



UDK 811.163.42'367.622

Izvorni znanstveni članak

Prihvaćen za tisak 1. 6. 2022.

<https://doi.org/10.29162/jez.2022.2>

**Tomislava Bošnjak Botica<sup>1</sup>**

**Jurica Polančec<sup>2</sup>**

**Robert Sviben<sup>2</sup>**

<sup>1</sup> Institut za hrvatski jezik i jezikoslovlje

<sup>2</sup> Filozofski fakultet Zagreb

# Korpusno istraživanje hrvatskih imenica s dugom i kratkom množinom\*

Ostvarenje dvostruke množine jednoga dijela dvosložnih i višesložnih imenica muškoga roda jedno je od zanimljivijih obilježja hrvatskoga imeničkoga sustava. U radu se na temelju korpusa analiziraju oblici tzv. duge i kratke množine 26 imenica muškoga roda podijeljenih u tri skupine: 1) dvosložne i trosložne imenice s nepostojanim *a*; 2) dvosložne imenice s kratkosilaznim naglaskom i dužinom; 3) složenice kojima je druga sastavnica jednosložna imenica. Izneseni podatci upućuju na šarolikost same pojave, ali i neujednačenost rječničkoga opisa te se iznose prijedlozi kako taj opis učiniti bližim stvarnoj jezičnoj uporabi odraženoj u korpusnim podacima. Uočavaju se i zanimljive oscilacije u zastupljenosti jednih ili drugih oblika s obzirom na konkretni padež te mogućnost nastanka mješovite množinske paradigme. Posebna se pozornost posvećuje metodologiji korpusnoga istraživanja ove pojave te kriterijima za stupnjevanje dubletnosti kroz tzv. pojase dubletnosti.

**Ključne riječi:** dvostruka množina, korpus, pojasi dubletnosti, mješovita množina, hrvatski jezik

\* Rad je nastao u sklopu projekta *Baza hrvatskih morfoloških dubleta*, a istraživanje je financirano projektom *Feast and Famine: Confronting overabundance and defectivity in language* (AH/T002859/1, Ujedinjeno Kraljevstvo). Autori zahvaljuju anonimnom recenzentu i urednicima na izrazito korisnim uputama za poboljšanje rada.

## 1. Uvod

Po dobro poznatu pravilu imenice muškoga roda a-sklonidbe u hrvatskom standardnom jeziku tvore množinu na jedan od dvaju načina. Jednosložne imenice tako pretežito tvore tzv. dugu množinu, u kojoj ispred množinskoga padežnoga nastavka u svim padežima dolazi sufixs *ov/ev*, poznatiji i kao množinski proširak, umetak ili formant. Pravilo je relativno čvrsto te imamo primjere kao što su *čepovi*, *ključevi*, *stolovi* itd., uz određen broj dobro dokumentiranih odstupanja (npr. *dani*, *konji*, *psi*) te nešto primjera dvostrukih oblika (npr. *brci/brkovi*, *druzi/drugovi*, *snijezi/snjegovi*), obično uvjetovanih kakvom funkcionalnom razlikom. Dvosložne i višesložne imenice ne podliježu tomu pravilu te množinu tvore jednostavnim dodavanjem množinskih nastavaka na osnovu (tzv. kratka množina): *mjeseci*, *papiri*, *prozori*, *dokumenti*, *početnici*, *istraživači* itd.

Ipak, stanovit broj dvosložnih imenica te neke složene odstupaju od spomenutoga pravila na način da dopuštaju i kratku i dugu množinu, pri čemu govorimo o dvostrukoj množini. Poznati su primjeri *pauci* i *paukovi*, *golubi* i *golubovi*, *slučaji* i *slučajevi* itd. Manji dio dvosložnih imenica odabirom jednoga ili drugoga množinskoga oblika ostvaruje značenjsko, stilsko ili područno razlikovanje, npr. u slučaju kratke množine *posli* u odnosu na *poslovi*.<sup>1</sup> Zabilježene su i dvosložne imenice koje po svemu sudeći tvore isključivo dugu množinu, npr. *mozak*, *ovan*, *dojam*, *pojam*, *zajam* itd.

Tema su ovoga rada upravo dvosložne i višesložne imenice koje mimo pravila uz očekivanu kratku množinu mogu tvoriti i množinu proširenu sufixsom *-ov/-ev-*. Ograničavamo se na imenice kojima dvostruke množinske oblike ne možemo (barem ne prije provedenoga istraživanja) dovesti u vezu s razlikom u značenju.

Središnji je koncept ovoga rada morfološka dubletnost (dvostrukost), a dvostruka je množina jedan od njegovih ostvaraja. Prema tome oblike poput *pauci* i *paukovi* nazivamo dubletama ili dvostrukim oblicima jer za isto značenje (funkciju) – nominativ množine – raspoložemo dvama oblicima. Izbor između dvaju oblika pritom nije uvjetovan nekim očitim funkcionalnim ili drugim čimbenikom.<sup>2</sup>

Istraživanje je zamišljeno kao pilot-studija u kojoj ćemo detaljno opisati mali uzorak imenica s dvostrukom (kratkom i dugom) množinom. Cilj je takva istraživanja korpusnom pretragom provjeriti koje imenice preferiraju duži, a koje kraći oblik, kakvi su omjeri među pronađenim potvrdama, ima li njihova distribucija veze s vrstom korpusa ili morfološkim ustrojem samih imenica te usporediti tako dobivene informacije s opisom u četirima rječnicima suvremenoga hrvatskoga jezi-

<sup>1</sup> HJP tako uz kratku množinu *posli* navodi odrednice *arhaično* i *regionalno*, a Hamm (1967: 70) dobro uočava da kratka množina ima „više preneseno, apstraktno pa i pejorativno značenje prema poslovi”.

<sup>2</sup> Postoje u hrvatskome i trostruki oblici, na primjer u genitivu množine nekih imenica e-sklonidbe (npr. *tréšnja*: *tréšanjā/tréšnjā/tréšnji*; izložba: *izložābā/izložbā/izložbi*).

ka.<sup>3</sup> U konačnici želimo razmotriti načine kako se relativno kompleksna slika koju daju korpusni podatci može vjernije prenijeti u rječnike i druge jezične priručnike. Drugim riječima, kako se dublete u rječnicima mogu predstaviti na način koji bolje odražava stvarnu jezičnu uporabu.

Rad je ustrojen na sljedeći način. Najprije se u § 2 opisuje istraživanje kroz građu i vrela. Zatim se u § 3 daje pregled literature relevantne za istraživanje: u § 3.1 predstavljaju se spoznaje o fenomenu duge i kratke množine u novijem (20-stoljetnom) standardu na temelju postojeće literature o hrvatskom jeziku, a u § 3.2 tako opisan fenomen dvostruke množine smješta se u širi teorijski okvir rasprave o dubletnosti i morfološkom preobilju. Budući da analiza postojeće literature pokazuje da je na temelju intuicije i jezičnoga znanja teško odrediti prihvatljivost mnogih imeničkih oblika, u radu pristupamo korpusno utemeljenu opisu uzorka od 26 imenica. Prije toga u § 4 podrobnije govorimo o dvostrukoj množini kao istraživačkom problemu, ponajprije o ulozi korpusa u ovom istraživanju, s posebnim osvrtom na ulogu korpusa u leksikografiji (§ 4.1), a potom u § 4.2 govorimo o pojasi dubletnosti, odnosno o jednom od načina kvantificiranja prihvatljivosti dubletnih oblika različitih imenica. Nakon temeljita razmatranja metodologije u § 5 predstavljamo rezultate korpusne pretrage: u § 5.1 za imenice s nepostojanim *a*, u § 5.2 za imenice s KS naglaskom i zanaglasnom dužinom te u § 5.3 za složenice. Rezultati se sažimaju i komentiraju u § 5.4. Rad se zaključuje u § 6 raspravom o rezultatima i nacrtom nastavka istraživanja.

## 2. Opis istraživanja

### 2.1. Građa

Istraživanje se temelji na uzorku od 26 imenica raspoređenih u tri skupine. Uzeto je po devet imenica iz skupine imenica s nepostojanim (sekundarnim, umetnutim) *a* i skupine imenica s kratkosilaznim (KS) naglaskom i zanaglasnom dužinom te osam imenica iz skupine složenica (tablica 1). Podjela na tri skupine imenica prilagođena je na temelju postojeće literature (npr. Orlandi 1963, Samardžija 1987).

Tablica 1. Razdioba imenica s dvostrukom množinom analiziranih u ovom radu

skupina imenica	imenice
imenice s nepostojanim <i>a</i>	<i>bljesak, gležanj, manjak, palac, pijetao, pladanj, pljusak, pucanj, stupanj</i>
imenice s KS naglaskom i zanaglasnom dužinom	<i>galeb, kesten, korijen, ležaj, pauk, plamen, pojas, prsten, slučaj</i>
složenice	<i>drvored, minibus, nogostup, parobrod, prototip, stereotip, suhozid, velegrad</i>

<sup>3</sup> Veliki rječnik hrvatskoga standardnog jezika (VRH), Rječnik hrvatskoga jezika (RHJ), Školski rječnik hrvatskoga jezika (ŠRHJ) i Hrvatski jezični portal (HJP).

Imenice su odabrane s popisâ imenica s dvostrukom množinom koje daju Samardžija (1987), Orlandi (1963), Lečić (2016) i još neka vrela. Pregledom tih popisa suzili smo izbor na 30-ak kandidata za svaku skupinu tako što smo isključili zastarjelice i imenice koje zbog svojih značenjskih obilježja u načelu ne dolaze u množini, npr. *značaj*. Dvosložne imenice za uzorak (tablica 1) odabrane su prema kriterijima posvjedočenosti obiju množina u rječnicima suvremenoga jezika<sup>4</sup> te zastupljenosti različitih sufikasa. Imenica *pijetao* ušla je u popis kao predstavnik skupine imenica s prelaskom završnoga *l* u *o*, koje su zbog te promjene u N. jd. prividno trosložne. U prvoj skupini imenica česti su sufiksi *-(a)nj* (*-nj-ø*) i *-(a)k* (*-k-ø*), dok su za drugu skupinu karakteristični sufiksi *-en* te u nešto manjoj mjeri *-aj*.<sup>5</sup>

Imenice iz treće skupine odabrane su većim dijelom tako da su za njih u anketi koja je prethodila istraživanju potvrđene obje množine. U tu smo skupinu ciljano uvrstili klasične složenice sa spojnikom (*drvored*, *nogostup*, *suhozid*), imenice s prefiksoidima, bilo stranim (*minibus*, *prototip*)<sup>6</sup> bilo domaćim (*velegrad*), te jednu prividnu složenicu (*stereotip*). Treba napomenuti da prema rječnicima za klasične složenice postoji malo naznaka da mogu tvoriti dugu množinu (što će se i potvrditi korpusnim podacima u § 5.3), ali su svejedno uvrštene u istraživanje da bi se omogućila usporedba s drugim tvorbenim tipovima koji su potencijalno skloniji dugoj množini.

## 2.2. Vrela

Kao izvor podataka na kojima temeljimo korpusni opis dubletnih paradigama odabrani su korpusi *Riznica* (*Hrvatski jezični korpus*) i potkorpus *forum.hr* mrežnoga korpusa *hrWaC*, osim za skupinu složenica, gdje je zbog premalo primjera u obzir uzet cijeli *hrWaC*. Osnovne podatke o obama korpusima daje tablica 2, a u nju su uvršteni i podatci za potkorpus *forum.hr*.<sup>7</sup>

<sup>4</sup> Od toga se načela odstupilo u slučaju imenice *manjak*, koja prema rječnicima može tvoriti samo kratku množinu, ali je preliminarna korpusna pretraga potvrdila i dugu množinu te je ta imenica uvrštena u uzorak.

<sup>5</sup> Sufikse radi lakšega praćenja bilježimo tradicionalno, bez nultih morfova, slijedeći npr. Babića (2002). Sukladno tomu nepostojani *a* koji se javlja uz neke sufikse bilježi se zagradama, a u zapisu sufikasa ne bilježi se zanaglasna dužina.

<sup>6</sup> Naravno, u navedenim je primjerima posrijedi izravno preuzimanje leksema, a ne domaća tvorba prefiksoidom.

<sup>7</sup> Za više detalja o korpusu *Riznica* v. Čavar i Brozović Rončević (2012). O kompiliranju *hrWaC-a* v. više u Ljubešić i Klubička (2014).

Tablica 2. Osnovni podatci o korpusima *Riznica* i *hrWaC* te o potkorpusu *forum.hr*

korpus	veličina (pojavnice)	starost tekstova	zastupljeni tekstovi
<i>Riznica</i> (Hrvatski jezični korpus)	101 782 863	od početka 20. stoljeća do ranih 2000-ih	književni, stručni i novinski tekstovi
hrWaC v2.2	1 397 757 548	21. stoljeće (do 2017. godine)	tekstovi na hrvatskom jeziku prikupljeni s vršne domene .hr
forum.hr (potkorpus hrWaCa)	241 694 709 (17,29 % hrWaCa)	21. stoljeće (do 2017. godine)	tekstovi na hrvatskom jeziku prikupljeni s poddomene forum.hr

Korpusni su izvori odabrani kao predstavnici standardnoga (*Riznica*) odnosno nestandardnoga (razgovornoga)<sup>8</sup> jezika (*forum.hr*) s ciljem da se utvrde eventualne razlike u statusu dvostrukih množinskih paradigama u tim dvama varijetetima hrvatskoga jezika. Korpus *Riznica* odabran je kao predstavnik standardnoga hrvatskoga jezika jer sadržava tekstove iz svih njegovih funkcionalnih stilova (osim razgovornoga<sup>9</sup>). On također sadržava i tekstove iz cijeloga 20. stoljeća iako pretežu tekstovi s njegova kraja, a stariji tekstovi gotovo isključivo pripadaju književnoumjetničkom stilu. Nasuprot njemu potkorpus *forum.hr* mrežnoga korpusa *hrWaC* odabran je kao predstavnik suvremenoga nestandardnoga hrvatskoga. Mrežni korpus *hrWaC* kao cjelina vrlo je raznolik te sadržava i tekstove koji pripadaju i standardnom i nestandardnom hrvatskom. Stoga je izdvojen potkorpus *forum.hr*, odnosno dio *hrWaC*-a s poddomene *forum.hr* koji čini 17,29 % *hrWaC*-a. Budući da su njegov sadržaj nestandardni jezični podatci u obliku neelektoriranih forumskih rasprava, on pruža uvid u nenadziranu i relativno spontanu jezičnu proizvodnju u čijem stvaranju sudjeluje velik broj korisnika iz cijele Hrvatske.

### 3. Dosadašnje spoznaje i teorijski okvir

U ovom se odjeljku daje deskriptivna i teorijska pozadina rada.

#### 3.1. Postojeći opisi

Pojava dvostrukih množinskih oblika prepoznata je u hrvatskim gramatikama odavno. Takve su imenice navodili već Babukić (2014 [1854]) i Veber Tkalčević (1876).

<sup>8</sup> Nestandardni ili razgovorni hrvatski možemo definirati kao općenacionalni varijetet hrvatskoga jezika koji u svojoj osnovi ima standardni jezik, ali „bez knjiških i strogo intelektualnih elemenata” (Brozović 2016: 394), a koji istovremeno sadržava brojne sebi svojstvene elemente (supstandardizme). Ti nestandardni elementi mogu biti prisutni na cijelom hrvatskom jezičnom području ili ograničeni na pojedine regije. U takvu viđenju razgovornoga jezika slijedimo Brozovića (2016: 385–402; usp. i Pasini 2018: 587–593; Škarić 2007: 24).

<sup>9</sup> Razgovorni stil *standardnoga* jezika redovito se navodi u literaturi iako je njegovo postojanje u praksi sporno jer se „razgovorni stil standardnog jezika rijetko ostvaruje i rijetko čuje” (Francić, Hudeček i Mihaljević 2005: 242).

Kasnije gramatike (Maretić 1899; 1963; Brabec, Hraste i Živković 1954; Babić i dr. 1991; Babić i dr. 2007; Raguž 1997; Barić i dr. 1997; Silić i Pranjković 2005) slijede sličnu praksu, a razlikuju se izborom i brojem primjera. Sažeto se o podacima iz gramatika može reći da potvrđuju postojanje dviju množina u imenica muškoga roda, da je duga ograničena uglavnom na jednosložne i dvosložne imenice te da neke imenice imaju samo kratku, neke samo dugu, a neke pak imaju oba oblika (usp. Samardžija 1987: 129).

Stariji i suvremeni rječnici dobro dokumentiraju pojavu, kombinirajući povijesno posvjedočene dublete sa suvremenima te, premda nesustavno, u morfološkom opisu navode oba oblika.

Temom dviju množina u proteklih se sedamdesetak godina bavilo nekoliko autora. O dugoj množini trosložnih i višesložnih imenica pisao je Klaić (1953). Njega su ponajprije zanimale akcenatske dublete među tim oblicima.

Zanimljiv i višestruko koristan prikaz povijesti duge množine donosi Orlandi (1963). Poznato je da je nastanak duge množine sekundaran u odnosu na kratku i da je isprva obuhvaćao ograničen dio imeničkoga sustava te se postupno širio, ne samo među jednosložnim nego i među dvosložnim imenicama muškoga roda, odnosno a-sklonidbe. U navedenom se radu pokazuje kako je išao smjer širenja duge množine s istočnoga štokavskoga na zapadni, a zaključci su relevantni i za današnje propitivanje nekih normativnih rješenja te predviđanje budućega smjera širenja duge množine na imenice s više od jednog sloga.

Gudkov (1964) piše o kolebanju množinskih oblika, ali u jednosložnih imenica muškoga roda, čemu pretpostavlja nekoliko čimbenika: dijalektna obilježja pisca/govornika, stilska obilježja teksta i sintaktički kontekst.

Najiscrpnije je dosad o ovoj temi pisao Samardžija (1987), koji nastoji pokazati koje imenice muškoga roda imaju samo kratku množinu, koje samo dugu, a koje obje. Premda Samardžija navodi da je odgovor na pitanje kad koji oblik dolazi potrebno ponajprije stranim govornicima hrvatskoga jer izvorni govornici „po osjećaju mogu rasuditi kad kako”, držimo da u mnogim primjerima, kad je riječ o pravim dubletama, to nije posve jasno ni izvornim govornicima. Od dvosložnih imenica izdvaja njih 114 koje mogu imati dugu množinu (što je tek oko 3 % ukupnoga broja dvosložnih imenica s njegova popisa), od kojih su oko dvije trećine (80) imenice s nepostojanim *a* (npr. *bubanj*, *gležanj*, *sajam*, *šiljak*, *vosak*, *žrvanj*).<sup>10</sup> Ostale određuje kao *imenice bez nepostojanog a* (npr. *jablan*, *kotar*, *kulen*, *pauk*), a mahom su posrijedi imenice s kratkosilaznim naglaskom i dužinom uz neke iznimke.<sup>11</sup> Sedam

<sup>10</sup> Gledajući sav popis, imenice s nepostojanim *a* većinom tvore kratku množinu, među njima je udio onih s dugom oko 22 % (Samardžija 1987: 133).

<sup>11</sup> Samardžija u tu skupinu uključuje i posuđenice sa suglasničkom skupinom na kraju: *dubl*, *singl*, *titl*, *vergl*, žanr te još poneku imenicu s drukčijim naglasnim obilježjem (1987: 34).

imenica s nepostojanim *a* ostvaruje samo dugu množinu, a u drugoj skupini takvih je dvanaest, pa je ukupno uzevši u Samardžije 95 imenica s mogućnošću jedne i druge množine (72 u prvoj te 22 u drugoj skupini). Među višesložnim imenicama s dugom množinom izdvaja one s prelaskom završnoga  $l > o$  (npr. *kotao*, *orao*, *posao*) te spominje *tvorenice* kojima je druga sastavnica riječ koja i sama može imati dugu množinu (npr. *polukat*, *veleum*) i čiji je popis otvoren.

Gledano po našim trima skupinama, ti se podatci mogu ukratko sažeti u tablici 3.

Tablica 3. Podatci o imenicama s dvostrukom množinom prema Samardžiji

skupina imenica	Samardžija 1987
imenice s nepostojanim <i>a</i>	72
imenice s KS naglaskom i zanaglasnom dužinom	22 (25) <sup>12</sup>
složnice	otvoren popis

U skupini imenica s nepostojanim *a* pojavljuju se sufiksi *-(a)c*, *-(a)k*, *-(a)m*, *-(a)nj* i *-(a)r*, dok se u skupini imenica s KS naglaskom mogu izdvojiti samo sufiksi *-aj* i *-en*.

Samardžijin je uvid izrazito koristan kao polazište s obzirom na količinu građe i temeljitost klasifikacije, ali ipak manjkav za prikaz suvremenoga stanja zbog naravi korištenih vrela i tada dostupnih istraživačkih metoda. Na dvama se mjestima spominje anketiranje (1987: 131, 136), ali se o tome ne donose nikakvi konkretniji podatci, stoga ne znamo koliko je iznesena klasifikacija ishod isključivo rječničke građe, a koliko je na nju bilo utjecaja jezične intuicije ili provjere među govornicima. Nije zanemariva ni vremenska dimenzija jer se naime pokazuje da iz današnje perspektive a) neke imenice s popisa nemaju dvostruku množinu koja im se pripisuje (npr. *kabel*, *metar*, *papar*), b) neke su zastarjele (nestandardne?) te je o njima nemoguće imati intuiciju ili za njih naći potvrde (npr. *ajam*, *čišanj*, *ramen*), c) neke nedostaju (npr. *slučaj*, *sokol*, *vitez*).

O korelaciji značenja i oblika duge i kratke množine pisao je Šipka (1999: 169–170). Težište je toga rada na značenjskim razlikama koje uvjetuju odabir jedne ili druge paradigme.

Marković (2012: 274) uglavnom potvrđuje prethodno iznesene spoznaje dopunjujući ih i novim primjerima. Od dvosložnih imenica koje mogu imati i nemati *-ov-*/*-ev-* navodi neke imenice s nepostojanim (sekundarnim) *a* te imenice naglasnoga

<sup>12</sup> Samardžija navodi *kläun*, *körjen* i *külen* bez zanaglasne dužine, no ona se u novoštokavskom često ostvaruje (za *körjēn* je bilježe i rječnici).

tipa *kāmēn – kāmēna*. Navodi da i neke trosložne i višesložne imenice mogu i imati i nemati *-ov-/-ev-*, točnije imenice kojima je treći slog nastao vokalizacijom završnog *l* (npr. *djetao, kotoa*) te složenice. Na primjeru nekih novoposuđenica (*SMS-ovi, selebovi*) pokazuje da je *-ov-/-ev-* jak signal množinskoga oblika.

Lečić (2016) u svojoj empirijski orijentiranoj studiji o dubletnosti u sustavu imenskih riječi raspravlja i o pojavi duge i kratke množine, koja se obrađuje na dva mjestima. Prvo se raspravlja općenito o dubletnosti u množinskim paradigma te se predstavljaju rezultati korpusne pretrage oblika za 36 jednosložnih, dvosložnih i višesložnih imenica (Lečić 2016: 84–88). Pretraga je obavljena u korpusu Riznica s ciljem da se uspoređi ponašanje istraženih imenica u književnom i neknjiževnom dijelu korpusa uz pretpostavku da će kratki oblici množine biti frekventniji u književnom dijelu Riznice. Autor dolazi do zaključka da žanr (književno prema neknjiževnom) ne objašnjava ponašanje imenica, već da objašnjenja za preferencije prema jednom ili drugom tipu množinske paradigme treba tražiti na razini pojedinih leksema.

Na drugom mjestu u radu Lečić uspoređuje ponašanje drugoga uzorka od 36 imenica s dugom i kratkom množinom u korpusu s rezultatima ankete o prihvatljivosti tih istih oblika među dvjestotinjak izvornih govornika hrvatskoga jezika (Lečić 2016: 164–177). Rezultati u općim crtama pokazuju da prihvatljivost među anketiranim govornicima korelira s ponašanjem imenica u korpusu (Lečić 2016: 172–173). Konkretnije, najbolje su ocijenjeni oblici duge množine onih imenica koje se već u korpusu pojavljuju ponajprije u dugoj množini, a isto vrijedi i za visoko ocijenjene oblike kratke množine. Obratan slučaj također vrijedi: dugi odnosno kratki oblici s niskim ocjenama vrlo su rijetki u korpusu. Oblici sa srednjim – ni visokim ni niskim – ocjenama pripadaju imenicama čiji su dubletni množini oblici otprilike jednakih udjela u korpusu.

Na temelju postojećih istraživanja spoznaje o dvostrukoj množini dvosložnih imenica muškoga roda u hrvatskom jeziku možemo svesti na ove tvrdnje:

- dvosložnih imenica s dugom i kratkom množinom ima oko devedeset
- rječnici se razlikuju s obzirom na navođenje obaju oblika
- za najveći dio imenica ne postoji jasna funkcionalna uvjetovanost jednoga ili drugoga oblika
- uočavaju se moguće tendencije s obzirom na fonološka, naglasna i tvorbeno obilježja (kategorizacija u skupine).

Svjedočimo također svojevrsnu konsenzusu u literaturi koji sugerira da se o uporabi duge i kratke množine mnogih imenica ne može postaviti čvrsto pravilo te od te pretpostavke polazimo u ovom radu.



### 3.2. Teorijski okvir

Postojanje dvostruke množine i dvostrukih oblika nije zanimljivo samo iz kroatiističke (deskriptivne i preskriptivne) perspektive nego i općenitije teorijski gledano intrigira lingviste stoga što se suprotstavlja uobičajenu načinu funkcioniranja jezika koji podrazumijeva ekonomičnost. Po načelu ekonomičnosti očekuje se da jednom obliku odgovara jedno značenje, odnosno jedna gramatička funkcija, čime se sprječava pojava više oblika istoga značenja. Nekoliko je naziva i tumačenja te pojave, najpoznatiji je tzv. učinak blokiranja ili zapreke, odnosno načelo morfološkoga blokiranja kojim postojanje jednoga oblika sprječava pojavu drugoga oblika iste funkcije (Aronoff 1976, a slično gledište zastupaju i drugi autori, npr. Carstairs-McCarthy 1987; Clark 1987; Kroch 1994). Velika većina jezičnih jedinica podliježe takvu uređenju (*konj* > *konji*/\**konjevi*; *ovan* > \**ovni/ovnovi*), pa će se postojeći dvostruki oblici obično povezivati s kakvom funkcionalnom razlikom (*sat* > *sati* ‘odsječci vremena od 60 minuta’ prema *satovi* ‘sprave za mjerenje vremena’). Međutim, iz vlastita jezičnoga iskustva i dostupnih jezičnih opisa znamo da postoje i dvostruki oblici. Dvostruka množina jedan je od takvih primjera u hrvatskome, ali nije jedini (*vježba* > G mn. *vježba/vježbi*; *kuhar* > I jd. *kuharom/kuharem*; *skitati* > prez. *skitam/skićem*; *dug* > komp. *duži/dulji* itd.). U jezicima koji su predmet najbrojnijih istraživanja, kakav je primjerice engleski, ta je pojava rijetko zastupljena (npr. nekoliko neprilagođenih imenica preuzetih iz latinskoga koleba se u množini, pokoji glagol ima mogućnost dvaju participa), u romanskim su jezicima također posrijedi izolirani primjeri, no slavenski su jezici, naprotiv, plodno tlo za brojne potvrde dvostrukosti cijelih paradigama ili njihovih dijelova.

Obilje pronađenih primjera morfoloških dvostrukosti dovelo je u posljednje vrijeme do preispitivanja mjesta ekonomičnosti, odnosno učinka blokiranja u suvremenim morfološkim pristupima. Slijedeći postavke kanonske tipologije (Corbett 2009), Thornton je razvila teoriju o morfološkom *preobilju* (engl. *overabundance*), odnosno supostojanju dvaju ili više oblika jedne riječi istoga gramatičkoga značenja, pri čemu upotreba jednoga ili drugoga (ili trećega) oblika nije uvjetovana nikakvim kontekstom (Thornton 2011). Riječ je o nekanonskoj pojavi, odnosno otklonu od kanoničnosti kao jedinstvenoga preslikavanja s oblika na funkciju i s funkcije na oblik jer dva ili više oblika mogu ispuniti isto mjesto (ćeliju, engl. *cell*) u paradigmi.

Morfološko preobilje kao teorijski koncept u načelu odgovara onomu što mi zovemo morfološkom dubletnošću i ta dva naziva ovdje smatramo istovrijednima. Osim preobilja i dubletnosti u literaturi se može naići i na termin *konkurentski* oblici (engl. *rival forms*, *competing forms*) (npr. Baayen i dr. 2013; Rainer i dr. 2019).

Istraživanja morfološkoga preobilja u jeziku intenzivirala su se posljednjih desetak godina, ali je težište još uvijek više na tvorbenoj nego flektivnoj morfologiji. Premda je pojedinačnih radova, kao što smo vidjeli, bilo i prije, o pojavi dvostrukih

oblika u hrvatskom u sklopu najnovijih teorija najprije je pisao Lečić (2015), analizirajući instrumental jednine imenica muškoga roda, a potom i ostale kategorije dubletnosti u imeničkom sustavu (Lečić 2016). Slijedili su i drugi radovi (npr. Bošnjak Botica i Hržica 2016; Bošnjak Botica, Budja i Hržica 2018; Bošnjak Botica i dr. 2019), kojima se pokazalo da je broj dubletnih odnosno preobilnih kategorija u hrvatskom izrazito velik – oko 40, a obuhvaćaju imenice, pridjeve (i priloge), glagole i zamjenice.<sup>13</sup>

Citirana su istraživanja potakla izradu Baze hrvatskih morfoloških dubleta, u kojoj se trenutačno nalazi oko 8000 unosa. Cilj joj je popisati i skupiti reprezentativan broj dvostrukih oblika preuzetih iz postojećih korpusa hrvatskog jezika kako bi se omogućilo istraživanje i opis same pojave te vidjelo njezino mjesto u ustroju jezičnoga sustava. Time se bitno olakšava pristup relevantnim podacima i provođenje metodološki jasno utemeljenih istraživanja.

Ovdje predstavljena studija pokušat će pokazati svrhovitost takva resursa. Naime, kao što ćemo vidjeti u pregledu rezultata svoga istraživanja (§ 5), postojeći rječnici suvremenoga jezika nude šarolike, a prema tome i manje pouzdane podatke o množinskim oblicima dvosložnih imenica. Stoga da bismo došli do podataka o tome koje imenice doista jesu dubletne u množini, posežemo za presjekom uporabne prakse koju nam pružaju veliki jezični korpusi. Bazu hrvatskih morfoloških dubleta pritom zamišljamo kao repozitorij korpusnih podataka koji na različite načine može poslužiti boljemu leksikografskom opisu množinskih oblika dvosložnih imenica.

## 4. Istraživački problem: rječnik i korpus

U ovom ćemo odjeljku reći nešto o odnosu korpusa i rječnika i ulozi korpusa u leksikografskom opisu dubletnih oblika.

### 4.1. *Treba li rječniku korpus i zašto*

Tradicionalni pristup opisu dubleta oslanja se ponajprije na rječničke potvrde (npr. Maretić 1963; Samardžija 1987),<sup>14</sup> što otvara pitanje izvora za potvrde koje nalazimo u rječnicima. Katkad su ti izvori relativno jasno specificirani,<sup>15</sup> katkad takva

<sup>13</sup> Dio radova bavio se pojedinim dvostrukim oblicima ili kategorijama bez uključivanja u spomenuti teorijski okvir (npr. Polančec 2017).

<sup>14</sup> Samardžija se primjerice oslanja na rječnike dviju Matica uz nespacificirana „dodatna anketiranja” (1987: 131). Slično tomu Maretić (1963: 143–146) kao svoje izvore navodi cijeli niz rječnika (Vukov, Ivekovićev, donekle i Akademijin) te jedan Daničićev rad.

<sup>15</sup> Primjerice, izbor građe za Ivekovićev i Brozov rječnik, o čemu se mnogo pisalo (v. npr. Mićanović 2006: 80–84).

specifikacija nedostaje.<sup>16</sup> U posljednje vrijeme leksikografska praksa slijedi suvremene trendove te se kao izvor građe navode postojeći tiskani rječnici, ali i elektronički korpusi.<sup>17</sup> Dapače, često se kao imperativ uzorna suvremenoga rječnika navodi upravo njegova utemeljenost na korpusu (Štrkalj Despot i Möhrs 2015: 346). O odnosu korpusa i leksikografije u hrvatskome opširnije piše Filipović Petrović (u tisku) s polazištem da suvremena deskriptivna leksikografija želi u korpusu doznati kako se jezik uobičajeno koristi. Napominje da je hrvatska leksikografija u tom smislu mnogo sklonija autoritetu pojedinca ili autoritetu jezika književnosti nego autoritetu jezičnoga dokaza, odnosno stvarne jezične uporabe kakva se ogleda u korpusu. Drži da je takav pristup povezan sa shvaćanjem rječnika kao mjesta na kojem će se poštivati standard i dobra upotreba u skladu s propisanom normom.

Iako se to nigdje izrijekom ne navodi, može se pretpostaviti da je u tradicionalnom leksikografskom opisu dubleta često bio presudan stav obrađivača, koji se dijelom oslanja na vlastitu intuiciju, a dijelom na iskustvo s jezičnom građom i pregledanim izvorima (npr. imenica *pauk* na HJP-u ima oblike u N mn. *pauci* te u G mn. *pauka/paukova*, dok ŠRHJ navodi samo oblike kratke množine). Sve to znači da su podatci o dubletama koje nalazimo u rječnicima uvjetovani nizom čimbenika: navodima u postojećim rječnicima, potvrđama u korpusima te nesumnjivo intuicijom obrađivača. K tomu, rječnici naravno ne mogu pružiti informaciju o brojčanom odnosu dubleta.<sup>18</sup> Dakle, uza sve prednosti rječnik je relativno nesavršen odraz uporabne prakse jer se u njemu presijecaju tradicija i inovacija te korpusno utemeljene potvrde i individualna intuicija. Problem je dakle nepreciznost (u kvantitativnom smislu), nemogućnost da razlučimo suvremeno stanje od ranijih razdoblja (često i ne pretjerano udaljenih) te parcijalnost u smislu da se rječnik katkada oslanja na intuiciju jednoga govornika (obrađivača). Ti se nedostaci rječnika pokušavaju ispraviti uvođenjem korpusno utemeljenih opisa. Premda su po mnogima dobro uravnoteženi jezični korpusi suvremenoga jezika vrlo relevantan pokazatelj njegove stvarne uporabe, ni oni nisu bez nedostataka (usp. Blagus Bartolec i Matas Ivanković 2019) te je nuždan oprez u interpretaciji rezultata.

Da se jezični opis s vremena na vrijeme treba revidirati, potvrđuju i razlike uočene na temelju Samardžijina popisa u odnosu na suvremene rječnike. Primjerice, Samardžija (1987: 133–134) navodi da dvostruku množinu imaju imenice *toranj* i *jelen* iako naši suvremeni rječnici bilježe samo dugu množinu za prvu imenicu (*tornjevi*),

<sup>16</sup> RHJ i *Veliki Anić* (koji je jezgra HJP-a) nigdje ne navode izvore za građu.

<sup>17</sup> Vidi primjerice predgovore u VRH-u i ŠRHJ-u te upute za budući *Hrvatski mrežni rječnik – Mrežnik* (Hudeček i Mihaljević 2018).

<sup>18</sup> Okvirna razlika u prihvatljivosti dubleta može se eventualno iščitati iz leksikografskih odrednica (npr. odrednica *knjiš*. 'knjiški' uz kratku množinu *vitezi* na HJP-u) ili poretkom navođenja množinskih oblika, npr. za imenicu *korijen* na HJP-u, gdje se prvo navodi po svemu sudeći običnija množina *korijeni*, a potom množina *korjenovi*.

a za drugu samo kratku (*jeleni*). Takvi su primjeri ipak manji problem jer se često mogu isključiti na temelju pregleda suvremenih rječnika, a redovito se podudaraju i s jezičnim osjećajem govornika suvremenoga jezika.

U tom su smislu mnogo kompleksniji slučajevi imenica koje se u suvremenim rječnicima navode kao dublete, a za koje među govornicima izostaje intuicija o tome koji je oblik prihvatljiviji ili jednostavno češći. Ovdje naravno ne govorimo o parovima kao što su *prozori* – \**prozorovi*, gdje nema ni govora o dubletnosti ili *posli* – *poslovi*, gdje postoji konsenzus o razlici u značenju i/ili stilskoj obilježenosti. Umjesto toga ponajprije je riječ o parovima poput *korijeni* i *korjenovi*, *pojasi* i *pojasovi/pojasevi*, gdje zacijelo postoji određena preferencija, ali koju je katkad teško specificirati. Problem se samo produbljuje predugim razmatranjem primjera, a čitatelji se i sami mogu uvjeriti da će na prvu vjerojatno preferirati oblik *korijeni* ili *pojasevi*, ali da nakon razmišljanja više neće biti sigurni kojemu dati prednost. To je zanimljivo stoga što u stvarnoj uporabi, odnosno u govoru, govornici uvijek moraju odabrati jedan od oblika i to rade nesvjesno.

Naime, može se pretpostaviti da u takvim primjerima intuicija varira od govornika do govornika te dovodi u sumnju već citiranu Samardžijinu tezu da odgovor na pitanje o prihvatljivosti dubletnih oblika „izvornim govornicima hrvatskoga može se i ne činiti bitnim, jer oni ‘po osjećaju’ mogu rasuditi kad kako” (1987: 130). Izvorni govornici zacijelo mogu prosuditi da su im oba oblika prihvatljiva, no ostaje pitanje kad i koliko će koji upotrijebiti. Naime, u stvarnoj ćemo uporabi nerijetko posvjedočiti tomu da se ista imenica u istom odlomku teksta i u istom kontekstu pojavljuje sad u dugoj, sad u kratkoj množini. Kao ilustracija nam može poslužiti imenica *plamen* u ulomku iz romana *Povratak Filipa Latinovicza* Miroslava Krležu:

Plavkasti **plamenovi** gorućeg alkohola oblizivali su njegovo tmasto lice, u sobi bilo je sasvim mračno i sve su stvari potitravale na modrikastom sjaju ognja u stakleno zdjeli na stolu. Baločanski je zurio u to zmijuljasto micanje **plamena** potpuno tupo i nepomično (...).

Jednak je primjer imenice *bagrem* u odlomku iz romana *Kužni pil* Pavla Pavličića:

Na lijevoj strani pogled su zaklanjale grane **bagremova** što su rasli na lesnom obronku, pa se nije moglo vidjeti daleko. Ali, zato su se ispod krošanja – kao da su **bagremi** ondje posađeni zato da bi bili vodiči promatračevu pogledu – nazirali gradski krovovi.

Da bi se primjereno obradilo pitanje imenica o čijim množinskim oblicima govornici nemaju jaku intuiciju ili ona varira, korpus je dobrodošao alat jer nam pomaže da kvantificiramo razliku u prihvatljivosti koju možda intuitivno osjećamo. Korpus prema već uvriježenu mišljenju jest superioran izvor jezičnih podataka u odnosu na procjenu individualnoga govornika (npr. Hunston 2002: 20). Procjena i jezična intuicija individualnoga govornika nisu u svojoj biti problematični, ali ne

mogu biti temelj metodološki ispravno postavljenih lingvističkih istraživanja (za recen-tan pregled vrlo bogate literature o tom pitanju v. Stefanowitsch 2020: 8–17).<sup>19</sup>

Superiornost korpusa očituje se u tome što nam pruža uvid u neku vrstu stvarne jezične uporabe, čime se zaobilazi problem subjektivnosti individualne procjene. Procjena prihvatljivosti individualnoga govornika naravno ne odražava stvarnu jezičnu uporabu zato što se stvarna jezična uporaba rađa iz presjeka jezičnih praksa velikoga broja govornika, gdje pritom na pojedinca utječe uporaba većine govornika u jednoj vrsti povratne sprege.

Na kraju treba naglasiti da jezični korpus, premda jest odraz uporabne prakse, nije istovremeno odraz (ne)prihvatljivosti dubleta. To proizlazi već iz činjenice da dubletne oblike u spontanom govoru nikad ne upotrebljavamo istovremeno, nego uvijek spontano biramo između dvaju oblika, što pak implicira da postoji neka razlika u prihvatljivosti između npr. paradigama *pauci* i *paukovi*, ali ne implicira nužno da je ijedna od njih apsolutno neprihvatljiva govornicima.<sup>20</sup>

Iz svega rečenog proizlazi da za potrebe ovoga rada korpus uzimamo kao odraz jezične intuicije *velike skupine* govornika ostvarene kroz spontane iskaze u stvarnom vremenu. Drugim riječima, za potrebe našega istraživanja korpus je primjeren alat jer nam omogućuje odgovore na dva pitanja: „Hoće li govornici spontano reći *pauci* ili *paukovi*?” (kvantificiranje prihvatljivosti) te „Ja bih rekla/rekao *pucnjevi*, ali je li to istina i za nekoga drugog?” (uvid u presjek govorničkih praksa). U nastavku teksta predstaviti ćemo metodu kvantificiranja prihvatljivosti koja se temelji na brojčanim omjerima potvrda u korpusu.

Postojanje dvaju ili više oblika jednoga značenja oduvijek je i među stručnjaci-ma i među prosječnim govornicima izazivalo potrebu nekakva rangiranja, traženja funkcionalnoga uporišta, a nerijetko i semantičkoga razlikovanja, što često u normativistici rezultira preporučivanjem jednoga od tih oblika. Naš je pristup ponaj-

<sup>19</sup> Ovdje treba istaknuti da je problem istraživanja koja uključuju govorničku procjenu ponajprije u tome što se oslanjaju na procjenu maloga broja stručnih govornika (najčešće samih lingvista). Taj se problem može zaobići na način da se istraživanje temeljeno na procjeni govornika oblikuje tako da uključuje velik broj ispitanika, i to laika (nelingvista, nefilologa). Takva su istraživanja popularna u posljednje vrijeme, primjerice za hrvatski je takvo istraživanje za dubletnost nekih oblika imenica proveo Lečić (2015; 2016). Za druge slavenske jezike relevantni su radovi Bermela i suradnika (npr. Bermel i Knittl 2012a; Bermel, Knittl i Russell 2015; Bermel, Knittl i Russell 2018). U svim se tim radovima mjeri podudarnost podataka iz korpusa s podacima koji se dobivaju izravno od govornika na temelju različitih vrsta upitnika. Usporedba dviju vrsta podataka u tim radovima pokazuje visok stupanj konvergencije između dviju vrsta podataka.

<sup>20</sup> Ovdje dolazi od izražaja još jedna razlika između procjene prihvatljivosti i stvarne jezične uporabe. Naime, procjenjujući prihvatljivost dubletnih oblika – npr. kad se pitamo zvuče li nam bolje *korijeni* ili *korjenovi* – činimo to na način koji ne odražava stvarnu jezičnu uporabu jer u stvarnoj uporabi, kako smo rekli, dubletne oblike uvijek biramo spontano, u trenutku, dok u procjeni oblika često s vremenom postajemo sve manje sigurni u prvotni odabir.

prije deskriptivni, ali ostavljamo otvorenim pitanje trebaju li se podatci o čestoti i uporabi pojedinih oblika ikako odraziti u njihovu (normativnom) leksikografskom opisu.

#### 4.2. Stupnjevi dubletnosti u korpusu: pojasi dubletnosti

Za potpuno razumijevanje načina na koji korpusni podatci doprinose istraživanju fenomena dvostruke množine trebat će dodatno precizirati, odnosno *operacionalizirati* koncept dvostrukosti odnosno dubletnosti, o kojem smo već govorili u uvodu. Ondje smo preliminarno definirali dublete kao paralelne (konkurentne) morfološke oblike (obično dvostruke) u kojih se izbor jednoga ili drugoga oblika ne može objasniti očitim značenjskim ili drugim čimbenicima. Međutim, nisu sve dublete jednake i na to nas upućuju upravo korpusni podatci.

Da bismo se u to uvjerali, promotrit ćemo rezultate pretrage za množinske paradigme imenica *pucanj* i *pauk* iz korpusa *Riznica* u tablici 4.<sup>21</sup>

Tablica 4. Udjeli kratke i duge množine u korpusu *Riznica* za imenice *pucanj* i *pauk*

imenica	kratka množina	duga množina	ukupna frekvencija
pucanj	114 (43,2 %)	150 (56,8 %)	264
pauk	182 (91,9 %)	16 (8,1 %)	198

Za te dvije imenice vidimo suprotan rezultat, za prvu je češća duga množina, za drugu kratka. Međutim, vidimo još jednu važnu razliku, a to je omjer dubletnih paradigama. Za imenicu *pucanj* razlika u korist češće paradigme mnogo je manja nego za imenicu *pauk*, gdje je češća paradigma mnogo dominantnija. Taj nam primjer zorno pokazuje problem prihvatljivosti – govori li nam korpus da je kratka množina imenice *pucanj* nekako prihvatljivija nego što je to duga množina imenice *pauk*? Odgovor na to pitanje nije jednostavno dati, no postojeća su istraživanja razvila način da s pomoću omjera frekvencija približno odredimo položaj slabije posvjedočene dublete, o čemu ćemo sada više reći.

Oslanjamo se na pregled problematike u Bermel, Knittl i Russel (2018: 204–205), gdje nalazimo različite korpusno utemeljene definicije dubletnosti proizašle većinom iz istraživanja na češkoj, u manjoj mjeri i poljskoj građi. Te se definicije oslanjaju na ideju da ovisno o omjeru potvrda dvostrukih oblika imenice možemo svrstati u različite **frekvencijske pojase dubletnosti**.

U ovom ćemo radu prihvatiti definiciju pojasa dubletnosti iz Bermel i Knittl (2012b: 251), po kojoj se imeničke dublete mogu podijeliti u četiri pojasa. Prvi pojas obuhvaća imenice u pojasu do omjera 69 : 31. Drugim riječima, prvi pojas obuhvaća

one parove konkurentskih oblika čiji frekventniji oblik ima udio od najviše 69 % (ali ne više od toga), a manje frekventan ima udio od najmanje 31 % (i ne manje od toga). Autori oblike unutar toga omjera smatraju jednakovrijednima (*equipollent*) i takve oblike ćemo ovdje smatrati pravim dubletama. Uz taj razlikuju još tri pojasa. U parova gdje su omjeri od 70 : 30 do 89 : 11 frekventniji se oblik smatra većinskim, a drugi manjinskim (*majority, minority*). Pri omjeru parova od 90 : 10 do 99 : 01 frekventniji oblik opisuje se kao neobilježen, a drugi kao obilježen (*unmarked, marked*). Oblik koji obuhvaća 99 % i više potvrda smatra se dominantnim (*dominant*), a njegov konkurent izoliranim (*isolated*). Predložene pojase sažima tablica 5.<sup>22</sup>

Tablica 5. Pojasi dubletnosti (prema Bermel i Knittl 2012b: 251)

deskriptor frekventnijega / manje frekventna oblika	udio frekventnijega oblika	udio manje frekventna oblika
jednakovrijedni oblici (prave dublete)	najviše 69 %	najmanje 30 %
većinski / manjinski oblik	70 – 90 %	10 – 29 %
neobilježeni / obilježeni oblik	91 – 99 %	1 – 9 %
dominantni / izolirani oblik	više od 99 %	manje od 1 %

Prema toj podjeli imenica *pauk* po podacima iz Riznice spadala bi u drugi pojas, odnosno njezina bi se kratka množina smatrala većinskim, a duga manjinskim oblikom. Imenica *pucanj* spadala bi pak po istim podacima u prvi pojas, odnosno njezini su oblici prave dublete, kod kojih su konkurentski oblici jednakovrijedni.

Upravo opisan pristup podjeli dvostrukih oblika izrazito je prikladan za precizniji leksikografski opis jer nam omogućuje da u skladu s korpusnim podacima imenicama dodijelimo različite empirijski utemeljene odrednice. U prikazu rezultata istraživanja u § 5 obrađene ćemo imenice odrediti s obzirom na pripadnost pojedinom pojasi dubletnosti te ćemo usporediti tako dobivene rezultate s opisom tih imenica u rječnicima suvremenoga jezika.

## 5. Rezultati

U ovom ćemo odjeljku opisati i interpretirati rezultate korpusne pretrage 26 imenica s dvostrukom (dugom i kratkom) množinom. Svaka od triju skupina imenica obrađuje se u zasebnom odjeljku, a zaključni se osvrt daje u § 5.4.

Za svaku imenicu unutar skupine prikazuje se udio jedne i druge množinske paradigme u pojedinom korpusu. Za svaku imenicu utvrđuje se također kojemu

<sup>22</sup> Određivanje točne granice (omjera) ovisi o različitim čimbenicima. Obično se promatra stilska i značenjska obilježenost pojedine inačice te njezina distribucija u kontekstu. Obilježenost i distribucija ograničena na specifične kontekste često koreliraju s njezinom potvrđenošću: veća obilježenost i ograničenija distribucija povlače za sobom i nižu frekvenciju u usporedbi s konkurentskom inačicom.

pojasu dubletnosti (§ 4.2) pripada s obzirom na korpus iz kojega su podatci. Podatci iz korpusa uspoređuju se s onima koje donose rječnici, odnosno posebno se ističu slučajji kad se podatci iz rječnika ne slažu s podacima iz korpusa. Daje se i kratki osvrt na relevantne sufikse i završetke.

Opisat ćemo ukratko i postupak prikupljanja podataka iz korpusa. Podatci iz dvaju korpusa prikupljeni su za sve imenice s pomoću regularnih izraza po modelu ‘puc(nji|anja|nje|njima)’ za kratku množinu i ‘pucnjev(i|a|e|ima)’ za dugu množinu. Na taj smo način dobili broj potvrda za oblike kratke množine, npr. *pucnji* (NV mn.), *pucanja* (G mn.), *pucnje* (A mn.) i *pucnjima* (DLI mn.), odnosno za oblike duge množine, npr. *pucnjevi* (NV mn.), *pucnjeva* (G mn.), *pucnjeve* (A mn.) i *pucnjevima* (DLI mn.).<sup>23</sup>

Tako prikupljeni konkordancijski redci pregledani su ručno kako bi se isključili različiti slučajevi homonimije. Gore prikazana imenica *pucanj* primjer je toga jer je njezin (pretpostavljeni) kratki oblik G mn. *pucanja* homoniman s G jd. glagolske imenice *pucanje*. Stoga smo sve retke s oblikom *pucanja* morali ručno pregledati, pri čemu je utvrđeno da svi pripadaju imenici *pucanje*, a nijedan imenici *pucanj*. Na sličan smo način pregledali i sve druge imenice, primjerice, neki oblici duge množine imenice *kesten* homonimni su s pridjevom *kestenov*, dok su oblici kratke množine imenice *plamen* homonimni s pridjevom *plamen*.

Posebno je važno istaknuti mogućnost homonimije između G jd. i G mn. imenica s KS naglaskom i zanaglasnom dužinom te složenica. Budući da se oblici genitiva jednine i množine tih imenica razlikuju samo u izostanku odnosno prisutnosti zanaglasne dužine koja se u pismu rijetko bilježi (npr. *galeba*) trebalo je ručno pregledati velik broj konkordancija koje su ciljale kratki oblik genitiva množine (*galeba*). Pregled je rezultirao i većim brojem dvoznačnih primjera koji se neće moći razriješiti kontekstom. Takvi su primjeri kao nerazrješivi isključeni iz frekvencija koje se navode u nastavku.

### 5.1. Imenice s nepostojanim a

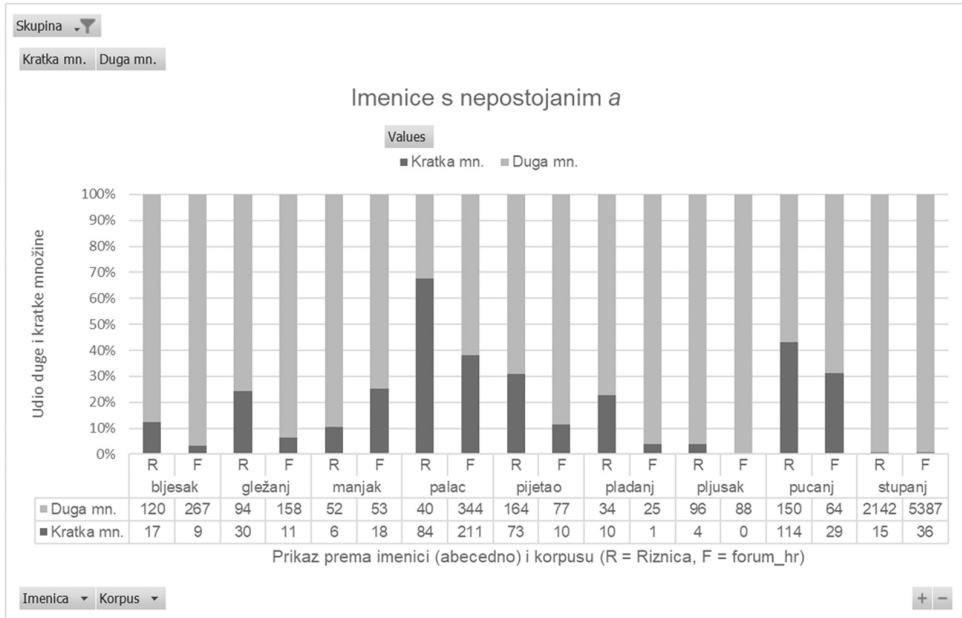
Analizirano je devet imenica iz ove skupine. To su, podsjetimo, imenice *bljesak*, *gležanj*, *manjak*, *palac*, *pijetao*, *pladanj*, *pljusak*, *pucanj* i *stupanj*.

Slika 1 zajednički donosi prikaz rezultata za svaku od imenica u obama korpusima. Imenice se navode abecedno. Uz svaku se imenicu uz udjele duge i kratke mno-

<sup>23</sup> Regularni su izrazi modificirani za pojedine imenice u kojih postoji varijacija u zapisivanju. U imenica s jatom (*bljesak*, *pijetao*) u korpusu smo pretražili sve relevantne oblike bez obzira na to kako je zapisan jat. Primjerice, regularni izraz ‘pi?jetlov(i|a|ima|e)’ nalazi sve oblike duge množine imenice *pijetao* bez obzira na pisanje jata: *pijetlovi* ili *pijetlovi* itd. Slično je učinjeno za imenice sa slovima s dijakritičkim znakovima jer se takve imenice u korpusima, a osobito u hrWaC-u, često pojavljuju u pojednostavljenu zapisu (*lezaji* umjesto *ležaji*).



žine navodi i njihova apsolutna frekvencija u pojedinom korpusu. Po istom ćemo modelu prikazivati i rezultate ostalih dviju skupina imenica u § 5.2, odnosno § 5.3.



Slika 1. Udjeli duge (sivo) i kratke (tamnosivo) množine za devet istraženih imenica s nepostojanim *a*

Prvo ćemo komentirati ovu skupinu imenica kao cjelinu. Usporedba rezultata u dvama korpusima pokazuje da od devet analiziranih imenica samo jedna imenica ima veći udio kratke nego duge množine (*palac*), i to samo u korpusu *Riznica*. Čini se da dubletne imenice s nepostojanim *a* općenito preferiraju dugu množinu. To je suprotno od ostalih (nedubletnih) imenica s nepostojanim *a*, koje u velikoj većini mogu tvoriti isključivo kratku množinu, npr. imenice *kupac*, *lanac*, *mačak*, *nokat*, *prošac*, *stalak*<sup>24</sup> itd. iako postoji i određen broj imenica koje tvore isključivo dugu množinu, npr. *češalj*, *dabar*, *dojam*, *lakat*, *mozak*, *toranj*, *vepar*, *sajam*, *zajam* itd. (usp. Maretić 1963: 145, § 138a; Samardžija 1987: 133–134).<sup>25</sup> Opažamo da imenice *stupanj* i *pljusak* u obama korpusima praktički pripadaju potonjoj skupini jer su po-

<sup>24</sup> Imenicu *stalak* u svojem popisu dubleta navodi Samardžija (1987) te smo njezin status pokušali utvrditi pretragom u korpusu. Ona je potvrdila podatke iz rječnika (pa i našu jezičnu intuiciju) da je riječ o imenici koja se ostvaruje isključivo u kratkoj množini. U *Riznici* nije nađen nijedan primjer duge množine uz 51 primjer kratke, dok je u cijelom *hrWaC-u* nađen tek 1 primjer duge množine (*stalkovi-ma*) nasuprot čak 1464 primjera kratke množine.

<sup>25</sup> Čini se da među dvosložnim imenicama s dugom množinom u cjelini ima malo primjera koji bi u potpunosti isključivali kratku množinu, npr. *ovan* > *ovnovi*/\**ovni*; *mozak* > *mozgovi*/\**mosci*. Druge imenice barem u nekim slučajevima dopuštaju kratku množinu. Takva je na primjer imenica *lakat*, u

tvrde njihove kratke množine izrazito rijetke. S tim se ne slažu rječnici suvremenoga jezika, koji za obje imenice navode i dugu i kratku množinu (v. niže).

Tablica 6 donosi abecedni popis devet analiziranih imenica u kojem se za svaku navodi udio kratke odnosno duge množine u obama korpusima te se navodi kojem pojasu dubletnosti pripada. Informacija o pojasu dubletnosti pomoći će nam da za pojedine imenice lakše uočimo razlike među dvama korpusima po udjelima kratke i duge množine.

Tablica 6. Udjeli kratke i duge množine za devet imenica s nepostojanim a te pripadajući pojasi dubletnosti

imenica	Riznica			forum.hr		
	% kr	% dg	pojas	% kr	% dg	pojas
bljesak	12,4%	87,6%	većinski dg	3,3%	96,7%	neobilježeni dg
gležanj	24,2%	75,8%	većinski dg	6,5%	93,5%	neobilježeni dg
manjak	10,3%	89,7%	većinski dg	25,4%	74,6%	većinski dg
<b>palac</b>	67,7%	32,3%	<b>prava dubleta</b>	38,0%	62,0%	<b>prava dubleta</b>
<b>pijetao</b>	30,8%	69,2%	<b>prava dubleta</b>	11,5%	88,5%	većinski dg
pladanj	22,7%	77,3%	većinski dg	3,8%	96,2%	neobilježeni dg
pljusak	4,0%	96,0%	neobilježeni dg	0,0%	100,0%	dominantni dg
<b>pucanj</b>	43,2%	56,8%	<b>prava dubleta</b>	31,2%	68,8%	<b>prava dubleta</b>
stupanj	0,7%	99,3%	dominantni dg	0,7%	99,3%	dominantni dg

Dvije imenice iz uzorka – *palac* i *pucanj* – pripadaju pojasu od 69 do 30 posto udjela jednoga oblika u obama korpusima, što ih čini pravim dubletama (te su stoga **masno otisnute**). Za imenicu *pijetao* to vrijedi samo u korpusu *Riznica*, i to vrlo rubno jer je imenica na granici dvaju pojasa (prava dubleta i većinski dg).

Status četiriju imenica jednak je u obama korpusima (bijela pozadina): *manjak* se nalazi u rasponu od 70 do 90 posto udjela dominantnoga oblika, imenice *palac* i *pucanj* u pojasu pravih dubleta, a *stupanj* ima udio kratke množine ispod 1 posto.

Mnogo je zanimljivije pogledati pet imenica koje se razlikuju po statusu: *bljesak*, *gležanj*, *pijetao*, *pladanj* i *pljusak* (sivo osjenčana pozadina). U svih je pet udio kratke (slabije preferirane) množine veći u Riznici, što bi moglo značiti da jezik predstavljen Riznicom ima veću sklonost kratkoj množini. Tomu je vjerojatno razlog prisutnost tekstova iz književnoumjetničkoga stila, što se može oprimirati imenicom *bljesak* gdje je od 17 potvrda njih 14 iz književnosti, a samo 3 dolaze iz novina *Vjesnik*. Od 14 primjera iz književnosti pola ih dolazi iz tekstova jednoga pisca (Josip

koje se kratka množina, izgleda, najbolje čuva u genitivu množine (*lakata*), a što opaža već Maretić, a preliminarnom korpusnom pretragom pokazuje i Lečić (2016: 87).

Kosor). Uz to ne treba isključiti ni vremensku dimenziju koja, po svemu sudeći, ide u smjeru podudaranja duge množine i vremenski bližega perioda. Treba međutim biti oprezan sa zaključcima o razlikama među korpusima s obzirom na to da je riječ o relativno slabo posvjedočenim imenicama: većina ih ima relativnu frekvenciju 1 ili manje u obama korpusima.

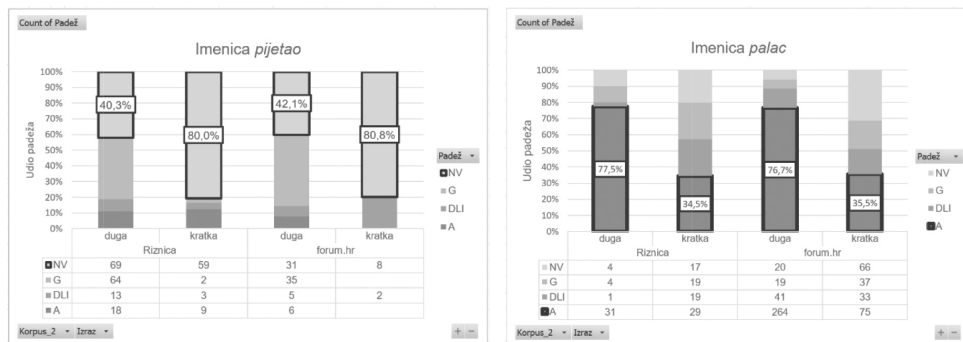
Usporedba tih podataka s informacijama iz četiriju rječnika suvremenoga hrvatskoga jezika pokazuje stanovitu neusklađenost. Za osam imenica (osim *manjak*) u svim se rječnicima navode kao moguće i duga i kratka množina.<sup>26</sup> Za imenicu *manjak* redovito se navodi kao mogućnost samo kratka množina iako je u korpusima dominantna duga množina. Redosljed navoda množine također se ne može pouzdano upotrijebiti u svim slučajevima. U RHJ-u se za svih osam imenica prvo navodi kratka množina, a slično je i u HJPu, gdje su dvije iznimke (*gležanj*, *pljusak*). Takav redosljed ne podudara se s podacima iz korpusa jer imenice ove skupine, kao što smo već rekli, preferiraju dugu množinu. U preostalim dvama rječnicima (ŠRHJ, VRH) redosljedi su mješoviti, ali ne čini se da održavaju uporabnu praksu predstavljenu korpusom. ŠRHJ primjerice za imenicu *gležanj* navodi prvo kratku množinu, a isto čini VRH za imenicu *stupanj* (obje imenice izrazito preferiraju dugu množinu u našim podacima). Ako bismo se doslovno držali gore navedenih podataka iz korpusa, za sve bi dubletne imenice iz ove skupine imalo smisla prvo navesti dugu množinu, a tek onda kratku. Jedina je iznimka imenica *palac*. U zaključcima rada (§ 6) raspraviti ćemo koliko takav pristup ima smisla.

U vezi s imenicama iz ove skupine spomenut ćemo još jedno zanimljivo opažanje. Uvidom u udjele pojedinih padeža unutar dviju množinskih paradigama može se uočiti da se u nekih imenica vrlo visok udio paradigme ostvaruje samo u jednom padežu. Spomenut ćemo na ovom mjestu dvije imenice u kojih je udio jednoga padeža veći od 70 posto paradigme. Kasnije ćemo nešto reći i o udjelima pojedinih padeža u imenica sa sufiksom *-(a)nj*.

Prva je imenica *pijetao*, u koje je u obama korpusima za najveći dio potvrda kratke množine odgovoran N mn. *pijetli*. Potvrda je malo u Riznici, ali tendencija koju tamo vidimo potvrđuje se u *forum.hr*-u jer je udio N mn. u kratkoj množini podjednak u obama korpusima (oko 80 %).

Imenica *palac* pokazuje sličan trend, samo što je za većinu potvrda njezine duge množine odgovoran akuzativ. Za njegovu su visoku frekvenciju odgovorni po svemu sudeći vrlo česti izrazi (frazemi) *držati palčeve* i *stisnuti palčeve*. U Riznici su ti izrazi odgovorni za 31 od 40 potvrda (77,5 %), a u *forum.hr* za 264 od 344 potvrde (76,7 %), kao što zorno pokazuje slika 2.

<sup>26</sup> U nekim se rječnicima imenica *bljesak* pojavljuje i kao *blijesak*, katkad i s drugim gramatičkim podacima.



Slika 2. Udio padeža u množinskim paradigmama imenica *pijetao* i *palac*

Upravo raspravljena opažanja pokazuju nam da neke imenice ne ostvaruju cjelovite množinske paradigme (u ovom slučaju kratke množine), već se među oblicima kratke/duge množine pretežito upotrebljava tek jedan padež. Stoga u nekim slučajevima možemo govoriti o **mješovitoj množinskoj paradigmi**, odnosno o paradigmi u kojoj se miješaju obrasci duge i kratke množine. Na isto nas upućuju i rječnici, posebice RHJ i HJP, u kojima povremeno nalazimo imenice u kojih se za N mn. navode oba množinska oblika, a za G mn. samo dugi. HJP tako navodi za *pucanj* N mn. *pūcnji/pūcnjevi*, a za G mn. samo *pūcnjēvā*. To se podudara i s našim podacima (v. niže u odjeljku o imenicama sa sufiksom *-(a)nj*). O konceptu mješovite množinske paradigme još ćemo govoriti u završnoj raspravi.

Na kraju ćemo se kratko osvrnuti na ponašanje imenica koje dijele isti sufiks. Tri imenice sa sufiksom *-(a)k* (*bljesak*, *manjak*, *pljusak*) relativno su slične po vrlo malim udjelima kratke množine, pri čemu se imenica *manjak* ističe njezinim relativno visokim udjelom u obama korpusima (R 10 %, F 25 %). Imenica na *-(a)k* ima velik broj<sup>27</sup> – npr. *prašak*, *ručak*, *zamak*, *jarak*, *čošak*, *točak*, *ručak*, *vosak*, *pupak* – te će biti potrebno opširnije istraživanje da bi se utvrdile eventualne tendencije u izboru njihove duge i kratke množine. Zasad možemo spomenuti imenicu *stalak*, za koju smo utvrdili da se isključivo pojavljuje u kratkoj množini<sup>28</sup> te imenicu *pljusak*, u koje nalazimo gotovo samo dugu množinu. Može se spekulirati o tome da razlika u preferencijama proizlazi iz toga što imenica *pljusak* dugom množinom izbjegava nastanak skupa *sc* u NV i DLI mn (*pljusci*, *pljuscima*), a isto bi onda vrijedilo i za kombinacije ostalih sibilanata i afrikata i *c* u imenica kao što su *prašak* i *ručak*. Tu će pretpostavku također tek trebati provjeriti.

Četiri imenice na *-(a)nj* (*gležanj*, *pucanj*, *pladanj*, *stupanj*) pokazuju određene međusobne razlike, koje se najbolje uočavaju kad se promotre udjeli pojedinih pa-

<sup>27</sup> Iscrpni popis može se naći u Matešić (1965–1967: 190–208, sv. II).

<sup>28</sup> V. bilj. 24.

deža tih imenica. Za imenicu *pucanj* već smo vidjeli da je riječ o imenici koja je prava dubleta u obama korpusima. Čini se da su za njezin status dublete odgovorne dvije suprotne tendencije. S jedne strane ta imenica prema našim podatcima uopće ne upotrebljava kratki G mn.,<sup>29</sup> za koji nema potvrde ni u jednom od dvaju korpusa, te se znatan udio oblika duge množine može pripisati G mn. *pucnjeva*. S druge strane velik broj potvrda kratke množine otpada na NV mn. *pucnji*, slično kao u imenice *pijetao* maloprije. Potonje vrijedi ponajprije za Riznicu, čime se može tumačiti veći udio kratke množine u Riznici (43,2 %) u odnosu na *forum.hr* (31,2 %).

Za imenicu *stupanj* već smo rekli da se tek iznimno pojavljuje u kratkoj množini, a taj ionako nizak broj potvrda gotovo isključivo dolazi od G mn., odnosno u obliku *stûpānjā*. Slično je i s imenicom *gležanj*. Imenice *gležanj* i *stupanj* ponašaju se suprotno od imenice *pucanj* te većina njihovih potvrda za kratku množinu dolazi upravo od oblika G mn. Imenica *pladanj* ne pokazuje nikakve tendencije. Izbjegavanje kratkoga G mn. imenice *pucanj* moglo bi se objasniti homonimijom s G jd. imenice *pucanje*, pogotovo zbog vrlo bliska značenja, ali isto objašnjenje ne vrijedi za imenicu *stupanj*, gdje bismo također mogli pronaći homonimiju s imenicom *stupanje*.

Općenito gledano čini se da ove imenice kao cjelina preferiraju dugu množinu (npr. imenica *toranj*) te bi prema tome imenica *pucanj* bila iznimka zbog svoje sklonosti kratkoj množini, osobito u NV mn. Imenica na *-anj* ograničen je broj<sup>30</sup> – npr. *bubanj*, *češanj*, *čučanj*, *oganj*, *ražanj*, *režanj*, *sužanj*, *svežanj*, *toranj*, *tutanj*, *žrvanj* – te će trebati dodatno istraživanje u tom smjeru.

Sa sufiksom *-(a)c* u uzorku je samo imenica *palac*. Ona uz iznimku A mn. preferira kratku množinu. Prema dostupnim podatcima čini se da takve imenice preferiraju upravo kratku množinu (usp. Maretić 1963: 145, § 138b), na što preliminarno upućuje pretraga imenica *jarac*, *kolac*, *lanac*, *ranac*, *šanac*, što će trebati potvrditi istraživanjem na većem broju imenica.<sup>31</sup>

Još spomenimo da, za razliku od iduće skupine, imenicama s nepostojanim *a* u ovoj skupini naglasno obilježje ne utječe na izbor množinske paradigme. Naime, za promatrane imenice važnijim se pokazuje glasovni sastav i tvorbeni sufiks. Kad tako ne bi bilo, očekivali bismo isto ponašanje imenica *bljësak* i *pălac* odnosno *stûpanj* i *pûcanj* (*pljusak* i *pladanj* dvojako se naglasno ostvaruju i u normativnim priručnicima i među govornicima: *pljûsak/pljûsak* i *plâdanj/plâdanj*).

<sup>29</sup> Usp. slično opažanje u Raguz (1997: 19).

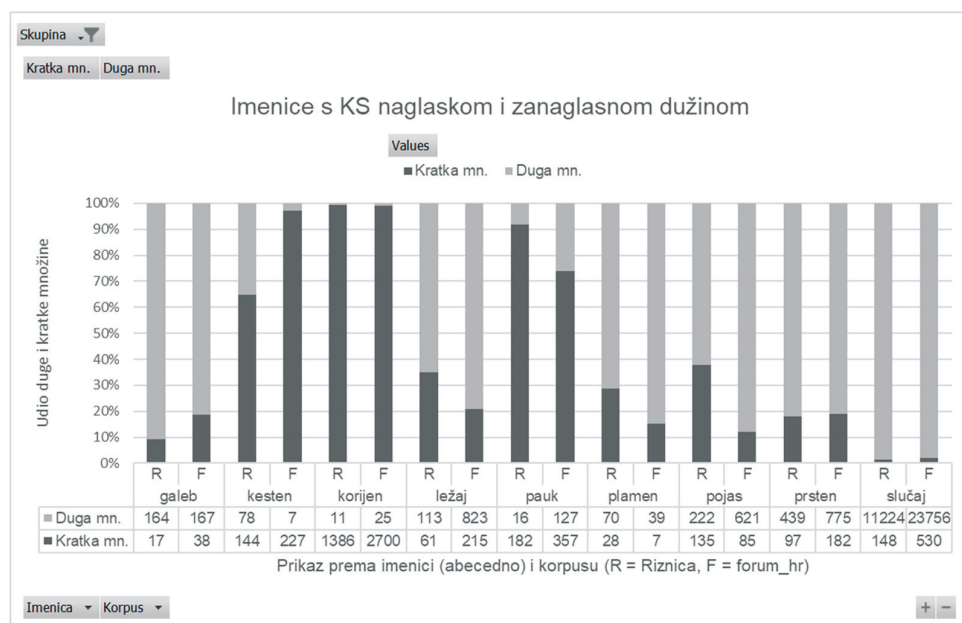
<sup>30</sup> V. Matešić (1965–1967: 281–282, sv. II). Usp. i Babić (2002: 358, § 1151).

<sup>31</sup> Imenice ove skupine relativno su brojne (usp. Matešić 1965–1967: 190–208, sv. I).

## 5.2. Imenice s KS naglaskom i zanaglasnom dužinom

Iz ove smo skupine izdvojili također devet imenica: *galeb*, *kesten*, *korijen*, *ležaj*, *pauk*, *plamen*, *pojas*, *prsten*, *slučaj*. Kao što je rečeno u uvodu, dvosložne imenice koje nemaju nepostojani *a*, a tvore dugu ili dvostruku množinu većinom pripadaju specifičnu prozodijskom tipu, odnosno imaju kratkosilazni naglasak na prvom slogu i zanaglasnu dužinu na drugom (*gālēb*), uz neke moguće iznimke.<sup>32</sup> Devet navedenih imenica uklapa se u taj prozodijski model.<sup>33</sup>

Slika 3 donosi prikaz rezultata za svaku od imenica u obama korpusima po modelu iz prethodnoga odjeljka (imenice se navode abecedno).



Slika 3. Udjeli duge (sivo) i kratke (tamnosivo) množine za devet istraženih imenica s KS naglaskom i zanaglasnom dužinom

Gledamo li skupinu kao cjelinu, imenice s kratkosilaznim naglaskom i dužinom raznolikije su od imenica s nepostojanim *a*. Dok je u tih imenica dominirala duga

<sup>32</sup> Samardžija (1987: 134) navodi sljedeće imenice koje se s različitih razloga ne uklapaju u navedeni prozodijski model: *dúbl*, *d̄io*, *kábel*, *kùlen*, *nàtpop*, *singl*, *titl*, *věo*, *vèrgl*, *žánr*. Za imenicu *kulen* v. fn. 12.

<sup>33</sup> Kao što uočava Maretić (1963: 145–146) imenice toga prozodijskoga tipa mogu se dodatno podijeliti s obzirom na to čuvaju li u genitivu jednine i drugdje u paradigmi zanaglasnu dužinu (npr. *pòjās/pòjāsa/pòjāsi*) ili ne (npr. *jāsēn/jāsena/jāseni*). Po svemu sudeći ta prozodijska razlika ne utječe na izbor duge i kratke množine. U našem uzorku tri imenice pripadaju tipu *jasen* (*kesten*, *plamen*, *prsten*), a ostale tipu *pojas* (u nekim se novoštokavskim govorima dužina pojavljuje u svim primjerima).

množina, od devet imenica s KS naglaskom i zanaglasnom dužinom u njima trima frekventnija je kratka množina (*kesten*, *korijen* i *pauk*), i to u obama korpusima. U ostatku imenica u većoj ili manjoj mjeri prevladava duga množina.

Na temelju ovolika uzorka nije moguće sa sigurnošću ustvrditi koliko imenice ovoga prozodijskoga tipa kao cjelina prihvaćaju dugu množinu. Tako znamo da postoje imenice koje ne dopuštaju dugu množinu i pojavljuju se isključivo u očekivanoj kratkoj množini, kao npr. *jelen*<sup>34</sup> (usp. i neke napomene niže u ovom odjeljku). Tim bismo imenicama mogli pribrojiti i imenicu *korijen* iz našega uzorka, kod koje su udjeli duge množine zanemarivi. S druge se strane mogu spomenuti i primjeri imenica koje se pojavljuju gotovo isključivo u dugoj množini, uz rijetke primjere kratke množine. Takva je imenica *slučaj* iz našega uzorka, ali i imenica *vitez*.<sup>35</sup>

Tablica 7 donosi abecedni popis devet analiziranih imenica s KS naglaskom i zanaglasnom dužinom. Za svaku se imenicu daje udio kratke odnosno duge množine u obama korpusima te se navodi pojas dubletnosti za svaki od korpusa.

Tablica 7. Udjeli kratke i duge množine za devet imenica s KS naglaskom i zanaglasnom dužinom te pripadajući pojasi dubletnosti

imenica	Riznica			forum.hr		
	% kr	% dg	pojas	% kr	% dg	pojas
galeb	9,4%	90,6%	neobilježeni dg	18,5%	81,5%	većinski dg
kesten	64,9%	35,1%	<b>prava dubleta</b>	97,0%	3,0%	neobilježeni kr
korijen	99,2%	0,8%	dominantni kr	99,1%	0,9%	dominantni kr
ležaj	35,1%	64,9%	<b>prava dubleta</b>	20,7%	79,3%	većinski dg
pauk	91,9%	8,1%	većinski kr	73,8%	26,2%	većinski kr
plamen	28,6%	71,4%	većinski dg	15,2%	84,8%	većinski dg
pojas	37,8%	62,2%	<b>prava dubleta</b>	12,0%	88,0%	većinski dg
prsten	18,1%	81,9%	većinski dg	19,0%	81,0%	većinski dg
slučaj	1,3%	98,7%	neobilježeni dg	2,2%	97,8%	neobilježeni dg

Općenito uzevši, udio većinske množinske paradigme u svih imenica može biti (približno) jednak u obama korpusima, što je slučaj u imenicama *korijen*, *pauk*, *prsten* i *slučaj*. Za četiri imenice udio dominantne paradigme veći je u forum.hr-u nego u Riznici (*kesten* 65 % prema 97 %, *ležaj* 65 % prema 79 %, *plamen* 71 % prema 85 %, *pojas* 62 % prema 88 %), samo je za imenicu *galeb* suprotno, njoj je udio dominantne

<sup>34</sup> Oblici duge množine imenice *jelen* nisu potvrđeni u Riznici, a u hrWaC-u je potvrđeno 20-ak primjera od 10-ak tisuća potvrda te imenice u tom korpusu. Imenica se navodi i s naglaskom *jèlen* (npr. u HJP-u).

<sup>35</sup> Usp. Maretićevu tvrdnju da sve imenice s KS naglaskom i dužinom koje imaju dugu množinu mogu imati i kratku (Maretić 1963: 145, § 139a). Slično vrijedi i za imenice s nepostojanim *a* (v. bilj. 25).

paradigme (duge množine) veći u Riznici nego u forum.hr-u. Vjerojatno je to stoga što se u forum.hr-u može naći više potvrda upotrebe ove imenice u značenju 'plaćeni ljubavnik, žigolo' a uz koju se veže kratka množina, i to ponajprije N mn. (tj. u tom se značenju pojavljuje oblik *galebi* prije nego *galebovi*). Time se otvara mogućnost utjecaja polisemije na distribuciju duge ili kratke množine.

Bijelo je osjenčano pet imenica čiji je status (u smislu pripadnosti pojasu dubletnosti) jednak u obama korpusima: *korijen*, *pauk*, *plamen*, *prsten* i *slučaj*. U prvih dviju dominira kratka množina (već smo rekli da imenica *korijen* gotovo da nema duge množine), a u imenica *plamen*, *prsten* i *slučaj* duga.

Status ostalih četiriju imenica razlikuje se u dvama korpusima. Imenice *kesten*, *ležaj* i *pojas* jedine su imenice iz ove skupine čije množinske paradigme imaju status pravih dubleta. Međutim, to vrijedi samo za korpus *Riznica*. Dubletni se status imenice *kesten* iz Riznice (iako s većim udjelom kratke množine) pretvara u izrazitu dominaciju kratke množine u forum.hr-u. Nešto je drukčija situacija s imenicom *ležaj*, gdje su udjeli duge množine nešto bliži (65 %, odnosno 79 %), ali se nalaze u dvama različitim pojasima. Konačno, duga množina dominira i u imenice *pojas* iako je mnogo bolje zastupljena u forum.hr-u.

U preostalih dviju imenica (*galeb* i *plamen*) dominira duga množina. Udjeli duge množine slični su za obje imenice u forum.hr-u, dok je imenici *plamen* u Riznici mnogo bolje potvrđena kratka množina nego imenici *galeb* te stoga imenica *plamen* u Riznici stoji na samoj granici statusa prave dublete.

Ako iznesene podatke usporedimo s navodima u rječnicima, dobivamo sliku sličnu onoj u prethodnoj skupini. Ondje smo vidjeli da se za gotovo sve imenice u rječnicima navode i duga i kratka množina kao moguće. Slično je i ovdje. Gotovo za sve imenice rječnici koji daju množinske oblike navode obje množine.<sup>36</sup> To vrijedi i za imenicu *slučaj*, koja je, kako smo vidjeli, vrlo slabo posvjedočena u kratkoj množini. Usporedbe radi spomenimo imenicu *jelen*, koja je, kako smo vidjeli, tek iznimno posvjedočena u dugoj množini.<sup>37</sup> Ta je činjenica odražena i u svim konzultiranim rječnicima, u kojima se za imenicu *jelen* navodi samo kratka množina. Pojava duge množine u korpusima može biti rezultat analogije prema skupini naglasno-morfološki sličnih imenica koje se odnose na životinje, a koje imaju (i) dugu množinu (*galeb*, *gavran*, *golub*, *labud*, *pauk*, *sokol*...). Stoga se može reći da rječnici uglavnom dobro bilježe mogućnost pojave alternativnih množinskih paradigama,

<sup>36</sup> Najdosljedniji je pritom VRH, koji za sve imenice navodi obje množine. ŠRHJ jedino za *pauk* navodi samo kratku množinu, a RHJ jedino za *plamen* ne navodi informaciju o množini (za sve ostale imenice daje obje množine). HJP za tri imenice (*kesten*, *plamen* i *pojas*) ne navodi informaciju o množini u glavnom tekstu članka, ali se u izborniku *Izvedeni oblici* mogu naći množinske paradigme, i to kratka za *kesten* te duga za *plamen* i *pojas*.

<sup>37</sup> Usp. bilj. 34.



ali ne navode precizno njihove odnose. Tomu dodajmo slučaj imenica *pojas* i *prsten*, za koje se u dvama najrecentnijim rječnicima (ŠRHJ, VRH) navodi značenjska razlika između dviju množina. Tako imenica *pojas* po rječničkoj uputi dugu množinu ostvaruje u značenju 'struk' i 'remen', dok se kratka množina upotrebljava u smislu dijela Zemljine površine ili nekoga geografskoga područja. Za imenicu *prsten* pak nalazimo razliku između prstena u značenju nakita (kratka množina) i prstena u značenju obruča ili svjetlosnoga kruga (duga množina). Navedene značenjske razlike nismo dodatno provjeravali u prikupljenim primjerima te ostaje otvorenim pitanje jesu li posrijedi čvrsta pravila ili tek tendencije.

Kad je posrijedi udio pojedinih padeža u paradigmi, u ovoj skupini imenica nema toliko izrazitih primjera dominacije jednoga padeža, stoga se na tome nećemo zadržavati.

Još ćemo se kratko osvrnuti na ulogu koju imaju slični sufiksi i završetci osnove u određivanju preferencija u izboru množinske paradigme. O trima netvorbenim imenicama (*galeb*, *pauk*, *pojas*) nema se što reći. Među ostalim imenicama možemo izdvojiti četiri imenice koje završavaju na *-en* (*kesten*, *korijen*, *plamen*, *prsten*) i dvije koje završavaju na *-aj* (*ležaj*, *slučaj*).

Imenice na *-en* možemo povezati tek uvjetno jer se *-en* ne može sinkronijski prepoznati kao sufiks osim u imenici *prsten* (Babić 2002: 359, § 1155) i eventualno *plamen* (Babić 2002: 329, § 1027).<sup>38</sup> Ipak, i u tim slučajevima izdvajanje sufiksa u sinkroniji ima tek uvjetno smisla jer je etimološki riječ o dvama različitim sufiksima<sup>39</sup> te su te tvorbe izrazito stare (nesumnjivo praslavenske). Stoga možemo zaključiti da je sličnost navedenih imenica u tvorbenom smislu slučajna, ali i bez toga ne možemo isključiti mogućnost analogije zasnovane na sličnosti završnoga dijela osnove. Podatci nam pokazuju da se dvije (dijakronijski) netvorbene imenice (*korijen*, *kesten*) ponašaju suprotno od preostalih dviju imenica (*plamen*, *prsten*): prve dvije izrazito preferiraju kratku množinu, a druge su dvije većinski posvjedočene u dugoj množini. Za bolje razumijevanje imenica na *-en* bit će stoga potrebno istražiti ponašanje svih takvih imenica, čiji je broj ograničen i koje, napomenimo još jednom, pripadaju istomu prozodijском obrascu. Osim navedenih četiriju imenica i već spomenute imenice *jelen* tu su još imenice *busen*, *greben*, *grmen* (zast.), *jasen*, *kamen*, *kremen*, *kulen*, *pramen*, *remen*, *spomen*, *stremen*, *stršljen*, *ugljen* i *znamen*.<sup>40</sup> Neke se od njih već u rječnicima navode s dvostrukom množinom (npr. *jasen*, *remen*, *stršljen*), dok se druge navode samo s kratkom (npr. *greben*, *spomen*, *znamen*).

<sup>38</sup> Prema tome i *kesten* i *korijen* možemo svrstati u skupinu netvorbenih imenica.

<sup>39</sup> U imenice *prsten* riječ je o sufiksu *-en* nejasna podrijetla, dok je u *plamen* imenički *-men*. Za obje imenice v. Skok (1973: s.v. prst, paliti) i Matasović (2014: 27, 127, fn. 31).

<sup>40</sup> Popis je sastavljen prema Matešić (1965–1967: 243–263, sv. II; usp. Matasović 2014: 24–25, 27–28, 125–126).

Za razliku od imenica na *-en*, u imenica *slučaj* i *ležaj* riječ je o istom te inače vrlo dobro posvjedočenu sufiksu *-aj*.<sup>41</sup> Obje imenice preferiraju dugu množinu iako je to mnogo izraženije u imenice *slučaj*. Sufiks *-aj* zanimljiv je zato što sve imenice s tim sufiksom pripadaju istomu prozodijskom obrascu, odnosno imaju KS naglasak i zanaglasnu dužinu (Babić 2002: 104–105, § 271–273). To znači da su dvosložne imenice s tim sufiksom mogući kandidati za buduće opširnije istraživanje.<sup>42</sup> Samo usputni pregled takvih imenica u rječnicima pokazuje da su neke od njih posvjedočene s dvostrukom množinom već u rječnicima, npr. *tečaj* i *lišaj* (u ŠRHJ-u i HJP-u). Ostatak se imenica u rječnicima bilježi samo s kratkom množinom (*drhtaj*, *gutljaj*<sup>43</sup>, *jecaj*, *titraj*, *treptaj*, *trzaj*, *vapaj* itd.). Jezična nam intuicija govori da te imenice nisu sklone dugoj množini, no to će trebati provjeriti i korpusnom pretragom. Sufiks *-aj* poučan nam je po tome što pokazuje da za imenice na *-aj* ni sufiks ni prozodijska obilježja nisu presudni za to hoće li imenica imati dvostruku množinu. Ostaje otvorenim pitanjem što onda jest, o čemu ćemo još nešto kratko reći u završnoj raspravi u § 6.

### 5.3. Složenice

U trećoj je skupini odabrano osam imenica: *drvored*, *minibus*, *nogostup*, *parobrod*, *prototip*, *stereotip*, *suhozid* i *velegrad*. Tim je imenicama zajamčena višesložnost i tvorbena složenost: bilo dvjema imeničkim sastavnicama povezanima spojnikom *-o-* bilo dvjema sastavnicama od kojih je prva svojevrсни prefiksoid. Imenice *stereotip* i *prototip* posuđene su u gotovu obliku kao složenice, a zbog svoje su se druge sastavnice našle u množinskom preobilju.

Navedenim je imenicama druga sastavnica samostalna imenica koja svoju množinu ostvaruje u dugom obliku, dakle sa sufiksom *-ov-*: *redovi*, *stupovi*, *brodovi*, *gradovi*, *busovi* (češće *busevi*), *tipovi*, što zacijelo utječe na nesigurnost u izboru množinske paradigme složenice. Višesložne imenice u hrvatskom jeziku općenito gledano ostvaruju samo regularnu (kratku) množinu, onu bez proširenja (*običaj* – *običaji* / *\*običajevi*), stoga oblici složenica s dugom množinom kao što su *parobrodovi*, *velegradovi* ili *stereotipovi* izlaze iz postojećega sustava fleksijskih tipova.

Zbog nedostatka ili maloga broja zasvjedočenih primjera u potkorpusu forum.hr za ovu smo skupinu imenica, za razliku od prethodnih dviju, podatke crpili iz cjelovitoga hrWaC-a. Njih, kao i za prethodno analizirane dvije skupine, uspoređujemo

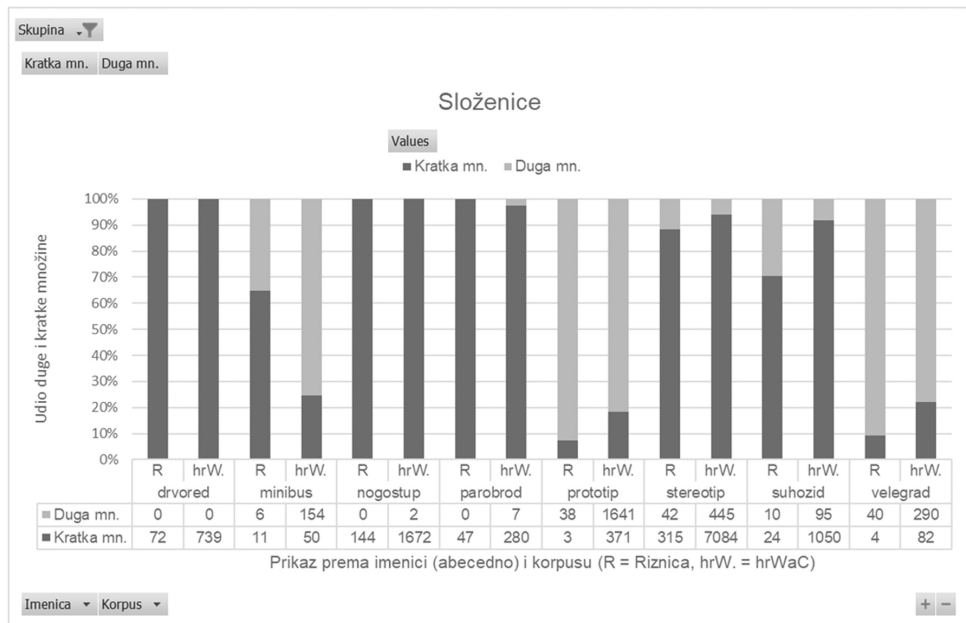
<sup>41</sup> Sufiks ima i jotirajući alomorf *-jaj*.

<sup>42</sup> Iscrpan popis može se naći u Matešić (1965–1967: 126–128, sv. II; usp. Babić 2002: 104–105, § 271–274).

<sup>43</sup> Posrijedi je sufiks *-ljaj*, no on se ne ponaša drukčije od sufiksa *-aj*.

s podacima iz korpusa *Riznica*. To nije metodološki idealno, ali omogućit će nam barem okvirnu usporedbu između dvaju korpusa.

Slika 4 donosi prikaz rezultata za svaku od imenica u obama korpusima po modelu iz prethodnih odjeljaka (imenice se navode abecedno).



Slika 4. Udjeli duge (sivo) i kratke (tamnosivo) množine za devet istraženih složenica

Gledajući cijelu skupinu, vidimo da od osam analiziranih imenica samo jedna (*drvored*) nema potvrde za dugu množinu ni u jednom korpusu. Budući da ta imenica ni u rječniku nije navedena kao imenica s dvostrukom množinom, možemo je isključiti iz budućih razmatranja. Vrlo blizu joj je i imenica *nogostup* koje također nema u Riznici u dugoj množini, a u drugom je korpusu zastupljena s tek dvama primjerima nasuprot gotovo 1700 primjera s kratkom, pa bismo, uzimajući u obzir pojase dubletnosti, i za nju mogli isključiti dvostrukost paradigme. Za ostalih šest imenica podatci su raznovrsni, kao što ćemo vidjeti u nastavku.

Pojedinačno je zastupljenost svih osam prikazana u nastavku abecednim redoslijedom (tablica 8). Analizu u nastavku treba uzeti sa zadržkom jer je broj potvrda u Riznici za neke od imenica vrlo nizak.

Tablica 8. Udjeli kratke i duge množine za osam istraženih složenica te pripadajući pojasi dubletnosti

imenica	Riznica			hrWaC (cijeli)		
	% kr	% dg	pojas	% kr	% dg	pojas
drvored	100,0%	0,0%	dominantni kr	100,0%	0,0%	dominantni kr
minibus	64,7%	35,3%	<b>prava dubleta</b>	24,5%	75,5%	većinski dg
nogostup	100,0%	0,0%	dominantni kr	99,9%	0,1%	dominantni kr
parobrod	100,0%	0,0%	dominantni kr	97,6%	2,4%	neobilježeni kr
prototip	7,3%	92,7%	neobilježeni dg	18,4%	81,6%	većinski dg
stereotip	88,2%	11,8%	većinski kr	94,1%	5,9%	neobilježeni kr
suhozid	70,6%	29,4%	većinski kr	91,7%	8,3%	neobilježeni kr
velegrad	9,1%	90,9%	neobilježeni dg	22,0%	78,0%	većinski dg

U prvim su se dvjema skupinama imenica korpusi podudarali u određenju prave dublete, dok je ovdje među njima razlika izrazitija. Dvije imenice koje smo već isključili iz rasprave (*drvored* i *nogostup*) ujedno su jedine koje u obama korpusima pripadaju istomu pojasi dubletnosti.

Imenica *minibus* u Riznici je prava dubleta, dok u drugom korpusu ne pripada tomu pojasi (statusu prave dublete u Riznici vrlo je blizu i imenica *suhozid*). Najbliže su statusu pravih dubleta u *hrWaC*-u množinske paradigme imenica *minibus* i *velegrad*. Imenicama *drvored* i *nogostup*, koje su gotovo isključivo jednoobrazne u obama korpusima, vrlo su bliske imenice *parobrod* s nešto zabilježenih primjera duge množine u *hrWaC*-u te *stereotip* s udjelom kratke množine od oko 90 posto.

Ukupno gledajući, za šest je imenica, osim imenica *velegrad* i *prototip*, udio kratke množine u Riznici mnogo veći od duge, što je vrijedilo i za prvu skupinu, a to povezujemo s vrstom građe i vremenom potvrde. Prikazani podatci potvrđuju rečeno za prvi korpus te se može reći da se načelno ta skupina drži primarne (kratke) množine. Imenice *velegrad* i *prototip* u tom se smislu razlikuju od ostatka te je po svojoj prilici u njih presudna uloga sastavnica *grad* odnosno *tip*, što se nazire i iz naglasnoga množinskoga lika *velegrādovi*, *prototipovi*. Situacija je slična i u *hrWaC*-u po tome što većina imenica preferira kratku množinu, ali se uz dvije već viđene iznimke (*velegrad* i *prototip*) javlja i treća (*minibus*). Imenica *minibus* zanimljiva je po tome što je njezino ponašanje prilično drukčije u dvama korpusima, ali vjerojatno je ponašanje zabilježeno u *hrWac*-u autentičnije s obzirom na mali broj potvrda u Riznici. Bliskost omjera s imenicama *velegrad* i *prototip* možda upućuje na zajednički prefiksoidni obrazac.

Suvremenim rječnički priručnici raznoliko navode podatke o množini izabranih imenica. Neke se imenice ne nalaze u nekom od vrela (npr. *minibus*, *prototip* i *ste-*

*reotip* u ŠRHJ-u), dok se gdjekad o njima ne donose podatci o množini (npr. *drvored*, *nogostup*, *parobrod*, *prototip* u RHJ-u).<sup>44</sup> Na osnovi dostupnih podataka može se reći da se u rječnicima imenice *drvored*, *nogostup*, *parobrod* i *suhozid* navode s kratkom množinom, dok se jedino imenici *velegrad* navode oba oblika (negdje se prvo navodi kratki, drugdje dugi). Imenica *prototip* u dvama vrelima u kojima joj se navodi množina pojavljuje se s dugim oblikom, dok se imenica *stereotip* ondje pojavljuje s kratkim. Zbirno uzevši, vidimo da su podatci iz rječnika, premda necjeloviti, za ovu skupinu imenica podudarniji s podacima iz korpusa nego što je to bilo za prve dvije skupine te možemo reći da stalne kratke množinske oblike imaju sve imenice osim imenice *velegrad*. Dupletni status imenice *minibus* u Riznici (blizu je tomu statusu i *suhozid*)<sup>45</sup> nije se potvrdio rječničkim opisom. Gledamo li po tvorbenom uzorku, moglo bi se očekivati da se vrlo slično ponašaju imenice sa spojnikom *drvo-* (*red*, *nogostup*, *suhozid* i *parobrod*), potom *velegrad* i *minibus* te *stereotip* i *prototip*, međutim, podatci to ne potvrđuju, što znači da na distribuciju tih oblika utječu drugi čimbenici. Uzmemo li u obzir vrstu prve sastavnice, mogli bismo na jednu stranu svrstati imenice *drvored*, *nogostup* i *parobrod* (imenica), na drugu *suhozid* (pridjev), na treću *velegrad*, *minibus* i *prototip* (prefiksoid), a zasebno se pokazuje imenica *stereotip*, što nije neobično s obzirom na njezin ustroj koji najviše odudara od kompozicionalnoga tipa složenice (*stereotip* nije nikakav tip). Rezultati dobiveni na osnovi korpusa upućuju na to da postoji stanovita paralela između tvorbenoga obilježja i izbora množine.

Vidjet ćemo što se može reći o njihovu naglasnom sastavu. Premda prva sastavnica nije osnova za naglasno obilježje složenice, većina ih u njoj ima kratki naglasak. Drugi dio složenice naizgled su samostalne imenice *red*, *stup*, *zid*, *brod*, *grad* i *tip* i sve su dugosilaznoga naglasaka, jedino se imenica *bus* (< *autobus*) u sustavu pojavljuje u dvjema inačicama: *būs* i *bûs*, s tim da je u kraćem liku mnogo češća. Jednak naglasni obrazac s kratkosilaznim naglaskom na prvom slogu i dužinom na posljednjem ostvaruju sve imenice (osim imenice *minibus*)<sup>46</sup>: *dřvorēd*, *nōgostŭp*, *sŭhozid*, *pàrobrōd*, *vělegrād*, *stěreotip*, *prōtotip*. Analiza naglasnih obilježja u množini nije se pokazala relevantnom.<sup>47</sup> Preostaje nam pogledati značenje odnosno tvorbenu motivaciju navedenih složenica: *drvored* 'pravilno zasađena stabla duž ulice, u parku i sl.'; *nogostup* 'malo izdignut dio ulice, sa strane, za pješake (od asfalta, kamenih ploča ili dr.)'; *suhozid* 'zid sagrađen od kamena bez vezivnoga gradiva'; *parobrod* 'brod

<sup>44</sup> Može se pretpostaviti da se tada radi o kratkoj množini.

<sup>45</sup> Potvrde su za *suhozidove* nađene u nekoliko hrvatskih autora (Laušić, Aralica, Pavličić, Božić), a za *minibuseve/minibusove* u *Vjesniku*.

<sup>46</sup> *Minibus* se u rječnicima navodi s kratkosilaznim naglaskom na prvom slogu, no među govornicima se ostvaruje i drukčije.

<sup>47</sup> Očekivano se u kratkoj množini čuva zanaglasna dužina (*vělegrādi*, *sŭhozīdi*). Rječnici mahom navode *vělegradovi*, dok se uporabno ostvaruje i lik *velegrādovi*.

s pogonom na paru'; velegrad 'veliki grad'; *minibus* 'autobus malih dimenzija, za prijevoz 12 putnika i više'; *stereotip* 'konvencionalno, formulirano, ob. vrlo pojednostavljeno mišljenje, koncepcija ili vjerovanje'; *prototip* 'osnovni model; praoblik, pralík, prauzor, original'. Uzimajući u obzir navedena značenja, imenice možemo podijeliti u tri skupine:

- a. drvored, nogostup
- b. minibus, parobrod, prototip, suhozid, velegrad
- c. stereotip.

Premda na prvu izgleda da je druga sastavnica imenica *drvored* i *nogostup* – *red* i *stup*, posrijedi su zapravo tvorenice od *redati*, *redanje* odnosno *stupati*, *stupanje*, dakle, nije težište na redu i stupu, nego na drveću koje se stavlja u red, koje se reda, te mjestu po kojem se stupa nogom. Kad to znamo, posve očekivano razumijemo gotovo potpun izostanak iz korpusa *drvoredova* i *nogostupova*. Nasuprot tomu u imenica *suhozid*, *velegrad*, *parobrod*, *minibus* i *prototip* druga je sastavnica imenica kojoj prva služi kao atribut da je поближе označi: kakav je zid, kakav je grad, kakav je brod odnosno bus, to je dakle vrsta zida, grada, broda, busa, čak i tipa (u smislu onoga koji je reprezentativni primjerak), dok za *drvored* ne možemo reći da je vrsta reda niti za *nogostup* da je vrsta stupa. Razumljivo je stoga da se u toj podskupini pojavljuju oblici duge množine jer govornika upućuju na glavu sintagme, temeljnu sastavnicu o kojoj je riječ: zidove, gradove, brodove, busove i tipove. Stanje bi se moglo protumačiti kao svojevrsan sukob dvaju načela: prvoga, gramatičkoga, po kojem višesložne imenice imaju množinu bez proširenja i drugoga, značenjskoga, po kojem glavnu riječ ima temeljna sastavnica. Ta se dva načela, kao što smo vidjeli, sukobljuju i u rječničkom opisu pa će tako npr. za imenice *suhozid* i *parobrod* množina ondje biti *suhozidi* i *parobrodi*, a za *velegrad* dvostruka *velegradi* i *velegradovi*.<sup>48</sup>

Zbog vrlo niska broja pojavnica analiza oblika po pojedinim padežima nije se mogla vjerodostojno provesti.

#### 5.4. Rasprava

Provjera korpusnih potvrda i usporedba s rječničkim opisom 26 odabranih imenica u množini pokazala je da postoji a) nepodudaranje u oblicima između korpusa, uvjetno rečeno, između jezičnih varijeteta (Riznice i hrWaC-a/forum.hr-a), b) nepodudaranje podataka iz korpusa s onima u rječnicima, c) nepodudaranje između četiriju rječnika i d) nepodudaranje opisa unutar istoga rječnika. To potvrđuje pretpostavku o ostvarenju dubletnosti među govornicima gotovo za sve izdvojene imenice, ali upućuje i na neujednačenost normativnoga opisa, odnosno kolebljivost u izboru jedne ili druge paradigme. Pojavnost oblika po padežima za prve dvije sku-

64 <sup>48</sup> Za imenicu *minibus* nema podataka o množini.

pine imenica otvorila je prostor pojavi *mješovite paradigme*. Naime, pokazalo se da su neke imenice češće ne samo u jednom padežu u jednom obliku, a u drugom padežu u drugom, nego i da se oblici duge množine pojavljuju u paradigmi kratke i obratno. Tako na primjer imenica *palac* dugu množinu ostvaruje najvećim dijelom zahvaljujući akuzativu množine (u izrazu *držati/stisnuti palčeve*), dok se oblik genitiva množine imenice *pucanj* ostvaruje isključivo u dugom obliku (*pucnjeva*), kako navode i rječnici, unatoč tomu što se u nominativu navode oba oblika, dugi i kratki. Dakle, oblici jedne ili druge paradigme nisu strogo odijeljeni, nego jedni ulaze u prostor drugih.<sup>49</sup>

Zbog specifičnih kombinacija dočetnih suglasnika i uloge nepostojanoga *a* (za imenice iz 1. skupine) te na koncu i zanaglasne dužine (ponajviše za imenice 2. skupine), koja nije svim govornicima razlikovno sredstvo, ondje već i u morfološkom opisu postoji neujednačenost. Primjere takva kolebanja nalazimo i u imenica koje ovdje nismo detaljnije analizirali.<sup>50</sup> Gledajući po pojasi dubletnosti, koje smo u § 4.2 definirali kao kriterije za određivanje razine dubletnosti (od prave dubletnosti do dominacije jednoga oblika), u prvoj skupini, među dvosložnim imenicama s nepostojanim *a*, proizlazi da su prave dublete u obama korpusa samo imenice *pucanj* i *palac*, dok su najudaljenije od dubletnoga statusa *pljusak* i *stupanj*. U ostalih prevladava većinski status jednoga oblika, i to duge množine. Utjecaj sufiksa na pojavu jednoga ili drugoga oblika ne može se pokazati na osnovi samo odabranih imenica, međutim, proširi li se istraživanje na veći broj imenica, mogu se očekivati stanovite tendencije.<sup>51</sup>

U drugoj skupini, koju čine dvosložne imenice s kratkosilaznim naglaskom i dužinom, pojasi dubletnosti pokazuju da bi prave dublete bile imenice *kesten*, *ležaj* i *pojas*, i to samo u korpusu Riznica, a blizu je tomu statusu i imenica *plamen*.

U skupini složenica prava je dubleta *minibus* i najbliže tomu *suhozid*, ali samo u jednom korpusu (Riznica), dok su u drugom omjeri udaljeniji. Većina složenica ima sklonost ka kratkoj množini, osim imenica *velegrad*, *minibus* i *prototip*.

<sup>49</sup> Slična zapažanja donosi Vukušić (1990) govoreći o množinskim oblicima imenica *vuk*, *vrag*, *rog* i *bog*.

<sup>50</sup> Na primjer, za imenicu *prašak* u jednom se vrelu navodi dvostruka množina *prašci/praškovi*, a u genitivu samo *prašaka*, tako i *kolci/kolčevi*, a u genitivu samo *kolaca*, dok za imenicu *ražanj* vrijedi obratno: u nominativu *ražnjevi /ražnji*, a u genitivu samo *ražnjeva*; za imenicu *ručak* nominativ množine glasi *ručkovi*, ali se u genitivu pojavljuje i kratki oblik, i to na prvom mjestu *ručaka/ručkova* itd.

<sup>51</sup> Po svemu sudeći imenice na *-anj* i *-alj* u velikoj su većini dugomnožinske (*češalj*, *čučanj*, *gležanj*, *pladanj*, *stupanj*, *pucanj*, *toranj*, *svežanj* itd.); imenica *sužanj* tu je iznimka jer pretežu kratki oblici, ali njezina uporaba nije posve vremenski i stilski neutralna te se uglavnom pojavljuje u kontekstu starijih razdoblja). Imenice na *-am* također su dugomnožinske (*pojam*, *dojam*). Prave dublete potencijalno su imenice na *-ak* (*ručak*, *prašak*, *zamak*, *točak*).

Premda su unutar sebe heterogene, tri skupine pokazuju ove tendencije: 1. imenice s nepostojanim *a*: većinom duga množina, 2. imenice s KS naglaskom: laganom većinom kratka, 3. složenice: laganom većinom kratka. Međutim, uzorak analiziranih imenica nedovoljan je da bi se ti rezultati pretočili u tvrdnje o morfološkom ponašanju pojedine skupine, stoga bi za takvo što trebalo uključiti mnogo veći postotak imenica (u idealnom slučaju sve) te načiniti statističku analizu po određenim parametrima (sufiks, naglasak, kombinacije dočetnih suglasnika itd.), kojima se mogu pridružiti i eksperimentalna istraživanja s varijablama kao što su stupanj prihvatljivosti oblika s obzirom na dob govornika, tip obrazovanja, organski idiom i sl. Za neke se imenice (*korijen, pojas, prsten, ležaj*) s obzirom na kolokacije nazire važnost značenjske diferencijacije te se u tom smislu može govoriti o tendencijama da se ona ostvari na morfološkom planu u izboru jedne ili druge množinske paradigme.

## 6. Zaključak

U istraživanje hrvatskih imenica s dvostrukim oblicima množine pošli smo od činjenice da je postojeći gramatički opis s više razloga nedostatan te da postoje različite jezične intuicije hrvatskih govornika koje se ogledaju u raznolikosti korpusnih primjera. Najprije smo na temelju pretrage korpusa htjeli vidjeti koje imenice preferiraju dugi, a koje kraći oblik te postoji li pritom kakva uvjetovanost. Dobiveni omjeri dvaju oblika u dvama različitim korpusima pokazala su da postoje razlike između korpusa u smislu da je korpus koji predstavlja književnu i standardnojezičnu građu nešto konzervativniji u potvrdi oblika, odnosno da ima veći udio kratke množine u odnosu na korpus neformalnijega, „polustandardnoga” jezika za prve dvije skupine. Podatci iz suvremenih rječnika također upućuju na čuvanje dvostrukosti u opisu i kad ona nije korpusno potvrđena te na mješovitost pojedinih paradigama.

Takvu pojavu tumačimo kao neminovan ishod jezične mijene koji se još nije stigao preklopiti sa suvremenim gramatičkim opisom. Duga je množina povijesno mlađa. Dijakronijski podatci o zastupljenosti dviju množina određenih imenica bez sumnje upućuju na smjer utjecaja s istoka štokavskoga prostora prema zapadu u vidu širenja duge, sekundarne, množine (usp. Orlandi 1963; Akademijin rječnik), a korpusni podatci pokazuju do koje je mjere taj proces zahvatio pojedinu imenicu odnosno skupinu općenito.

Sufiksi *-ov/-ev-* rabe se kao svojevrsni množinski markeri i govorniku osiguravaju maksimalnu nedvosmislenost i preciznost te je to mogući razlog njihova širenja. Proces zacijelo neće ići u suprotnom smjeru, osim ako se on normativno ne počne izrazito promicati, što dosad nije bio slučaj. U tom je smislu važno progovoriti i o odnosu standardnojezične norme prema korpusu. Govoreći o korpusima iz perspektive korisnika, Blagus Bartolec i Matas Ivanković ističu da su korpusni tekstovi te pogreške i odstupanja od neutralne, kodificirane norme kao njihov sastav-



ni dio dobar pokazatelj stvarne jezične upotrebe u odnosu na upotrebu propisanu normom (Blagus Bartolec i Matas Ivanković 2019: 40), čime se možda implicira da bi korpus služio kao svojevrsni korektiv koji bi trebao utjecati na to da se normativni opis ne udalji previše od uporabe.

Istraživanje na temelju korpusa trebalo nam je osigurati koliko-toliko objektivan sud o stvarnoj zastupljenosti pojedinih množinskih oblika, međutim, pitanje je što će se s dobivenim podacima učiniti i koliko će oni biti presudni za gramatički opis navedenih imenica.<sup>52</sup> Ako je za neku imenicu u korpusu utvrđena golema prevlast jednoga oblika nad drugim, treba li to utjecati i na normativnu preporuku? Suvremeni rječnici vole istaknuti utemeljenost na korpusu, ali je ta utemeljenost katkad više konzultacijska negoli stvarno provedena (usp. Hudeček i Mihajević 2018). Rječnički i gramatički opisi obično slijede nekakav općeprihvaćen stav o tome što jest dopušteno i kako bi koji oblik trebao glasiti, a koji se često preuzimao iz prethodnih izdanja i ranijih izvora bez provjere stvarne potvrđenosti među govornicima. Postoje mišljenja da tako i treba biti, odnosno da norma ne treba slijediti govorničku praksu, nego polaziti od utvrđenih činjenica jezičnoga standarda. Većinom se to dvoje podudara, no primjeri mnogih dubleta i inačica gotovo na svim jezičnim razinama pokazuju da postoje i stanovita nepodudaranja. Dostupnost, reprezentativnost i veličina današnjih korpusa ne mogu se pak zanemariti pri prosudbi o statusu nekoga oblika, što je prije bilo moguće, pa se time pomalo mijenja i stav o načinu izrade i ulozi normativnih priručnika. To naravno ne znači jednak status svakoga oblika koji se u korpusu pojavi pa normativni status na osnovi temeljite znanstvene prosudbe koja uključuje i jezičnu tradiciju i suvremene inovacije proizišle iz uzusa nema alternativu. Konačno, kako navodi Hunston (2002: 22–23), korpus nam ne daje informacije o tome je li što u jeziku moguće ili nemoguće, nego što jest ili nije frekventno na temelju svoga sadržaja, kao uzorka jezika, te velik broj primjera koji nam on osigurava tek treba interpretirati i opisati na osnovi jezične intuicije. Upravo u sprezi jednoga i drugoga vidimo najbolji put suvremenoga opisa jezičnih činjenica.

Na osnovi maloga uzorka imenica koje ostvaruju dva množinska tipa pokazuje se stanovito nepodudaranje leksikografskoga opisa s korpusnim podacima. Tomu je nepodudaranju više razloga: starost rječničke građe, oslanjanje isključivo na prethodne opise, vođenje vlastitom intuicijom, nekonzistentnost opisa za isti morfološki obrazac itd. S druge strane ni korpusi nisu idealni pokazatelji stvarne upotrebe neke jezične jedinice znamo li kako se oblikuju i na temelju kakve građe su nastali. Nepodudarnost rezultata između dvaju vrela ipak nije takva da bi dovela u pitanje vjerodostojnost postojećega leksikografskoga opisa, no upućuje na važnost praćenja suvremenoga jezičnoga stanja, bilo to i u manje formalnim varijetetima. Drugo je pi-

<sup>52</sup> Postoje ocjene da su gramatički podatci u korpusno utemeljenu rječniku mnogo pouzdaniji i detaljniji nego u drugim rječnicima (usp. Štrkalj Despot i Möhrs 2015: 343).

tanje kako bi (i bi li) suvremeni leksikografski priručnici trebali biti normirani i kako će se pristupiti pojavama koje normativno imaju dvojbenu status. Naime, slijedimo li trend suvremene leksikografije po kojem rječnik ne treba propisivati upotrebu, nego informirati korisnika na temelju jezičnih podataka o uporabi (usp. Filipović Petrović u tisku), pitanje o normativnom presuđivanju postaje neumjesnim.

Nama je pak važno pitanje kako kvantitativne odnose koje korpus prikazuje preračunati u prihvatljivost, odnosno kako najbolje opisati ono što nalazimo u korpusu. Ako su oba množinska oblika neke imenice normativno prihvatljiva, onda je korpus izvrstan alat koji će nam pomoći da tu prihvatljivost kvantificiramo, da je izrazimo u okviru nekih vrijednosti kroz presjek govorničkih praksi. Mislimo da je uvođenje pojasa dubletnosti dobar način da se ponajprije opiše stanje svake imenice (a onda i skupine), a potom se dobiveni podatci mogu izravno ili posredno povezati s njezinim opisom u rječniku ili gramatici. Izravno se to može učiniti tako da se na primjer u navođenju oblika najprije navede onaj koji je dominantan ili neobilježen (pa i većinski) po pojasi dubletnosti. Time se taj oblik nužno ne preporučuje, nego se ističe njegova prevlast u uporabi, što može biti korisna informacija pogotovo inojezičnim govornicima hrvatskoga, koji ne polaze od vlastite jezične intuicije.<sup>53</sup> Posredno se rezultati korpusnih istraživanja preslikani u pojase dubletnosti mogu povezati s (normativnim) leksikografskim opisom preko baze podataka koja osim popisa dubletnih leksema sadržava izdvojene korpusne potvrde obaju množinskih oblika. Upravo u tome vidimo najbolju primjenu tako provedenih istraživanja jer će baza filtrirati i organizirati podatke o bročanim odnosima zatečenima u korpusu i tako omogućiti oslonac leksikografskom opisu s jedne strane, a s druge svakomu zainteresiranijem korisniku pružiti dodatnu informaciju o statusu nekoga oblika (s obzirom na vrstu teksta, prostor, vrijeme, autora...)<sup>54</sup> U tom je smislu s obzirom na veličinu i sve veću raznovrsnost građe koju sadržava korpus danas nezaobilazno vrelo suvremenoga leksikografskoga opisa, no ne tako da se njegovi podatci uzimaju kao jedino relevantni i da se potpuno zanemari dosadašnja leksikografska praksa zasnovana na podacima iz jezične tradicije i relevantnih znanstvenih studija (među kojima i onih nastalih na introspekciji ili intuiciji jednoga govornika), nego da on služi kao svojevrsna točka provjere na temelju koje će se pojedini podatci moći preciznije opisati.

Obrada rezultata upozorila je na važnu pojavu o kojoj se u okviru hrvatske fleksije malo govorilo, a to je miješanje paradigama, odnosno upotreba oblika jedne množine za jedan padež, a obiju za ostale padeže. Mješovita paradigma upućuje nas na pojavu tzv. praznina (engl. *gaps*) unutar paradigme koje se onda ispunjavaju oblicima druge paradigme. U morfologiji je trenutno vrlo aktualna tema defektivnosti

<sup>53</sup> Gramatički opis (npr. u priručniku) pritom ima više prostora da se takvi podatci u nj uključe.

<sup>54</sup> Predviđeno je naime povezivanje Hrvatskoga mrežnoga rječnika s Bazom hrvatskih morfoloških dubleta. Točni detalji još se razrađuju.

(nepotpunosti) s jedne strane (npr. Sims 2019) i preobilja s druge u kojoj se ta pojava sa zanimanjem promatra, a raspravlja se i o samom statusu oblika koji nije potvrđen na nekom mjestu unutar paradigme.<sup>55</sup> Veći uzorak naših imenica moći će ponuditi neke odgovore i u tom pogledu.

Analiza na temelju izdvojenih sufikasa (uključujući širu građu) pokazala je tek neke tendencije (npr. sufixi *-am*, *-alj* i *-anj* mahom su dugomnožinski, dok su *-ac* i *-ak* većinski prave dublete), ali nedovoljno da bismo taj kriterij mogli smatrati presudnim za izbor jednoga ili drugoga tipa množine. Naglasna obilježja potvrdila su se kao relevantna za drugu skupinu imenica, za koju se to i očekuje, ali nisu imala utjecaja na ponašanje preostalih dviju skupina.

U odgovoru na pitanje što je presudno za izbor jednoga ili drugoga množinskoga oblika osim uočena trenda širenja duge množine po načelu veće markiranosti, pogotovo u slučajima gdje razlikovnost oblika ovisi o ostvarenju zanaglasne dužine,<sup>56</sup> zacijelo ulogu ima i postupanje po analogiji sa sličnim imenicama koje su se već ostvarile u dugoj množini. Skupina složenica svojim je sastavom heterogenija te status pojedinih imenica u njoj ovisi o tvorbenoj podlozi i jačini temeljne sastavnice. Ipak, ta je skupina najpogodnija za moguće normativno rješenje koje bi se gramatički moglo najbolje opravdati, a to je da se isključi ostvarivanje duge množine na višesložnim imenicama bez obzira na njihov sastav.

Supostojanje dvaju oblika, odnosno postojanje morfoloških dubleta u jeziku može trajati vrlo dugo, dapače, možda je to trajno obilježje jezika te se nikad ne razriješi u potpunosti, samo se broj kategorija i pojedinačnih leksema u kojima se dubletnost ostvaruje može povećavati ili smanjivati. Mali udio tzv. pravih dubleta među ovdje analiziranim imenicama<sup>57</sup> pokazuje da su većinom posrijedi omjeri u kojima je jedna paradigma većinska ili dominantna te je velika vjerojatnost da će ona u nekom času sasvim potisnuti konkurenciju.<sup>58</sup> Međutim, istodobno se mogu pojaviti nova *preobilna* mjesta te će se u budućnosti sigurno i o njima raspravljati.<sup>59</sup>

<sup>55</sup> Jedan od glavnih ciljeva projekta *Feast and Famine: Confronting Overabundance and Defectivity in Language* jest istražiti suodnos tih dviju pojava i način na koji funkcioniraju u morfološki bogatim jezicima. Raspravlja se npr. o statusu oblika koji je u sustavu moguć i gramatički ga opis podrazumijeva, ali se u realnoj jezičnoj praksi ne ostvaruje, tj. nema potvrde u korpusu.

<sup>56</sup> Slično vrijedi za skupinu imenica ženskoga roda s dubletnim genitivom množine koja je pod utjecajem širenja morfema *-i*, što je tema zasebnoga rada.

<sup>57</sup> Tj. imenica kod kojih su dvije paradigme približno podjednako zastupljene. Pritom treba istaknuti da se prema svemu što znamo iz dostupne literature može očekivati da je imenicâ s podjednakim udjelom množinskih paradigama mnogo manje nego imenicâ kod kojih jedna od množinskih paradigama znatno preteže, tj. uključuje više od 90 % potvrda množine (Bermel, Knittl i Russell 2018: 206).

<sup>58</sup> Djelomično se to može spriječiti normativnim preporukama i izravnim intervencijama u lektoriране tekstove ako bi to bio interes hrvatske norme.

<sup>59</sup> O nekima već jest (usp. Kapović 2018).

U ovom radu nije bilo mjesta da se navedeni primjeri provjere na govornicima s obzirom na njihovu regionalnu odnosno dijalektnu pozadinu, čak ni dobnu stratifikaciju. Naime, uočavaju se vrlo zanimljive naznake upotrebe jednih ili drugih oblika upravo na temelju tih obilježja.

S obzirom na brojne nepoznanice o tome kako istraživati dubletne paradigme imenica odlučili smo istraživanje oblikovati u obliku pilot-studije u kojoj smo detaljno opisali mali uzorak imenica s dvostrukom (kratkom i dugom) množinom. Takav pristup omogućio nam je stjecanje uvida u brojna obilježja tih imenica koja se prepoznaju njihovim ponašanjem u korpusu. Na temelju tih uvida moći ćemo bolje planirati buduće istraživanje koje bi trebalo obuhvatiti sve zabilježene imenice s dvostrukom (dugom i kratkom) množinom.

Svakako će biti zanimljivo korpusnom istraživanju dodati eksperimentalno te na većem uzorku ispitanika provjeriti ovdje uočene trendove.

## Literatura

- Akademijin rječnik = Daničić, Đuro i dr. (ur.). 1880–1976. *Rječnik hrvatskoga ili srpskoga jezika* (23 sveska). Zagreb: Jugoslavenska akademija znanosti i umjetnosti. Dostupno na: <https://dizbi.hazu.hr/a/?pr=i&id=196759>. Datum posjete stranici: 31. svibnja 2022.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge: MIT Press.
- Baayen, R. Harald; Endresen, Anna; Janda, Laura; Makarova, Anastasia; Nessel, Tore. 2013. Making choices in Russian: pros and cons of statistical methods for rival forms. *Russian Linguistics* 37(3). 253–291. <https://doi.org/10.1007/s11185-013-9118-6>.
- Babić, Stjepan; Brozović, Dalibor; Škarić, Ivo; Težak, Stjepko. 2007. *Glasovi i oblici hrvatskoga književnoga jezika*. Zagreb: Nakladni zavod Globus.
- Babić, Stjepan; Brozović, Dalibor; Moguš, Milan; Pavešić, Slavko; Škarić, Ivo; Težak, Stjepko. 1991. *Povijesni pregled, glasovi i oblici hrvatskoga književnog jezika: nacrti za gramatiku*. Zagreb: Hrvatska akademija znanosti i umjetnosti – Nakladni zavod Globus.
- Babić, Stjepan. 2002. *Tvorba riječi u hrvatskome književnome jeziku*. 3. poboljšano izd. Zagreb: Hrvatska akademija znanosti i umjetnosti – Nakladni zavod Globus.
- Babukić, Vjekoslav. 2014 [1854]. *Ilirska slovnica*. Pretisak. Zagreb: Institut za hrvatski jezik i jezikoslovlje.
- Barić, Eugenija; Lončarić, Mijo; Malić, Dragica; Pavešić, Slavko; Peti, Mirko; Zečević Vesna; Znika, Marija. 1997. *Hrvatska gramatika*. 2. izd. Zagreb: Školska knjiga.
- Bermel, Neil; Knittl, Luděk. 2012a. Morphosyntactic variation and syntactic constructions in Czech nominal declension: corpus frequency and native-speaker judgments. *Russian Linguistics* 36(1). 91–119. <https://doi.org/10.1007/s11185-011-9083-x>.
- Bermel, Neil; Knittl, Luděk. 2012b. Corpus frequency and acceptability judgments: A study of morphosyntactic variants in Czech. *Corpus Linguistics and Linguistic Theory*. 8(2). 241–275. <https://doi.org/10.1515/cllt-2012-0010>.
- Bermel, Neil; Knittl, Luděk; Russell, Jean. 2015. Morphological variation and sensitivity to frequency of forms among native speakers of Czech. *Russian Linguistics* 39(3). 283–308. <https://doi.org/10.1007/s11185-015-9149-2>.

- Bermel, Neil; Knittl, Luděk; Russell, Jean. 2018. Frequency data from corpora partially explain native-speaker ratings and choices in overabundant paradigm cells. *Corpus Linguistics and Linguistic Theory*. 14(2). 197–231. <https://doi.org/10.1515/cllt-2016-0032>.
- Blagus Bartolec, Goranka; Matas Ivanković, Ivana. 2019. Korpus umom korisnika (na što treba pripaziti u korpusno utemeljenom istraživanju). U Matešić, Mihaela; Vlastelić, Anastazija (ur.), *Jezik i um: Zbornik radova s međunarodnoga znanstvenog skupa Hrvatskoga društva za primijenjenu lingvistiku održanoga od 3. do 5. svibnja 2018.*, 31–42. Zagreb: Srednja Europa.
- Bošnjak Botica, Tomislava; Hržica, Gordana. 2016. Overabundance in Croatian Dual-Class Verbs. *Fluminensia* 28(1). 83–106.
- Bošnjak Botica, Tomislava; Budja, Jurica; Hržica, Gordana. 2018. Rodovna i sklonidbena dvojnost imenica u hrvatskom jeziku. *Rasprave*. 44(1). 57–88. <https://doi.org/10.31724/rihjj.44.1.2>.
- Bošnjak Botica, Tomislava; Lečić, Dario; Hržica, Gordana; Budja, Jurica. 2019. Baza hrvatskih morfoloških dubleta (DvojBa). *Suvremena lingvistika* 45(87). 95–108. <https://doi.org/10.22210/suvlin.2019.087.03>.
- Brabec, Ivan; Hraste, Mate; Živković, Sreten. 1954. *Gramatika hrvatskoga ili srpskog jezika*. 2. prerađeno i nadopunjeno izd. Zagreb: Školska knjiga.
- Brozović, Dalibor. 2016. *Jezik današnji: 1965-1968*. Priredio Krešimir Mićanović. Zagreb: Disput.
- Carstairs-McCarthy, Andrew. 1987. *Allomorphy in inflexion*. London: Croom Helm.
- Clark, Eve V. 1987. The Principle of Contrast: A constraint on language acquisition. U MacWhinney, Brian (ur.), *Mechanisms of Language Acquisition*, 157–193. Hillsdale: Erlbaum.
- Corbett, Greville G. 2009. Canonical inflectional classes. U Montermini, Fabio; Boyé, Gilles; Tseng, Jesse (ur.), *Selected proceedings of the 6th Décembrettes: Morphology in Bordeaux*, 1–11. Somerville: Cascadia Proceedings Project. <http://www.lingref.com/cpp/decemb/6/paper2231.pdf>.
- Ćavar, Damir; Brozović Rončević, Dunja. 2012. Riznica: The Croatian Language Corpus. *Prace Filologiczne* (63). 51–66.
- Filipović Petrović, Ivana. u tisku. Što je hrvatskoj leksikografiji korpus. U *Sedmi hrvatski slavistički kongres: zbornik radova sa znanstvenoga skupa s međunarodnim sudjelovanjem održanoga u Šibeniku od 25. do 28. rujna 2019.*
- Frančić, Anđela; Hudeček, Lana; Mihaljević, Milica. 2005. *Normativnost i višefunkcionalnost u hrvatskome standardnom jeziku*. Zagreb: Hrvatska sveučilišna naklada.
- Gudkov, Vladimir. 1964. Duga i kratka množina kao diferencijalni semantički znak. *Jezik* 12(1). 6–11.
- Hamm, Josip. 1967. *Kratka gramatika hrvatskosrpskog književnog jezika: za strance*. Zagreb: Školska knjiga.
- HER = Matasović, Ranko; Jojić, Ljiljana (ur.). 2002. *Hrvatski enciklopedijski rječnik*. Zagreb: Novi Liber. <https://hjp.znanje.hr/>.
- HJP = v. HER
- Hudeček, Lana; Mihaljević, Milica. 2018. *Hrvatski mrežni rječnik – Mrežnik. Upute za obrađivače. Radna inačica*. Zagreb: Institut za hrvatski jezik i jezikoslovlje. <http://ihjj.hr/mreznik/uploads/upute.pdf>.
- Hunston, Susan. 2002. *Corpora in Applied Linguistics*. Cambridge: Cambridge University Press.
- Kapović, Mate. 2018. Širenje nastavka -i u genitivu množine e–deklinacije u suvremenoj štokavštini. *Suvremena lingvistika* 44. 39–72.
- Klaić, Bratoljub. 1953. O dugoj množini trosložnih i višesložnih imenica. *Jezik* 2(3). 74–78.

- Kroch, Anthony. 1994. Morphosyntactic Variation. U Beals, Katharine; Denton, Jeannette; Knippen, Robert; Melnar, Lynette; Suzuki, Hisami; Zeinfeld, Erica (ur.), *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, 180–201. Chicago: Chicago Linguistic Society.
- Lečić, Dario. 2015. Morphological doublets in Croatian: the case of the instrumental singular. *Russian Linguistics* 39(3). 375–393. <https://doi.org/10.1007/s11185-015-9152-7>.
- Lečić, Dario. 2016. *Morphological Doublets in Croatian: A multi-methodological analysis*. Neobjavljena doktorska disertacija, Sheffield: University of Sheffield. Dostupno na: <https://etheses.whiterose.ac.uk/16068/>. Datum posjete stranici: 31. svibnja 2022.
- Ljubešić, Nikola i Filip Klubička. 2014. bs,hr,srWaC - Web Corpora of Bosnian, Croatian and Serbian. U *Proceedings of the 9th Web as Corpus Workshop (WaC-9)*, 29–35. Gothenburg, Sweden: Association for Computational Linguistics. <https://doi.org/10.3115/v1/W14-0405>. Dostupno na: <https://www.aclweb.org/anthology/W14-0405>. Datum posjete stranici: 31. svibnja 2022.
- Maretić, Tomo. 1899. *Gramatika hrvatskoga jezika: za niže razrede srednjih škola*. Zagreb: L. Hartman.
- Maretić, Tomo. 1963. *Gramatika hrvatskoga ili srpskoga književnog jezika*. 3., nepromijenjeno izd. Zagreb: Matica hrvatska.
- Marković, Ivan. 2012. *Uvod u jezičnu morfologiju*. Zagreb: Disput.
- Matasović, Ranko. 2014. *Slavic nominal word-formation: Proto-Indo-European origins and historical development*. Heidelberg: Universitätsverlag Winter.
- Matešić, Josip. 1965–1967. *Rückläufiges Wörterbuch des Serbokroatischen* (2 sveska). Wiesbaden: Otto Harrassowitz.
- Mićanović, Krešimir. 2006. *Hrvatski s naglaskom: standard i jezični varijeteti*. Zagreb: Disput.
- Orlandi, Roberto. 1963. II plurale breve e lungo in serbo-croato. *Ricerche slavistiche* 11. 3–33.
- Pasini, Dinka. 2018. Supstandardni varijeteti hrvatskoga jezika u 20. stoljeću. U Bičanić, Ante; Samardžija, Marko; Pranjković, Ivo (ur.), *Povijest hrvatskoga jezika. Knj. 5, 20. stoljeće: prvi dio*, 580–629. Zagreb: Društvo za promicanje hrvatske kulture i znanosti CROATICA.
- Polančec, Jurica. 2017. Naknadni prijelaz iza glasova *s t z* u hrvatskom jeziku. *Rasprave* 43(1). 197–225.
- Raguž, Dragutin. 1997. *Praktična hrvatska gramatika*. Zagreb: Medicinska naklada.
- Rainer, Franz; Gardani, Francesco; Dressler, Wolfgang; Luschützky, Hans Christian (ur.). 2019. *Competition in Inflection and Word-Formation*. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-030-02550-2>.
- RHJ = Šonje, Jure (ur.). 2000. *Rječnik hrvatskoga jezika*. Zagreb: Leksikografski zavod Miroslav Krleža – Školska knjiga.
- Samardžija, Marko. 1987. Duga i kratka množina u hrvatskom književnom jeziku. *Jezik* 35(5). 129–136.
- Silić, Josip; Pranjković, Ivo. 2005. *Gramatika hrvatskoga jezika: za gimnazije i visoka učilišta*. Zagreb: Školska knjiga.
- Sims, Andrea 2019. *Inflectional defectiveness*. Cambridge: Cambridge University Press.
- Skok, Petar. 1973. *Etimologijski rječnik hrvatskoga ili srpskoga jezika* (3 sveska). Zagreb: Jugoslavenska akademija znanosti i umjetnosti.
- Stefanowitsch, Anatol. 2020. *Corpus linguistics: A guide to the methodology*. Berlin: Language Science Press. Dostupno na: <https://langsci-press.org/catalog/book/148>. Datum posjete stranici: 31. svibnja 2022.

- Šipka, Milan. 1999. Korelacije značenja i oblika duge i kratke množine imenica muškog roda u srpskom i/ili hrvatskom jeziku. U: Tošović, Branko (ur.), *Die grammatischen Korrelationen: Grazer Linguistische Slawistentage (GraLiS-1999), 4.-6. November 1999*, 165–176. Graz: Karl-Franzens-Universität Graz, Grazer Linguistische Slawistentage.
- Škarić, Ivo. 2007. Fonetika hrvatskoga književnoga jezika. U: *Glasovi i oblici hrvatskoga književnoga jezika*, 17–157. Zagreb: Nakladni zavod Globus.
- ŠRHJ = Jozić, Željko; Hudeček, Lana; Mihaljević, Milica (ur.). 2012. *Školski rječnik hrvatskoga jezika*. Zagreb: Školska knjiga – Institut za hrvatski jezik i jezikoslovlje. Dostupno na: <http://rjecnik.hr/>. Datum posjete stranici: 31. svibnja 2022.
- Štrkalj Despot, Kristina; Möhrs, Christine. 2015. Pogled u e-leksikografiju. *Rasprave*. 41(2). 329–353.
- Thornton, Anna M. 2011. Overabundance (Multiple Forms Realizing the Same Cell): A Non-Canonical Phenomenon in Italian Verb Morphology. U Maiden, Martin (ur.), *Morphological autonomy: perspectives from Romance inflectional morphology*, 358–381. Oxford: Oxford University Press.
- Veber Tkalčević, Adolfo. 1876. *Slovnica hrvatska: za srednja učilišta*. U Zagrebu: vlastita naklada.
- VRH = Jojić, Ljiljana (ur.). 2015. *Veliki rječnik hrvatskoga standardnog jezika*. Zagreb: Školska knjiga.
- Vukušić, Stjepan. 1990. Sklonidba imenica vuk, vrag, rog, bog u kratkoj množini. *Jezik* 38(3). 70–72.

## CROATIAN MASCULINE NOUNS WITH COMPETING PLURAL PARADIGMS: A CORPUS STUDY

Competing plural paradigms found in a segment of the disyllabic and polysyllabic masculine nouns are one of the most interesting features of the Croatian noun declension system. The paper analyzes the competition between the so-called long and short plural forms of 26 masculine nouns divided into three groups: 1) disyllabic and trisyllabic nouns with a “fleeting *a*”; 2) disyllabic nouns with the short falling accent and a post-accentual length; 3) compounds where the second component is a monosyllabic noun. The presented data paint a picture of a complex and heterogeneous phenomenon and reveal the inconsistencies with respect to how the analyzed nouns are presented in the contemporary Croatian dictionaries. Suggestions are made as to how to make the presentation of such nouns in dictionaries more consistent with the actual usage as reflected in the present corpus findings. The data also revealed the existence of nouns that display what could be called “mixed plural paradigms”, where individual cases within the plural paradigms of such nouns diverge as to their preferences for the forms of the two competing plural paradigms (e.g., a noun may prefer the short plural for one case, and the long plural for the rest of the paradigm). Special attention is paid to the methodology in corpus research of competition in morphology, particularly to the criteria for distinguishing different gradations of competition by making use of frequency bands.

**Keywords:** competing plural paradigms, corpus-based approach, frequency bands, mixed plural paradigms, Croatian language

Adrese autora:

**Tomislava Bošnjak Botica**

Institut za hrvatski jezik i jezikoslovlje Zagreb  
HR – 10 000 Zagreb, Republike Austrije 16  
tbosnjak@ihjj.hr

**Jurica Polančec**

Filozofski fakultet Zagreb  
HR – 10 000 Zagreb, Ivana Lučića 3  
jpolance@m.ffzg.hr

**Robert Sviben**

Filozofski fakultet Zagreb  
HR – 10 000 Zagreb, Ivana Lučića 3  
robertsviben3@gmail.com