mali bisidnik za tō jošť rozajanskě pisanjě*. Padova: CLEUP.
- STEENWIJK, Han, 2010: Resian Specifications. Tomaž Erjavec (ur.): *MULTEXT-East Morphosyntactic Specifications, Version 5*. <http://nl.ijs.si/ME/V5/msd/html/msd-sl-rozaj.html>
- ŠEKLI, Matej, 2008: *Zemljepisna in osebna lastna imena v kraju Livek in njegovi okolici*. Ljubljana: Založba ZRC, ZRC SAZU.
- ŠUMENJAK, Klara, 2013: *Opis govora Koprive na Krasu na osnovi dialektološkega korpusa. Doktorska disertacija*. Koper: Univerza na Primorskem, Fakulteta za humanistične vede.
- TOPORIŠIČ, Jože, 2000: *Slovenska slovnica*. Maribor: Obzorja.

## Vir

- Mat II. = BAUDOUIN DE COURTENAY, Jan, 1904: *Materialen zur südslavischen Dialektologie und Ethnographie II. Sprachproben in den Mundarten der Slaven von Torre in Nordost-Italien*. Sankt Petersburg: Buchdruckerei der Kaiserlichen Akademie der Wissenschaften.

Besedilo je bilo pripravljeno z vnašalnim sistemom ZRCola, ki ga je na ZRC SAZU razvil Peter Weiss.