

IZPELJAVA SAMOSTALNIKOV IZ GLAGOLOV: KORPUSNA PRODUKTIVNOST IN SLOVNIČNI OPIS

Nataša LOGAR

Fakulteta za družbene vede, Univerza v Ljubljani
Ljubljana, Slovenija

Logar, N. (2018): Izpeljava samostalnikov iz glagolov: korpusna produktivnost in slovnični opis. Slovenščina 2.0, 6 (2): 96–126.

DOI: <http://dx.doi.org/10.4312/sl02.0.2018.2.96-126>.

V prispevku predstavljamo razumevanje besedotvorne produktivnosti, kakršnega je na korpusni metodologiji konec 80. in v začetku 90. let 20. stoletja utemeljil Harald Baayen. Baayenov način izračunavanja produktivnosti *P* smo preizkusili na desetih slovenskih priponskih obrazilih, ki tvorijo samostalniške navadne izpeljanke iz glagolov, in sicer izpeljanke moškega spola s pomenom človeškega vršilca dejanja. Šlo je za obrazila *-vec*, *-telj*, *-ež*, *-uh*, *-avh*, *-ar*, *-lec*, *-ač*, *-aj* in *-ant*. Polavtomatski postopek pridobivanja podatkov smo izvedli na korpusu pisne slovenščine Kres ter med drugim ugotovili, da sta število tvorjenk, katerih del je določeno obrazilo, in produktivnost *P* tega istega obrazila lahko (zelo) različna. Npr.: največ tvorjenk izmed naštetih je tvorilo obrazilo *-lec*, medtem ko je bil *P* najvišji pri obrazilu *-ež*. Na podlagi podatkov iz Kresa smo nato prikazali še seznam obravnavanih obrazil skupaj s tvorjenkami, tj. seznam, kakršen bi lahko bil del nove slovnice slovenščine, ter ga na kratko komentirali, v zaključni razpravi pa smo se nato opredelili še do nekaterih prednosti in slabosti prikazanega pristopa.

Ključne besede: besedotvorje, produktivnost, H. Baayen, priponsko obrazilo, tvorjenka, slovnica, korpus

1 UVOD

Ustvarjanje besed z različnimi priponami je staro, kolikor daleč nazaj sledimo razvoj jezika. Prvotne pripone so večinoma omrtvele in se ohranile samo kot okamnine, besed ne določajo več ne po pomenu ne po obliki. Nastajale pa so nove pripone,

nekatero z zelo jasnim pomenom, druge z manj določenim, zato so nekatere v jeziku zelo tvorne in razširjene, druge manj (Bajec in dr. 1956: 70–71).

Navedena misel iz prve izdaje Slovnice štirih dobro povzema problematiko, ki ji raziskovalci besedotvorja kateregakoli jezika posvečajo precej pozornosti, namreč tematiko besedotvorne – pa sicer tudi širše, morfološke – produktivnosti (Baayen, Lieber 1991: 801; Bauer 2005: 315). Avtorji razprave o besedotvorni produktivnosti pogosto začenjajo z opozorilom o nenatančni rabi tega izraza in njegovi neenotni definiciji (npr. Aronoff 1976: 35; Lieber 1992: 1–9; Bauer 2005: 315, 330; Scherer 2015: 1781, 1782), večinsko pa se strinjajo, da gre za pojav z lastnostmi kontinuuma (Aronoff, Anshen 1998: 242–243). Na enem skrajnem koncu tega kontinuuma so obrazila (če za potrebe prispevka ostanemo le pri teh), ki jih pri tvorbi novih besed ne uporabljamo več – so "omrtvela", kot so se izrazili Bajec, Kolarič in Rupel; na skrajnem drugem koncu so obrazila, ki v tvorbeni postopek vstopajo zelo pogosto in so del zelo številnih tvorjenk – so torej "zelo tvorna in razširjena", medtem ko so vmes obrazila, ki niso niti zelo pogosta niti zelo redka (kot bi se po Likertu izrazili družboslovni metodologi). Pri tem je ključen izziv, na katerega so skušali odgovoriti nekateri, predvsem pa Baayen (1992; več gl. v nadaljevanju), kako to "kontinuumsko" lastnost izmeriti.

Tudi v slovenskem jezikoslovju je običajno, da raziskovalci besedotvornih pojavov svoje ugotovitve o besedotvornih morfemih, besedotvornih vrstah ali pa tvorjenkah iz njih interpretirajo z vidika njihove večje oz. manjše produktivnosti. Prim. nekaj takih ocen (poud. N. L.):

Nasploh lahko rečemo, da najdemo glagolske zloženke tipa *listopad* že v starocerkvenoslovanščini, v večji ali manjši meri pa so uporabljene v vseh slovanskih jezikih. Vsaj za slovenščino velja, da **tvorbena v glavnem niso več žive**, zlasti pa se ne spreminja obseg skladenjskopodstavnih glagolov (Vidovič Muha 1988: 53).

Tudi posnemovalni medmet *tik tak* je **produktiven** za glagol *tiktakati* in iz njega nastale tvorjenke *tiktakanje*, *tiktakajoč* in računalniški žargonizem *tiktakalo*

'processor' (Stramljič Breznik 2013: 104).

Sklenemo lahko, da je sklapljanje v slovenščini **manj produktiven** /.../, vendar pa enostaven in učinkovit način tvorjenja stilno zaznamovanih besed (Logar 2005: 191).

V zvezi s pomenotvornimi postopki velja omeniti tudi ugotovitev, da so tvorjenke, pri katerih se je zgodil pomenski prenos, **tvorbena manj produktivne** od tistih, pri katerih pomenskega prenosa ni (Kern 2017: 228).

Ko smo deloma že videli, se pri takih in podobnih interpretacijah poleg izraza produktivnost (oz. njegove sopomenke *tvornost*) uporabljajo še sorodne besede, kot so živost, razširjenost, aktualnost in pogostost. To je še dodatno opazno v naslednjem odstavku z naslovom *Živost besedotvornih vrst in sredstev* v Toporišič (2000: 160–161; poud. N. L.):

Besedotvorne vrste in sredstva **niso vsi enako živi in pogosti**. Od besedotvornih vrst je **najbolj pogosta** izpeljava, pri glagolih tudi sestava, nato pridejo zloženke in na koncu sklopi. Vse vrste pa **so žive**. To se ne da reči za vsa besedotvorna sredstva, npr. za vsa priponska obrazila. **Popolnoma mrtvo** je npr. priponsko obrazilo *-sm*, kakor ga imamo v besedi *pesem* 'to, kar se poje'; tudi *-ezen/-azen* je **zelo redko**. Prav tako **netvorno** je npr. priponsko obrazilo *-t*, kot ga imamo v samostalnikih tipa *povest* 'to, kar se pove'. **Zelo tvorno** pa je npr. priponsko obrazilo *-lec* (*bralec*, *čistilec*); z njim **je izpeljanih veliko besed**, in še vedno jih delamo. **Tvornost ponskih obrazil** lahko **zamre** samo za kak pomenski tip, za druge pomene pa je **še zmeraj zelo živa**; tako je s priponskima obraziloma *-(i)ca* in *-ka*, ko zaznamujeta ženski par moškemu, npr. *Smrekarica* /.../, namesto česar se danes, posebno v meščanskih krogih, oficialno rabijo posamostaljene pridevniške besede s pripono *-ova/-eva*, tj. *Smrekarjeva*.

V prispevku se bomo naprej odzvali na obe zgornji zadregi in skušali odgovoriti na vprašanja,

a) kaj je besedotvorna produktivnost (in kako se – vsaj v ožjem pomenu – loči od živosti, razširjenosti, aktualnosti in pogostosti oz. kaj ima z njimi skupnega) ter

b) kako besedotvorno produktivnost izmeriti (torej kako pri vsakem besedotvornem pojavu natančneje določiti njegovo mesto med "zelo tvornimi" in "omrtvelimi").

V nadaljevanju bomo na podlagi izračuna produktivnosti izbranih desetih priponskih obrazil ter na podlagi drugih podatkov o njih prikazali še,

c) kakšen bi lahko bil del besedotvornega poglavja v novi slovnici slovenščine, če bi k njegovi pripravi pristopili korpusno.

Pri tem smo za izhodišče izbrali Toporišičevo *Slovensko slovnico* (2000) in dve tuji slovnici (obe korpusni): Cvrček (2010): *Mluvnice současné češtiny 1, Jak se píše a jak se mluví*, ter Biber in dr. (1999): *Longman grammar of spoken and written English*.

2 BESEDOTVORNA PRODUKTIVNOST: DEFINICIJA IN IZRAČUN

Bauer (2005; gl. pregled tudi v Scherer 2015) obrat k resnejšemu proučevanju morfološke produktivnosti pripisuje Chomskyjevemu prispevku *Remarks on nominalization* iz leta 1970, čeprav zametke prepozna že v zgodnejših delih, najbolj izrazito pri Schultinku (1961)¹ in Zimmerju (1964).² Tako Bauer (prav tam: 318–321) kot Baayen in Lieber (1991) pa prvi poskus kvantitativnega formaliziranja produktivnosti pripisujejo Aronoffovemu delu *Word formation in generative grammar* (1976).

Aronoff v svoji knjigi poglavje z naslovom *Productivity* (35–45) začinja z ugotovitvijo, da prav v produktivnosti korenini naslednje "čudno in trdovratno dejstvo: čeprav so v morfologiji možne mnoge stvari, so nekatere bolj možne kot druge" (prav tam: 35). Na primeru angleških samostalniških priponskih

¹ Schultink, Henk (1961): Produktiviteit als morfologisch fenomeen. *Forum der Letteren* 2, 110–125.

² Zimmer, Karl E. (1964): Affixal negation in English and other languages. *Supplement to Word* 20.

obrazil *-ness* in *-ity*, ki (oz. ko) se pripenjata na pridevnike, ki se končajo na *-ous* (npr. *fabulousness* : *fabulosity*), je nato dokazoval, da dotedanje početje raziskovalcev, tj. ugotavljanje produktivnosti s preprostim štetjem besed z določenim obrazilom – po logiki katerega je bolj produktivno tisto obrazilo, s katerim je tvorjenih več besed – ni skladno z intuitivnim občutkom o večji oz. manjši produktivnosti (v smislu: *perspectiveness* in *perceptivity* sta sicer obe možni, ampak druga se sliši "čudno" ali "izumetničeno" (37)), hkrati pa je poudaril, da tako štetje tudi "ni fer" (36). In sicer že zato ne, ker ne upošteva morfoloških omejitev, npr.: obrazili *-ment* in *-ion* v angleščini obe tvorita samostalnike iz glagolov, vendar pa se *-ion* razvršča samo na latinične glagole, zato bo njegova produktivnost v smislu zgolj štetja besed, ki jih tvori, vedno manjša od produktivnosti obrazila *-ment* (o tem gl. tudi Aronoff in Anshen 1998: 243–244). Aronoffov predlog za izračun indeksa produktivnosti je bil zato naslednji: prešteti je treba število besed, za katere menimo, da bi iz njih z določenim obrazilom lahko nastala tvorjenka, nato preštejemo število dejanskih tvorjenk, ki so nastale na ta način, ter na koncu izračunamo še razmerje med obema vsotama. Indeks produktivnost bi bil v tem primeru torej razmerje med možnimi tvorjenkami z določenim obrazilom in dejanskimi tvorjenkami iz njih.

Avtor se je sicer že sam zavedal, da ima tudi njegov način izračunavanja produktivnosti pasti (npr. to, da številne besede niso uslovarjene, torej jih pri računanju produktivnosti nimamo na seznamu), dokončno pa sta njegov predlog zavrnila Baayen in Lieber (1991), ki sta se problematike lotila s korpusno metodologijo.³ Baayen in Lieber sta pri tem izhajala iz naslednje

³ V Aronoff, Anshen (1998) je razvidno, da je tudi Aronoff pozneje v celoti sprejel Baayenov način merjenja morfološke produktivnosti. Treba je še dodati, da gre kljub dvojnemu avtorstvu razprave Baayen in Lieber (1991) metodologijo pripisati Baayenu, ki je na to temo

(opomba se nadaljuje na naslednji strani)

Schultinkove (1961) definicije produktivnosti (nav. po Baayen, Lieber 1991: 808): *morfološka produktivnost je zmožnost jezikovnih govorcev, da nenamerno tvorijo številne besede, ki se jih načeloma niti ne da dokončno prešteti*. Pri tej definiciji sta se jima zdeli ključni: (a) nenamernost postopka ("če je besedotvorni postopek zares produktiven, potem bodo tvorjenke, ki nastanejo v njem, povsem neopazne" (prav tam: 808)) in (b) omejena zmožnost prešetja ("zelo produktivni besedotvorni postopki bodo načeloma dali neskončno število tvorjenk, medtem ko bodo neproduktivni besedotvorni postopki dali le določeno število tvorjenk, torej število, ki se ga da prešteti" (prav tam)). Njuna formula za izračun morfološke produktivnosti, v katero sta skušala zajeti oboje, se tako glasi:

$$P = \frac{n_1}{N},$$

pri čemer je n_1 število različnih tvorjenk z določenim obrazilom, ki se v vzorcu (korpusu) pojavijo natančno enkrat (hapax legomena), in N skupno število pojavitev vseh tvorjenk s tem obrazilom.⁴ " P je /torej/ ocena verjetnosti, da bomo naleteli na novo, doslej neopaženo tvorjenko, če je velikost vzorca opazovanih tipov tvorjenk enaka N " (prav tam: 809–810). Kar hkrati pomeni, da lahko napovemo tudi, po kateri stopnji bi število novih tvorjenk tega tipa naraščalo, če bi se odločili korpus povečati (prav tam: 811), seveda pa je to naraščanje tangentno, kar pomeni, da P kaže tudi, kako hitro se bo določen način tvorjenja v vzorcu izčrpal. Če je P velik, lahko sklepamo, da je število

leta 1989 doktoriral (Vrije Universiteit, Amsterdam), prim. tudi Lieber (1992: 2, 4), Baayen (1992) in poznejše avtorje, ki izrecno pišejo le o Baayenovem merjenju produktivnosti (npr. Bauer 2005; Schreier 2015). Sicer pa je Baayen s korpusnim pristopom izrazito drugačen od predhodnikov (tudi Aronoffa, kot smo videli) po tem, da so ti produktivnost merili na osnovi slovarskih geslovníkov, ker boljšega vira do takrat pač še ni bilo (k njemu se je pozneje na primeru slovarja *The Oxford English Dictionary* pri proučevanju diahronih morfoloških postopkov sicer spet vrnil Plag (1999)).

⁴ Baayenovih formul je v resnici več, a kot je razvidno tudi iz Bauer (2005), je ta ključna.

tvorjenk opazovanega tipa, ki so ostale zunaj vzorca, še veliko; če je *P* majhen, pa smo v vzorec zajeli skoraj vse tvorjenke tega tipa vsaj enkrat. V prvem primeru gre za produktiven besedotvorni postopek, za katerega je značilno veliko, morda celo neskončno število možnih tvorjenk; v zadnjem primeru pa gre za neproduktiven postopek, pri katerem nastane majhno in seveda končno število tvorjenk (prav tam: 813). Velik *P* posledično dobimo, če je število enkratnih tvorjenk zelo visoko ali če je v vzorcu sicer veliko različnih tvorjenk, a nimajo izrazito visoke pogostosti (prav tam: 815).

Na Baayenov način izračunavanja besedotvorne produktivnosti so se najbolj izrecno odzvali van Marle (1992) ter Frauenfelder in Schreuder (1991). Van Marle je ob izraziti pohvali Baayenu, da je v središče raziskovanja problematike s korpusnim pristopom postavil performanco (prav tam: 152), kritično opozoril na to, da Baayenova formula ni izračun zmožnosti določenih tvorbenih postopkov, da "dajo" določeno število tvorjenk, temveč je kvečjemu izračun stopnje, do katere lahko realno pričakujemo, da bodo določeni besedotvorni postopki podlaga za nove tvorjenke (prav tam: 152). Frauenfelder in Schreuder pa sta bila do Baayenove formule kritična z vidika psiholingvistike: osrednja pomanjkljivost, ki sta jo našla, je bilo Baayenovo neupoštevanje tega, da govorec ob tvorbi novih besed procesira tudi fonološke značilnosti ter semantično transparentnost podstavne ter nove besede, medtem ko Baayenova formula, utemeljena na korpusni pogostosti, upošteva samo hitrost tvorčevega spominskega dostopa do tam že obstoječih jezikovnih podatkov. Na obe kritiki je Baayen obsežno odgovoril v Baayen (1993) in tam ponovno zagovarjal svoj način izračunavanja produktivnosti na podlagi podatkov iz rabe (torej pogostosti), pri čemer zgornja formula po njegovem mnenju ustrezno odgovarja na vprašanje – kot ga je povzel Bauer (2005: 326), kolikšen del tvorjenk, ki so nastale po določenem postopku, je hapaxov. Z odgovorom na to vprašanje pa dobimo tudi podatek, ki ga lahko primerjamo z istovrstnimi podatki drugih besedotvornih pojavov in posledično ločimo tiste pojave, ki so v določenem korpusu bolj produktivni (imajo več hapaxov na skupno število

pojavitve vseh tvorjenk z obravnavanim obrazilom), od tistih, ki so manj produktivni.

Baayenova metodologija je bila v nadaljnjih letih širše sprejeta (prim. npr. že omenjena Aronoff, Anshen 1998; Bauer 2005; Säily 2011; Booij 2012), zato jo bomo uporabili tudi v tukajšnji analizi, skupaj z njim (in poznejšimi avtorji, ki so mu sledili) pa prevzemamo tudi zgoraj že navedeno Schultinkovo definicijo produktivnosti, ki se v natančnem prevodu iz nizozemščine glasi:⁵

morfološka produktivnost je zmožnost jezikovnih govorcev, da s pomočjo morfološkega postopka, ki temelji na oblikovnem in pomenskem ujemanju nekaterih njim znanih besed, nenamerno tvorijo številne nadaljnje besede, ki se jih načeloma niti ne da dokončno prešteti.

S tem obravnavani pojem za potrebe tega prispevka terminološko ožimo na kvantitativni korpusni pristop in ga ločimo od sorodnih izrazov oz. ga z njimi povezujemo zgolj na naslednji način: *pogostost* upoštevamo tako, kot jo prikazuje Baayenova formula; *razširjenost* razumemo kot lastnost, ki izhaja iz korpusno ugotovljene prisotnosti obrazil (če seveda obravnavamo ta) po različnih jezikovnih zvrsteh, besedilnih vrstah, področjih rabe, tematikah ipd. (angl. *domain*, o tem gl. v Plag 1999; tudi Baayen 2009: 19–23); medtem ko *živost* razumemo kot sinonimno poimenovanje za bodisi ničto vrednost produktivnosti *P* (t. i. *mrtvo* morfemsko obrazilo) bodisi vsakršno drugo vrednost *P* (*živo* morfemsko obrazilo).

3 BESEDOTVORNA PRODUKTIVNOST: ANALIZA KORPUSNIH PODATKOV

Analizo smo izvedli v skladu z Baayen in Lieber (1992; za nemščino gl. npr. Evert, Lüdeling (2001), za poljščino Górski (2013: 70–74)), in sicer na korpusu Kres (Logar Berginc in dr. 2012: 77–97). Kres vsebuje 100 milijonov besed oz.

⁵ Za prevod se zahvaljujem lekt. dr. Aniti Srebnik s Filozofske fakultete v Ljubljani.

dobrih 120 milijonov pojavnic iz besedil, ki so izšla med letoma 1990 in 2011. V celotni obseg korpusa 35 % besed prinašajo knjige (od tega 17 % leposlovje, 18 % stvarna besedila), 40 % periodični tisk (20 % časopisi in 20 % revije), 20 % internetna besedila (visoko brani novičarski portali ter predstavitevne spletne strani podjetij in ustanov), preostalih 5 % pa zapisi sej državnega zbora, podnapisi in postproduksijska besedila RTV Slovenija.

Priponska obrazila, ki smo jih za izračun produktivnosti naključno izbrali iz Toporišič (2000: 161–163), so bila naslednja: *-vec*, *-telj*, *-ež*, *-uh*, *-avh*, *-ar*, *-lec*, *-ač*, *-aj* in *-ant*. Vsa se razvrščajo na glagolsko podstavo in tvorijo samostalnik moškega spola s pomenom človeškega vršilca dejanja.

Do korpusa smo dostopali v orodju Sketch Engine (Kilgarriff in dr. 2004), ki ga za slovenščino podpira Center za jezikovne vire in tehnologije Univerze v Ljubljani.⁶ Sezname smo pridobili z iskalnimi pogoji

```
[lemma=".+vec" &tag="Som.*"],
```

```
[lemma=".+telj" &tag="Som.*"] itd.,
```

nato pa smo na podlagi ogleda konkordanc ročno izločili najprej netvorjene besede, dalje pa še tvorjenke, ki niso iz glagola ali nimajo pomena človeškega vršilca dejanja, zloženke in druge neizpeljavne besedotvorne vrste ter zatipkane besede in napačno označena lastna imena. Pri manjšem delu tvorjenk smo sprejeli še naslednji dve odločitvi:

a) Kjer sta bili možni dve poti nastanka tvorjenke, tj. bodisi iz glagola (*žvižgač* < *žvižgati* < *žvižg*) bodisi iz samostalnika (*žvižgač* < *žvižg* < *žvižgati*), smo dosledno upoštevali prvo možnost in smo torej tvorjenko na naš seznam uvrstili.

b) Izmed tvorjenk, ki so v konkordancah izkazovale tako poimenovanje

⁶ <https://www.cjvt.si/>

človeškega kot nečloveškega (npr. *nakladač*, *rezkar*, *pomivalec*, *upravljavec*, *nosilec*), smo na seznam uvrstili vse tiste, ki so lastnost človeškega izkazovale vsaj enkrat. Pri tem pojavitev s kategorijo nečloveško nismo odšteli – če bi namreč hoteli biti v takem odštevanju dosledni, bi morali natančno ročno pregledati prav vse konkordance iskanih tvorjenk, to pa bi nam vzelo izredno veliko časa (vseh konkordanc iskanih tvorjenk je bilo namreč skoraj 365.000, med njimi so nekatere, pri katerih je že hiter ogled konkordanc potrdil, da imajo pomen tako človeškega kot nečloveškega, zelo pogoste, npr. že omenjeni *nosilec* ima 5.546 zadetkov, *povzročitelj* jih ima 1.542, *upravljavec* 1.407 itd.). Ocenili smo, da ima prisotnost kategorije nečloveško v tem obsegu na končne rezultate te študije le zanemarljiv vpliv.

Tabela 1 prikazuje številske podatke opisanega polavtomatskega postopka in končni izračun produktivnosti P po Baayenovi formuli.

Priponsko obrazilo	Število pojavitev	Število različnic	Število iskanih tvorjenk med različnicami	Število vseh pojavitev iskanih tvorjenk	Število hapaxov med iskanimi tvorjenkami	P
<i>-vec</i>	85.549	2.153	50	35.916	17	0,00047
<i>-telj</i>	105.609	461	168	101.201	41	0,00041
<i>-ež</i>	58.324	926	14	1.092	5	0,0046
<i>-uh</i>	38.462	210	15	5.175	2	0,00039
<i>-avh</i>	28	11	0	0	0	0
<i>-ar</i>	262.210	4.004	77	19.746	14	0,00071
<i>-lec</i>	239.456	1.948	662	179.718	213	0,0012

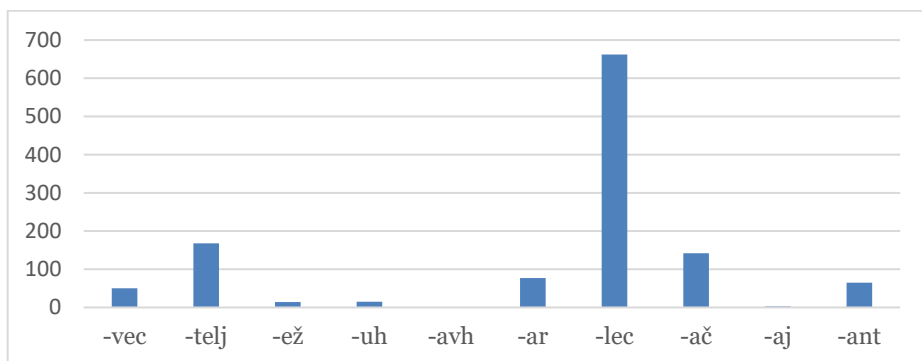
-ač	22.882	652	142	12.879	34	0,0026
-aj	130.632	1.053	3	1.251	0	0
-ant	34.971	812	65	7.981	10	0,0013

Tabela 1: Število pojavitev samostalniških tvorjenk iz glagolov, moškega spola in s pomenom človeškega vršilca dejanja z 10 izbranimi priponskimi obrazili v korpusu Kres ter njihova produktivnost *P*.

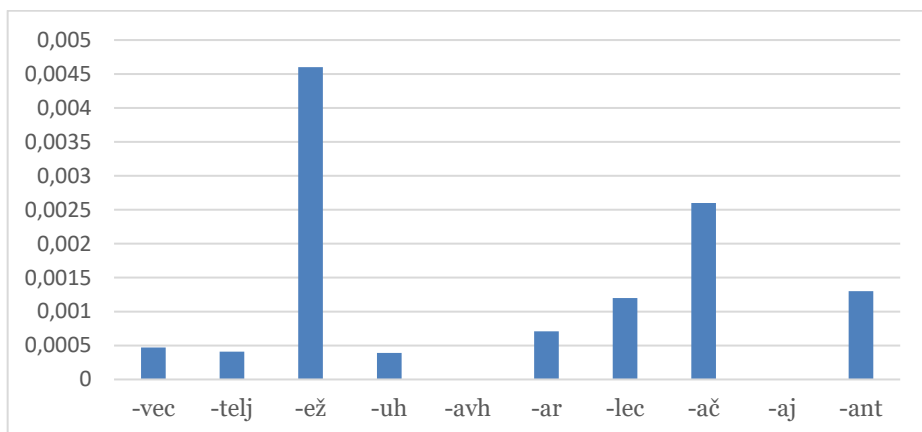
V Tabeli 1 je razvidno, da je bilo v Kresu daleč največ različnih besed z izbranim pomenom tvorjenih z obrazilom *-lec* (gl. 4. stolpec in Sliko 1), npr. *igralec*, *prebivalec*, *obiskovalec*, *bralec*, *gledalec*, *izvajalec*; sledili sta obrazili *-telj* in *-ač* (*prijatelj*, *učitelj*, *vlagatelj*, *pisatelj*, *voditelj*, *ljubitelj*; *navijač*, *dirkač*, *potapljač*, *tekač*, *berač*, *kovač* itd.), nato pa *-ant* in *-vec* (*diplomant*, *reprezentant*, *maturant*, *muzikant*, *projektant*, *emigrant*; *delavec*, *pevec*, *volivec*, *upravljavec*, *razstavljavec*, *pivec*) ter *-uh* in *-ež* (*varuh*, *potepuh*, *ovaduh*, *požeruh*, *oderuh*, *stremuh*; *vedež*, *klatež*, *hlinež*, *burkež*, *rogovilež*, *kurbež*). Obrazilo *-aj* v iskanem pomenu v Kresu izkazujejo le 3 tvorjenke (*strežaj*, *čuvaj*, *točaj*), z obrazilom *-avh* pa v Kresu ni nobene iskane tvorjenke. Tudi sicer so razlike v njihovi pogostosti zelo velike (*točaj* ima npr. 146 pojavitev, *igralec* pa kar 20.608).

Ugotovljeno obenem še ne pomeni, da je *-lec* v izbranem vzorcu tudi najbolj produktivno obrazilo, tj. najbolj produktivno v pomenu *P*. Če smo natančni, smo potrdili le naslednje: za *-lec* je značilno, da se na glagolsko podstavo razvršča z malo omejitvami (je pomensko nevtralen in fonetično široko družljiv, ker se pripenja na končni samoglasnik glagolske osnove *-a-* ali *-i-* (Toporišič 2000: 163)), ugodno pa je tudi, da ima slovenščina zelo veliko zelo pogostih glagolov, ki nase pač lahko sprejmejo to obrazilo. Taka spodbudna podlaga seveda prinese veliko število tvorjenk, vendar pa – če pogledamo z druge strani – to še ne pomeni, da bo obrazilo *-lec* tudi najbolj verjetna prva izbira pri tvorbi

naslednje nove slovenske besede (in vsake nadaljnje) s pomenom človeškega vršilca dejanja moškega spola. Razmerje med številom hapaxov (6. stolpec Tabele 1) sicer na prvi pogled med izbranimi obrazili kaže približno enaka razmerja kot število različnih tvorjenk, vendar pa izračun vrednosti P , torej produktivnosti po Baayenovi formuli (7. stolpec v Tabeli 1), izpostavlja druge "dobitnike medalj" (Slika 2).



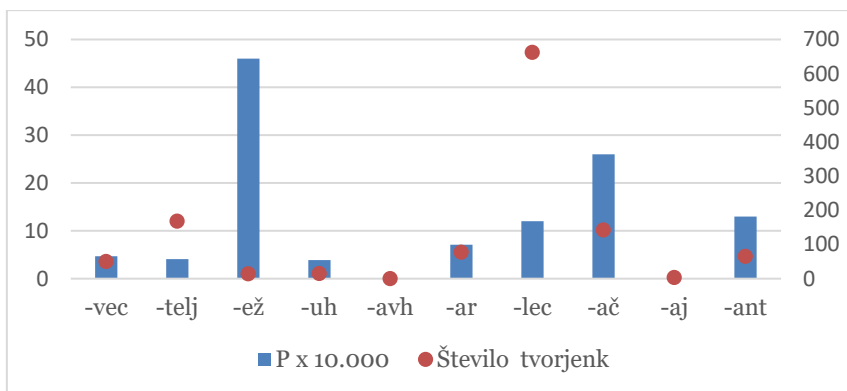
Slika 1: Število različnih samostalniških tvorjenk iz glagolov, moškega spola in s pomenom človeškega vršilca dejanja z 10 izbranimi priponskimi obrazili v korpusu Kres.



Slika 2: Produktivnost P 10 izbranih priponskih obrazil samostalniških tvorjenk iz glagolov, moškega spola in s pomenom človeškega vršilca dejanja v korpusu Kres.

Slika 2 kaže, da je v izbranem vzorcu najbolj produktivno obrazilo *-ež*; med le 14 tvorjenkami s skupnim razmeroma nizkim številom vseh pojavitev (1.092) je namreč v Kresu kar 5 hapaxov, ki se končajo na *-ež*: *prdež*, *pofuklež*, *pizdež*, *buljež* in *štrbunkež*. Dalje smo izluščili kar 34 hapaxov z obrazilom *-ač*: *čofotač*, *zavirač*, *zapravljač*, *stokač*, *savnač*, *ravsac* itd., kar to obrazilo po produktivnosti *P* uvršča na drugo mesto, sledi pa obrazilo *-ant* (npr. *rehabilitant*, *nategant*, *malverzant*, *mahinant*, *konzultant*, *kapitulant*). Obrazilo *-lec* je tako šele na 4. mestu s sicer v absolutnem številu velikim številom hapaxov (213, npr. *zalagalec*, *terjalec*, *prebujevalec*, *zajtrkovalec*, *točkovalec*, *sesljalec*), na 5. mestu je *-ar* (*savnar*, *plužar*, *pizdar*, *linčar*, *driblar*, *balincar*), nato pa s skoraj enakimi vrednostmi sledijo še *-vec*, *-telj* in *-uh* (*zaklinjavec*, *pospravljavec*, *hlastavec*, *premagavec*, *izjavljavec*, *belivec*; *žalitelj*, *pesnitelj*, *mrzitelj*, *obdavčitelj*, *opremitelj*, *prilastitelj*; *zaležuh*, *oprezuh*). Poleg obrazila *-avh*, za katerega nismo našli nobene tvorjenke z obravnavanim pomenom, nobenega hapaxa ni imelo še obrazilo *-aj*.

Razmerja med produktivnostjo *P* in številom različnih tvorjenk s posameznimi obrazili, ki smo jih zgoraj obravnavali ločeno, spodaj združeno prikazujemo še na Sliki 3.



Slika 3: Razmerja med produktivnostjo *P* in številom različnih tvorjenk z 10 izbranimi priponskimi obrazili samostalniških tvorjenk iz glagolov, moškega spola in s pomenom človeškega vršilca dejanja v korpusu Kres.

Na Sliki 3 je hitro opazno, da so razlike v razmerjih največje pri obrazilih *-ež* (v smislu veliko večje produktivnosti *P* kot števila različnih tvorjenk) ter *-ač* in *-ant* (zopet v smislu večje produktivnosti) – v drugi smeri (torej kot precej manjša produktivnost v primerjavi s številom različnih tvorjenk) pa pri *-lec* in *-telj*. Če torej za kazalnik produktivnosti vzamemo razmerje med številom hapaxov, ki imajo določeno obrazilo, in številom pojavitev vseh tvorjenk, ki imajo to obrazilo, potem je izmed 10 izbranih obrazil, ki v slovenščini iz glagolov tvorijo samostalnike moškega spola s pomenom vršilca dejanja, najbolj produktivno – ponavljamo – obrazilo *-ež*. Ali še drugače: sodeč po deležu hapaxov med vsemi tvorjenkami z izbranimi obrazili, je najbolj verjetno, da bo imela naslednja nova tvorjenka iz glagola in z obravnavanim pomenom obrazilo *-ež*, sledil mu bo *-ač*, nato *-ant*, *-lec* itd.

Na zgornje dejstvo seveda lahko vplivajo tudi značilnosti podstavnih glagolov (glagolska pripona, fonetične značilnosti, naglas). Hiter pregled je pokazal, da se vsa obravnavana obrazila razvrščajo na besedotvorne podstave iz glagolov na *-a-ti* in *-i-ti* (lahko tudi *-eva/-ova-ti* in *-e-ti*), izjema je le obrazilo *-ant*. To obrazilo je edino, ki se izmed 10 v našem naboru pripenja na besedotvorno podstavo, ki je nastala iz prevzetih glagolov na *-ira-ti*, npr. *emigrant* < *emigrirati*, *projektant* < *projektirati*, *simulant* < *simulirati* (pravzaprav je sama tvorjenka največkrat pogostejša od postavnega glagola ali pa ga celo edina dokazuje, npr. *visitant* < **vizitirati*, čeprav slednje velja tudi za tvorjenke (sploh hapaxe) z drugimi obrazili, npr. *popackovalec* < **popackovati*). V nekaj primerih se je *-ant* sicer razvrstil še na besedotvorno podstavo iz glagolov na *-a-ti* oz. *-ava-ti* (*prevarant*, *zabušant*, *zafrkant*, *zajebant* in *nategant*). Pri vseh drugih obrazilih pa, kot rečeno, v zvezi z glagolsko pripono podstavnega glagola ni posebnosti. Tudi fonetičnih omejitev na morfemskem šivu pri izbranih obrazilih ni, z izjemo obrazil *-lec* in *-vec*, pri katerih je drugi v prednosti, če se *l* ali *lj* pojavi že v besedotvorni podstavi (npr. *pripravljavec* : *pripravljalec*). Z izjemami v mislih lahko torej sklepamo, da je primerjava izračunane produktivnosti med izbranimi obrazili relevantna.

4 SLOVNIČNI OPIS

Zdi se, da je produktivnost v širšem pomenu besede že v samem izhodišču tudi v slovničnih opisih slovenskega besedotvorja. Tako npr. Breznik v uvodu v ta del svoje slovnice ugotavlja: "Večina slovenskih besed je nastala po izpeljavi" (1934: 158), tej ugotovitvi pa nato neposredno sledi prav obravnava omenjene besedotvorne vrste. Tudi sicer se slovnična poglavja o besedotvorju v slovenščini praviloma začenjajo z izpeljavo (gl. npr. Janežič, Sket 1900: 124–157; Breznik 1934: 158–179; Toporišič 2000: 161–232; Herrity 2000: 347–362), zaključujejo pa z besedotvorno vrsto, po kateri nastane najmanj tvorjenk – v zadnji Slovenski slovnici (Toporišič 2000) je to pri tvorbi samostalnikov, pridevnikov in prislovov sklapljanje.

4.1 Seznam obrazil

Slovnične opise besedotvornih značilnosti večinoma sestavljajo sezname obrazil skupaj z zgledi, kakršne za slovenščino, češčino in angleščino na primeru priponskih obrazil za tvorbo samostalnikov prikazujejo spodnje slike. Ti sezname niso osnovani na vseobsegajočnosti in izčrpnosti, temveč zopet – tako je mogoče sklepati – prav na lastnosti, o kateri razmišljamo tu: produktivnost – oz. če smo natančni: avtorji v odstavkih pred temi sezname ali neposredno za njimi pišejo, kot smo deloma že videli, o veliki oz. manjši tvornosti, produktivnosti, pogostosti, tipičnosti in živosti (Toporišič 2000: 160–161; Cvrček 2010: 81–124; Biber in dr. 1999: 320–325).

Priponsko obrazilo	Zgled
-ø	<i>prêrok, spàk, pék (zelo zelo redko);</i>
-a	<i>slúga;</i>
-ja	<i>vódja;</i>
-c	<i>ígrc;</i>
-ec [æc]	<i>bórec, hválec, govórec, jézdec, písec, vídec, jédec, lóvec, védec, strélec, jókec;</i>
-(á)lec [æc]	<i>brálec, dvigálec, malikoválec, sesálec, godálec;</i>
-(í)lec [æc]	<i>branílec, gasílec, pletílec, drobílec, gnetílec, volílec;</i>

-í nec [əɕ]	<i>mrgolí nec;</i>
-unec [əɕ]	<i>begúnec;</i>
-vec [əɕ]	<i>pívec, brívec, pévec, mlévec, klávec;</i>

Slika 4: Priponska obrazila za izpeljavo samostalnikov s pomenom vršilca dejanja moškega spola in zgledi v Toporišič (2000: 161–162), del.

Priponsko obrazilo	Zgled
-a	<i>bambula, něha, osma, chvála, čiča;</i>
-ba	<i>honba;</i>
-oba	<i>staroba;</i>
-tba	<i>kletba;</i>
-da	<i>strejda, pravda, obejda;</i>
-ajda	<i>žákajda;</i>
-nda	<i>šeptanda;</i>
-anda	<i>fešanda;</i>
-uha	<i>ostruha</i>
-echa	<i>macecha</i>

Slika 5: Priponska obrazila za izpeljavo samostalnikov in zgledi v Cvrček (2010: 93), del.

Priponsko obrazilo	Pomen(i)	Zgled
-age	'collection of N'	<i>baggage, leafage, plumage</i>
	'action/result of V'	<i>breakage, haulage, wastage</i>
	'cost of N/V-ing'	<i>brokerage, haulage, postage</i>
	'measure in N-s'	<i>acreage, mileage, tonnage</i>
	'place for N'	<i>hermitage, orphanage, vicarage</i>
-al	'action/instance of V-ing'	<i>arrival, burial, denial</i>
	-an, -ian	'person who lives in N'
'language of N'		<i>Estonian, Korean</i>
'person associated with N'		<i>Darwinian, Victorian</i>

-ance, -ence	'action or state of V-ing'	<i>adherence, assistance, resemblance</i>
	'state of being A'	<i>abundance, dependence, ignorance</i>
-ant, -ent	'person who V-s'	<i>assistant, consultant, student</i>
	'something used for V-ing'	<i>coolant, defoliant, intoxicant</i>
-cy	'state or quality of being A/N'	<i>accuracy, adequacy, infancy</i>
-dom	'state of being A/N'	<i>boredom, freedom, stardom</i>
-ee	'person who has been or is to be V-ed'	<i>employee, evacuee, trainee</i>
	'person to whom something has been or is to be V-ed'	<i>assignee, licensee</i>
	'person who V-s or has V-ed'	<i>escapee, retiree, standee</i>
	'person who is A'	<i>absentee, devotee</i>
-er, -or	'person who V-s'	<i>advertiser, driver, governor</i>
	'something used for V-ing'	<i>computer, filler, silencer</i>
	'person concerned with N'	<i>astronomer, footballer, geographer</i>
-ery, ry	'person living in N'	<i>cottager, Londoner, New Yorker</i>
	'action/instance of V-ing'	<i>bribery, robbery, thievery</i>

Slika 6: "Nekatera značilna /angl. *common*/" priponska obrazila za izpeljavo samostalnikov, njihov pomen in zgledi v Biber in dr. (1999: 321), del.

V novi slovnici slovenščine bi bilo treba take prikaze pripraviti na korpusnih podatkih, pri čemer predlagamo, da bi veljalo naslednje:

a) pri izboru obrazil in obrazilnih morfemov se upošteva njihova produktivnost, izračunana po Baayenovi (ali kateri drugi, v prihodnje nastali) formuli, kar pomeni, da se obrazila in obrazilni morfemi z ničelno ali zelo nizko vrednostjo *P* (meja je dogovorna) izpustijo;

b) pri zaporedju obrazil in obrazilnih morfemov se upošteva število različnih tvorjenk, katerih del so (in ne npr. abeceda);

c) pri zgledih se upoštevata število in besedilna razpršenost korpusnih pojavitev tvorjenk, pri čemer enkratnih tvorjenk ali tvorjenk z nizkim številom pojavitev (meja je zopet dogovorna) med zgledi ne navajamo.

V skladu s povedanim bi bila slovnicična slika tu obravnavanih priponskih obrazil skupaj z zgledi naslednja (Slika 7):

Priponsko obrazilo	Zgled
-lec	<i>igravec, prebivalec, obiskovalec, bralec, gledalec, izvajalec, raziskovalec, proizvajalec, nosilec, zmagovalec</i>
-telj	<i>prijatelj, učitelj, vlagatelj, pisatelj, voditelj, ljubitelj, predlagatelj, skladatelj, ravnatelj, ustanovitelj</i>
-ač	<i>navijač, dirkač, potapljač, tekač, berač, kovač, krojač, veslač, jahač, nosač</i>
-ar	<i>slikar, vladar, natakhar, kuhar, smučar, stražar, šolar, ropar, zidar, romar</i>
-ant	<i>diplomant, reprezentant, maturant, muzikant, projektant, emigrant, komandant, protestant, prevarant, migrant</i>
-vec	<i>delavec, pevec, volivec, upravljavec, razstavljavec, pivec, brivec, razpravljavec, klavec, pripravljavec</i>
-uh	<i>varuh, potepuh, ovaduh, požeruh, oderuh</i>
-ež	<i>vedež, klatež, hlinež, burkež, rogovilež</i>

Slika 7: Izbrana priponska obrazila za izpeljavo samostalnikov s pomenom vršilca dejanja moškega spola in zgledi – prikaz na podlagi podatkov iz korpusa Kres.

S Slike 7 smo izpustili obrazili *-avh* in *-aj*, ki v Kresu za iskani tip tvorjenk

izkazujeta ničelno produktivnost *P* (gl. Tabela 1); obrazila smo razvrstili v skladu z različnim številom tvorjenk, katerih del so (največ jih ima *-lec*, sledi *-telj* itd.); zgledi pa so navedeni po korpusni absolutni pogostosti, pri čemer smo s številom zgledov (10 : 5 – števili sta arbitrarni) posredno nakazali še, kateri dve obrazili sta del znatno manjšega nabora tvorjenk, ki so obenem tudi manj pogoste.⁷

Hkrati naj opozorimo, da smo za razliko od Toporišičeve slovnice naglas obrazil zaenkrat zanemarili (prim. npr. *slik-ár* : *kúh-ar*), ker menimo, da bi bile tu nujno potrebne predhodne sodobne fonetične in fonološke potrditve (Toporišič npr. loči med *-ánt*, npr. *diplománt*, in *-ànt*, npr. *zabušànt* – prim. Jurgec (2011), ki ugotavlja, da je treba ločiti med samoglasnikom *a* in (novim) srednjim nizkim samoglasnikom *ʌ*, ki se razlikujeta po formantnih frekvencah, torej kvaliteti, ne po kvantiteti (dolžini)). Kljub zgledu v Biber in dr. (1999) pa smo se odločili, da ne vključimo informacije o pomenu obrazil, ker je ta okvirno vsebovan že v poimenovanju celotne skupine: človeški vršilec dejanja.

4.2 Komentar seznama obrazil

Avtorji slovnice sezname obrazil komentirajo predvsem z vidika omejitev njihove pojavnosti in razvrščanja ter modifikacij. Kot npr.:

Priponsko obrazilo *-ec* krni glagolsko pripono in se sklaplja s soglasniškim izglasjem podstave (*govor-iti* + *-ec* > *govorec*) (Toporišič 2000: 163).

Razlog, zaradi katerega so predpanska obrazila manj produktivna kot priponska, je verjetno v tem, da je večina prvih grškega izvora, medtem ko so skoraj vsa priponska obrazila, naštetá v 4.8.1.2, po izvoru romanska ali germanska (Biber in dr. 1999: 324).

Pri tvorbi besed iz samostalnikov lahko nastanejo podstave, ki se na besedotvorni

⁷ V eventualni elektronski izdaji so seveda možni še daljši sezname tako obrazil kot zgledov ter preklop v korpusne pojavitve, pa seveda tudi v druge spletno dostopne jezikovne vire.

meji zaključujejo s podvojenimi glasovi (-ss-). Podvojeni -ss- je na takem mestu predvidljiv, tako priponsko obrazilo v skladu s tradicijo razumemo kot obrazilo *-ský*: / *Brandýs(ø)* – podstava: *Brandýs* > *brandýský*; priponsko obrazilo *-sk(ý)* (Cvrček 2010: 87).

V slovnici avtorjev Biber in dr. (1999) seznamu obrazil, katerega del smo prikazali na Sliki 6, neposredno sledi še poglavje z naslovom *Pogostost značilnih priponskih obrazil za izpeljavo samostalnikov*. Poglavje začenja trditev: "Distribucija izpeljanih/⁸/ samostalnikov, kot tudi samostalnikov nasploh, je po zvrsteh /angl. *registers*/ zelo različna" (prav tam: 322). V nadaljevanju avtorji navajajo ugotovitve, kakršni sta npr. naslednji dve: (a) pogovorna zvrst ima daleč najmanj izpeljanih samostalnikov in pa (b) z eno izjemo (*-ness* pri leposlovju) relativna pogostost izpeljanih samostalnikov od pogovorne zvrsti prek leposlovja in časopisja do akademskih pisnih besedil hitro raste (prav tam). Komentar dopolnjuje graf, v katerem je prikazano število tvorjenk s štirimi najbolj značilnimi priponskimi obrazili (*-tion*, *-ity*, *-ism* in *-ness*), ki tvorijo abstraktne samostalnike, in iz katerega je npr. razvidno, da je tvorjenk z obrazilom *-tion* v leposlovju trikrat toliko kot v pogovorni zvrsti, nadalje v časopisju 9-krat toliko, v akademskih besedilih pa kar 22-krat toliko.

Tudi zadnja Slovenska slovnica povsem brez zvrstnih komentarjev ni, kot smo videli že zgoraj v Uvodu v navedku iz Toporišič (2004: 160–161) v zvezi s tvorjenko *Smrekarjeva* (: *Smrekarica*), ki da se rabi v "meščanskih krogih, oficialno" (prav tam: 161), še pogostejše pa so v Toporišičevi slovnici stilne (ekspresivne) opombe tipa: priponska obrazila z *u* (npr. *smrduh*) so slabšalna, podobno se slabšalnost "drži ustnične prvine tudi v priponskih obrazilih *-avs*, *-avt* ipd." (prav tam: 164).

Oboji – pojavno-razvrstilni in zvrstno-stilni – komentarji obrazil so za opis besedotvornih pojavov pomembni, zato bi jih bilo smiselno ohraniti tudi v novi

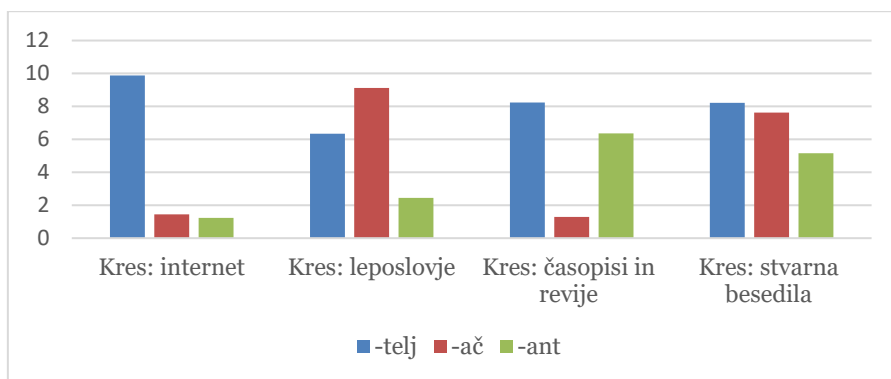
⁸ Angl. termin *derived nouns* sicer vključuje tudi sestavljenke.

slovnici, pri čemer bi morali avtorji pri zvrstno-stilnih komentarjih natančno upoštevati značilnosti besedilne in tematske zgradbe uporabljenega korpusa (oz. korpusov).

Zgolj za ponazoritev grafičnega elementa, omenjenega v zvezi s slovnico Biber in dr. (1999), prikazujemo še podkorpusno razporeditev treh naših obrazil, in sicer obrazil *-telj*, *-ač* in *-ant* (Tabela 2 in Slika 8).

Priponsko obrazilo	Kres: internet	Kres: leposlovje	Kres: časopisi in revije	Kres: stvarna besedila
<i>-telj</i>	9,88	6,34	8,24	8,22
<i>-ač</i>	1,45	9,12	1,29	7,62
<i>-ant</i>	1,23	2,45	6,36	5,16

Tabela 2: Število pojavitev samostalniških tvorjenk iz glagolov, moškega spola in s pomenom človeškega vršilca dejanja s 3 izbranimi priponskimi obrazili v delih korpusa Kres (na 100.000 pojavnic).



Slika 8: Število pojavitev samostalniških tvorjenk iz glagolov, moškega spola in s pomenom človeškega vršilca dejanja s 3 izbranimi priponskimi obrazili v delih korpusa Kres (na 100.000 pojavnic).

Iz Tabele 2 in s Slike 8 je razvidno, da so tvorjenke z obrazilom *-telj* z iskanim pomenom najbolj pogoste v internetnem delu Kresa, najmanj pa v leposlovju, skoraj enak delež imajo v periodiki in priročnikih; vendar pa tu razlike niso zelo velike. Večje so pri tvorjenkah z obrazilom *-ač*, ki jih je v internetnih besedilih ter v časopisih in revijah izrazito manj kot v leposlovju, pa tudi v stvarnih besedilih so precej pogoste. Tvorjenke z obrazilom *-ant* so najbolj pogoste v revijah in časopisih, sledijo stvarna besedila, malo in še manj pa jih je v leposlovju ter v internetnem delu Kresa.

Nasplošno naj še dodamo, da je take prikaze na podlagi korpusnih podatkov mogoče pripraviti za številne in zelo različne besedotvorne pojave, zato bi bilo treba razmisliti o tem, kako njihov prikaz v temeljnem jezikovnem viru, slovnici, omejiti na najbolj izstopajoče in povedne; vsekakor pa bi bilo treba pri tem poleg tvorjenk iz predvidljivih tvorbeno-pretvorbenih postopkov upoštevati tudi tvorjenke, ki niso systemske (Logar 2006; Voršič 2013). Odprto puščamo tudi vprašanje širše zgradbe celotnega slovnicega besedotvornega poglavja, zdi pa se, da bi tudi prikaz besedotvornih vrst nasploh lahko sledil korpusnim podatkom.

5 RAZPRAVA

Uporaba velikih korpusov ima v analizi vsakršnih jezikovnih pojavov mnoge prednosti (npr. Leech 1992; Tognini - Bonelli 1993; Čermák 1995/2005; Biber in dr. 1998; Teubert 1999/2005; Kennedy 1999), v analizi besedotvorne ravnine jezika pa hkrati tudi eno pomembno slabost: velik obseg ročnega dela. Bogato oblikoslovje v slovenščini dopolnjuje prav tako pestra besedotvorna morfološkost s sedmimi jezikovnosistemskimi besedotvornimi vrstami (Vidovič Muha 1988) in močno stopenjskostjo (Kern 2017). Poleg tega ima slovenščina glasovno ničto priponsko obrazilo (tudi pripono ter medponsko obrazilo in medpono) ter obrazilne morfeme, ki so v deležu, ki ga ne moremo zanemariti, še končniško homonimni. Naš izbor obrazil in tip tvorjenk, ki smo ga analizirali, je bil v tem smislu razmeroma nezahteven: omejili smo se na

neničta obrazila, ki niso končniško homonimna (kot je npr. prav obrazilo *-ø* ali pa obrazilo *-a*), in se ustavili pri številki 10 (Toporišičev seznam je veliko daljši: vsebuje kar 54 morfemov (2000: 161–163)),⁹ iskali smo samostalniške tvorjenke (ne pa npr. prislovnih, ki bi prinesle še izločanje napačno oblikoskladenjsko označenih pridevnikov (prim. Grčar in dr. 2012)). Poleg tega smo izbrali izpeljavno besedotvorno vrsto, zanemarili pa besedotvornomorfemsko zahtevnejše (vsaj kar se avtomatskega pridobivanja podatkov tiče) zlaganje, pa tudi sestavljanje (sploh glagolsko). Kot že omenjeno, ne smemo pozabiti na napake pri korpusnem označevanju: te sicer niso predstavljale težav v primerih, ko so napačne "tvorjenke" prišle na seznam (takrat smo jih pač ročno izločili), zelo verjetno pa obstaja v korpusu še kakšna tvorjenka, ki se je zaradi napake v označitvi izognila priklicu (sploh če je šlo za hapaxe, za katere smo videli, da so pri izračunu produktivnosti ključni). Več ročnega dela obenem pomeni tudi večji vpliv človeške napake (brez dvoma tudi v tem prispevku). Opozoriti je treba tudi na to, da smo tu ravnali po načelu delnega korpusnega pristopa (Gorjanc 2005: 24), medtem ko bi bil za končni slovnični opis ustreznejši popolni korpusni pristop s kolikor se da predprocesiranimi podatki.¹⁰

Baayenove formule za izračun morfološke produktivnosti še danes veljajo za sicer ne idealne, ob natančni interpretaciji pa vendarle znanstveno sprejemljive. Pri tem ne smemo spregledati avtorjevega opozorila (Baayen,

⁹ V lasten zagovor lahko povemo, da so se tudi vsi raziskovalci, na katere smo se z literaturo opirali pri obravnavi besedotvorne produktivnosti, omejili le na nekaj obrazil (od 2 do pribl. 10).

¹⁰ Možnost večjega avtomatskega predprocesiranja podatkov za izračun morfološke produktivnosti sta na primeru nemščine testirala Evert in Lüdeling (2001) ter pri tem ugotovila, da sistema, ki bi dal za ta namen dovolj kakovostne podatke brez ročnega pregleda, še ni. Za slovenščino si je mogoče obetati vsaj delno tovrstno pomoč z zbirkami podatkov, ki bodo nastale pri projektu *Nova slovnica sodobne standardne slovenščine: viri in metode* (ARRS, 2017–2019, <http://slovnica.ijs.si/>).

Lieber 1998: 820; tudi Bauer 2005: 325), da produktivnost P velja samo za določen korpus, kar pomeni, da P -ji, izračunani na podatkih iz različnih korpusov, med sabo niso primerljivi. Vpliv korpusne zgradbe na to meritev je torej izrazit.

Baayen na več mestih utemeljuje, da so njegove formule precej skladne z jeziko(slov)no intuicijo (Baayen, Lieber 1991: 801, 808, 811, 817; Baayen 1992: 11; Baayen 1993: 81). Ker gre za subjektivno lastnost, deloma priučeno na preteklih načinih interpretiranja raziskovalnih ugotovitev (pri jezikoslovcih) bodisi temelječo na jezikovnem občutku (pri teh in vseh drugih govorcih jezika), se je z avtorjem mogoče tako strinjati kot mu nasprotovati. V slovenskem prostoru vrednosti P kot "matematično izražene jezikoslovne predstave morfološke produktivnosti" (Baayen, Lieber 1991: 811) do zdaj še nismo preizkusili, zato je na neki način tudi nismo navajeni, a vrednost zgoraj prikazanega je vsekakor v objektivnosti, ki deluje prepričljivo. Vseeno pa – če se nadalje še enkrat navežemo na zgornjo Sliko 7 – vrednosti P tu nismo izrazito postavili v središče slovničnega opisa. Ocenili smo, da je pri prikazu obrazil še vedno boljše izhajati iz števila in besedilne razpršenosti korpusno izkazanih tvorjenk, ker je ta količinskost za opis tehtnejša (in priznajmo: tudi intuitivnejša). Pri tem smo skladni z Biber in dr. (1999), ki pod naslovom *Produktivnost značilnih samostalniških izpeljank* (prav tam: 323) naslovni termin definirajo kot skupno število različnih tvorjenk, ki imajo določeno obrazilo, in posledično med drugim npr. povzemajo, da je "obrazilo *-tion* daleč najbolj produktivno tako v smislu skupnega števila samostalnikov, ki jih tvori, kot v skupini relativno redkih tvorb (tj. izpeljanih samostalnikov, ki se pojavijo manj kot desetkrat na milijon besed)" (prav tam). Kljub temu pa se produktivnost P zdi upoštevanja vredna pri izbiri prikazanih obrazil (natančneje: opustitvi neproduktivnih) in pri njihovi zvrstno-stilni interpretaciji (sploh ko gre za sopomenska obrazila). Je pa – ponavljamo – izredno pomembno, kakšen je korpus, iz katerega za take izračune pridobivamo podatke.

Kljub temu, da smo se tu ukvarjali le s priponskimi obrazili, velja ob koncu razprave omeniti še to, da bi Baayenu lahko sledili tudi pri izračunu produktivnosti posameznih besedotvornih vrst. Čeprav bi tak podvig zahteval ročni pregled še veliko večje količine podatkov, bi bil zanimiv (in morda v marsičem presenetljiv).

6 SKLEP

Obravnava besedotvornih prvin slovenščine ima močno tradicijo in je tudi v sodobnosti eno živahnejših raziskovalnih področij (prim. npr. Michelizza 2008; Vidovič Muha 2009; Gložančev in dr. 2009; Stramljič Breznik 2013; Stramljič Breznik 2016; Kern 2014; Štumberger 2015; Žele 2016; Sicherl, Žele 2018). Veliki korpusi lahko tovrstne študije vsekakor podatkovno okrepijo – raziskovalce pa spodbudijo tudi k pripravi novih teoretičnih izhodišč –, kar pa še ne pomeni, da olajšajo analitični del takih raziskav. A to je lahko hkrati tako slabost kot prednost; namreč v domeni, ki je bila v času prve izdaje zadnje Slovenske slovnice (Toporišič 1976) še povsem jezikoslovna, lahko zdaj poznavalcem morfemske zgradbe jezika pomagajo tudi računalničarji, metodologi in matematiki.

Kot je bilo v prispevku razvidno, je pri korpusnem pristopu k besedotvorju malo zadreg le pri naboru zgledov. Če prvi del analize podatkov nekako še zmorejo strokovnjaki za morfematiko, pa bi bilo treba k pripravi komentarja obrazil, zajetih v slovnični opis, povabiti tudi poznavalce stilistike, besedilne zvrstnosti, normativistike in etimologije. Pri prihodnjem slovničnem prikazu besedotvornega dela slovenščine tako brez timskega dela ne bo šlo – vsaj če se odgovorno zavežemo nalogi, da delamo v vseh pogledih sodoben jezikovni opis (in predpis).

ZAHVALA

Članek je nastal v okviru programskega financiranja ARRS. Raziskovalni program št. P6-0215 je sofinancirala Javna agencija za raziskovalno dejavnost

Republike Slovenije iz državnega proračuna.

LITERATURA

Aronoff, M. (1976): *Word formation in generative grammar*. Cambridge: MIT Press.

Aronoff, M., in Anshen, F. (1998): Morphology and the lexicon: Lexicalization and productivity. V A. Spencer in A. M. Zwicky (ur.): *The handbook of morphology*: 237–246. Oxford: Blackwell Publishers.

Baayen, H. (1993): On frequency, transparency and productivity. V G. Booij in J. van Marle J. (ur.): *Yearbook of morphology 1992*: 181–208. Dordrecht: Springer.

Baayen, H. (1994): Productivity in language production. *Language and cognitive processes*, 9 (3): 447–469.

Baayen, H. (2009): *Corpus linguistics in morphology: morphological productivity*. V A. Lüdeling in M. Kytö (ur.): *Corpus linguistics: An international handbook*: 900–919. Berlin; New York: Mouton de Gruyter.

Baayen, H., in Lieber, R. (1991): Productivity and English derivation: A corpus-based study. *Linguistics*, 29 (5): 801–843.

Bajec, A., Kolarič, R., in Rupel, M. (1956): *Slovenska slovnica*. Ljubljana: Državna založba Slovenije.

Bauer, L. (2001): *Morphological productivity*. Cambridge: Cambridge University Press.

Bauer, L. (2005): Productivity: Theories. V P. Štekauer in R. Lieber (ur.): *Handbook of word-formation*: 315–334. Dordrecht: Springer.

Biber, D., Conrad, S., in Reppen, R. (1998): *Corpus linguistics: Investigating language structure and use*. Cambridge: Cambridge University Press.

- Biber, D., Johansson, S., Leech, G., Conrad, S., in Finegan, E. (1999): *Longman grammar of spoken and written English*. Essex: Pearson Education.
- Booij, G. (2012): *The grammar of words: An introduction to linguistic morphology*. Oxford: Oxford University Press.
- Breznik, A. (1934): *Slovenska slovnica za srednje šole*. Celje: Družba sv. Mohorja.
- Center za jezikovne vire in tehnologije, Univerza v Ljubljani. Dostopno prek: <https://www.cjvt.si/> (22. junij 2018).
- Cvrček, V. (2010): *Mluvnice současné češtiny: 1, Jak se píše a jak se mluví*. Praga: Karolinum.
- Čermák, F. (1995/2005): Jezikovni korpus: sredstvo in vir spoznanj. V V. Gorjanc in S. Krek (ur.): *Študije o korpusnem jezikoslovju*: 137–171. Ljubljana: Krtina./Slovo a slovesnost 56. 119–140.
- Evert, S., in Lüdeling, A. (2001): Measuring morphological productivity: Is automatic preprocessing sufficient?. V P. Rayson idr. (ur.): *Proceedings of the Corpus Linguistics 2001 Conference*: 167–175. Lancaster: UCREL.
- Frauenfelder, U. H., in Schreuder, R. (1992): Constraining psycholinguistic models of morphological processing and representation: The role of productivity. V G. Booij in J. van Marle (ur.): *Yearbook of Morphology 1991*: 165–183. Dordrecht: Springer.
- Gložančev, A., Jakopin, P., Michelizza, M., Uršič, L., in Žele, A., ur. (2009): *Novejša slovenska leksika: v povezavi s spletnimi jezikovnimi viri*. Ljubljana: Založba ZRC, ZRC SAZU.
- Gorjanc, V. (2005): *Uvod v korpusno jezikoslovje*. Domžale: Založba Izolit.
- Górski, R. (2013): Polish derived expressive adjectives, a corpus-based study.

- V M. Polczyńska, Ł. Piotr Pakuła in D. Jaworska (ur.): *Young linguists' insights: Taking interdisciplinary approaches to the core*: 65–80. Varšava: Anglistyki UAM.
- Grčar, M., Krek, S., in Dobrovoljc, K. (2012): *Obeliks: statistični oblikoskladenjski označevalnik in lematizator za slovenski jezik*. V T. Erjavec in J. Žganec Gros (ur.): *Zbornik Osme konference Jezikovne tehnologije*: 89–94. Ljubljana: Institut "Jožef Stefan".
- Herrity, P. (2016): *Slovene: A comprehensive grammar*. London; New York: Routledge.
- Janežič, A., in Sket, J. (1900): *Slovenska slovnica*. Celovec: Družba sv. Mohorja.
- Jurgec, P. (2011): Slovenščina ima 9 samoglasnikov. *Slavistična revija*, 59 (3): 243–268.
- Kennedy, G. (1999): *An introduction to corpus linguistics*. London; New York: Longman.
- Kern, B. (2014): Zagadnienia słowotwórcze w wybranych podręcznikach do nauczania języków słoweńskiego i polskiego jako obcych. V E. Pałuszyńska (ur.): *Glottodydaktyka – media – komunikacja, Negocjowanie znaczeń*: 211–220. Łódź: Wydawnictwo Uniwersytetu Łódzkiego.
- Kern, B. (2017): *Stopenjsko besedotvorje*. Ljubljana: Založba ZRC, ZRC SAZU.
- Kilgarriff, A., Rychlý, P., Smrz, P., in Tugwell, D. (2004): *The Sketch Engine. Proceedings of the 11th EURALEX International Congress*: 105–116. Lorient: Universite de Bretagne-Sud.
- Leech, G. (1992): *Corpora and theories of linguistic performance*. V J. Svartvik (ur.): *Directions in corpus linguistics*: 105–122. Berlin: Mouton de Gruyter.

- Lieber, R. (1992): *Deconstructing morphology: Word formation in syntactic theory*. Chicago; London: The University of Chicago Press.
- Logar, N. (2005): Besedotvorni sklopi. *Slavistična revija*, 53 (2): 171–192.
- Logar, N. (2006): Stilno zaznamovane nove tvorjenke – tipologija. *Slavistična revija*, 54 (pos. številka): 87–101.
- Michelizza, M. (2008): Nove tvorjenke v spletnih besedilih (primer Wikipedije). V M. Košuta (ur.): *Slovenščina med kulturami*: 328–338. Celovec; Ljubljana: Slavistično društvo Slovenije.
- Nova slovnica sodobne standardne slovenščine: viri in metode*. Dostopno prek: <http://slovnica.ijs.si/> (22. junij 2018).
- Plag, I. (1999): *Morphological productivity: Structural constraints in English derivation*. Berlin; New York: Mouton de Gruyter.
- Säily, T. (2011): Variation in morphological productivity in the BNC: Sociolinguistic and methodological considerations. *Corpus linguistics and linguistic theory*, 7 (1): 119–141.
- Scherer, C. (2015): Change in productivity. V P. O. Müller idr. (ur.): *Word-formation: An international handbook of the languages of Europe: 1781–1793*. Berlin; Boston: De Gruyter Mouton.
- Sicherl, E., in Žele, A. (2018): Prekrivanke v slovenščini z vidika vpliva angleškega jezika. *Jezik in slovstvo*, 63 (1): 75–88.
- Stramljič Breznik, I. (2013): Besedotvorna produktivnost posnemovalnih medmetov za zvoke naprav v slovenščini. *Slavia Meridionalis*, 13: 97–107.
- Stramljič Breznik, I. (2016): Slovene. V P. O. Müller (ur.): *Word-formation: An international handbook of the languages of Europe: 2979–2998*. Berlin; Boston: De Gruyter Mouton.
- Stramljič Breznik, I., in Voršič, I. (2011): Word-formational productivity of the

- Slovene language in the case of sports neologisms. *Linguistica*, 51: 23–38.
- Štumberger, S. (2015): Besedotvorje novejšje slovenske leksike: medponskoobrazilne zloženke. V D. Zuljan Kumar in H. Dobrovoljc (ur.): *Zbornik prispevkov s simpozija 2013*: 155–163. Nova Gorica: Založba Univerze.
- Teubert, W. (1999/2005): Korpusno jezikoslovje in leksikografija. V V. Gorjanc in S. Krek (ur.): *Študije o korpusnem jezikoslovju*: 103–136. Ljubljana: Krtina./Korpuslinguistik und Lexikographie. Deutsche Sprache 4.
- Tognini - Bonelli, E. (1993): *Corpus linguistics at work*. Amsterdam, Philadelphia: John Benjamins.
- Toporišič, J. (2000): *Slovenska slovnica*. Maribor: Založba Obzorja.
- Van Marle, J. (1992): The relationship between morphological productivity and frequency: A comment on Baayen's performance-oriented conception of morphological productivity. V G. E. Booij in J. van Marle (ur.): *Yearbook of Morphology 1991*: 151–163. Dordrecht: Springer.
- Vidovič Muha, A. (1988): *Slovensko skladijsko besedotvorje ob primerih zloženek*. Ljubljana: Znanstveni inštitut Filozofske fakultete.
- Vidovič Muha, A. (2009): Skladijska interpretacija glagolskih predponskih obrazil – vprašanje propozicije. *Slavistična revija*, 57 (2): 251–261.
- Voršič, I. (2013): *Sistemska in nesistemska leksikalna tvorba v novejšem besedju slovenskega jezika: Doktorska disertacija*. Maribor.
- Žele, A. (2016): Aktualni jezikovni načini izražanja v slovenščini: sklapljanje kot naravni in aktualni odraz nepretrganosti govora v narečnem in knjižnem jeziku. *Annales*, 26 (4): 709–714.

NOUN DERIVATIVES FROM VERBS: CORPUS WORD-FORMATION PRODUCTIVITY AND GRAMMATICAL DESCRIPTION

This paper discusses the understanding of word-formation productivity, as was founded by Harald Baayen at the end of the 1980s and in the early 1990s. Baayen's method of calculating productivity P is based on corpus linguistics methodology, therefore by using data from the Kres corpus, a corpus of written Slovene, we tested the method on the following ten Slovene suffixes, which all form noun derivatives from verbs (gender: masculine, meaning: human agent): *-vec*, *-telj*, *-ež*, *-uh*, *-avh*, *-ar*, *-lec*, *-ač*, *-aj*, and *-ant*. Among other findings, one was indicative: the number of derivatives with a certain suffix, and the productivity P of that same suffix could be (quite) diverse. For example: the suffix that had the largest number of derivatives was *-lec*, while the suffix *-ež* had the highest P score. On the basis of the data from the corpus, we then presented a list of the ten aforementioned suffixes along with example derivatives, i.e. the list, that could be part of a new grammar of Slovene. We also briefly commented on it and then, in the final discussion, pointed out some advantages and disadvantages of the approach presented.

Keywords: word-formation, productivity, H. Baayen, suffix, derivative, grammar, corpus

To delo je ponujeno pod licenco Creative Commons: Priznanje avtorstva-Deljenje pod enakimi pogoji 4.0 Mednarodna.

This work is licensed under the Creative Commons Attribution-ShareAlike 4.0 International.

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

