

## VPLIV BESEDILNIH TIPOV NA POJAVLJANJE NOMINALIZACIJE V SLOVENŠČINI: KORPUSNA RAZISKAVA

Tamara Mikolič Južnič  
Filozofska fakulteta, Ljubljana

UDK 811.163.6'42'322'367.622'367.625/.45

V prispevku skušamo opredeliti nominalizacijo in njeno pogostnost v različnih slovenskih besedilnih tipih. Pojavljanje nominalizacije, ki je definirana v okviru sistemsko-funkcijske slovnice kot vrsta slovnične metafore, opazujemo v namensko zgrajenem korpusu različnih besedilnih tipov (literarnih, znanstvenih, strokovnih), podatke pa primerjamo s splošnim jezikom, kot ga izkazuje korpus JOS 100k.

nominalizacija, besedilne vrste, korpusna analiza, slovenščina

The paper presents an overview of nominalisation and its frequency in certain Slovene text types. The presence of nominalisation, which is defined within systemic functional grammar as a type of grammatical metaphor, is observed in a specifically compiled corpus of various text types (literary, academic, technical) and is compared with the standard language as shown in the JOS 100k corpus.

nominalisation, text types, corpus analysis, Slovene

### 1 Uvod

Kontrastivnim slovničarjem in prevodolovcem, ki se ukvarjajo s slovenščino, se pri delu pogosto dogaja, da pri obravnavanju določenih značilnosti na stiku dveh ali več jezikov potrebujejo raziskovalno in bibliografsko podporo svojih ugotovitev. Take podpore navadno ni težko najti za tuje jezike, kot so angleščina, italijanščina, španščina ali nemščina, medtem ko je stanje za slovenščino pogosto precej problematično: interesi slovenistov so se v zadnjih letih (desetletjih?) osredotočili na številna zanimiva in perspektivna področja, vendar med njimi praktično ni nikogar, ki bi se ukvarjal z določenimi slovničnimi vprašanji, zlasti ob korpusni podpori. Tako na primer o nominalizaciji (oz. posamostaljenju) v slovenskem bibliografskem sistemu COBISS najdemo natanko 35 bibliografskih enot, od katerih jih je za slovenščino znanstveno relevantnih manj kot deset.

Tako stanje predstavlja velik izziv, saj od raziskovalcev zahteva, da se v posamezno področje poglobijo tudi na slovenski strani. V tem duhu pričujoči prispevek skuša osvetliti nekatere razlike, ki se pojavljajo v pogostnosti nominalizacije v različnih besedilnih tipih. Specifično se primerja stanje v korpusu JOS 100k (Erjavec idr. 2010) s stanjem v namensko sestavljenem korpusu z različnih področij. Izhaja se iz hipoteze, da se nominalizacija pogosteje pojavlja v žanrih, kjer se je zgodovinsko pojavila prej (znanstveni diskurz, strokovni jezik), manj pogosto pa v ostalih žanrih, v nespecializiranem, splošnem jeziku in literaturi. Ocena pogostnosti nominalizacij temelji na pojavljanju nominalizacij z najpogosteje uporabljenimi obrazili (Mikolič Južnič 2007). Odgovor na zastavljeno vprašanje bo tako lahko ponudil boljše razumevanje dogajanja v prevodih iz različnih jezikov v slovenščino, kot tudi iz slovenščine v različne tuje jezike.

## 2 Kaj je nominalizacija

Nominalizacija, kot jo razumemo v tem prispevku, je definirana v okviru modela sistemsko-funkcijske slovnice, v skladu s katerim je to vrsta predstavne slovnice metafore (ideational grammatical metaphor, Halliday, Matthiessen 2004), pri kateri glagolski dogodek (process), ki ga skladno (congruently) ubesedimo z glagolom, metaforično – tj. s slovnico metaforo – izrazimo s samostalnikom, pri čemer spremenimo seveda tudi vlogo ostalih elementov, vezanih na sam glagolski dogodek. Na primer, namesto da se izrazimo *Presenetilo me je, da je odšel*, uporabimo formulacijo *Presenetil me je njegov odhod*, kjer je beseda *odhod* nominalizacija, vršilec dejanja pa je izražen s svojilnim pridevnikom *njegov*. Tak način izražanja naj bi se prvič začel dosledno pojavljati v znanstvenih spisih Galiela Galilei in njegovih sodobnikov ter se od tam razširil tako na druga področja človeškega znanja kot tudi na druge evropske jezike (Halliday, Matthiessen 1999).

Še dandanes se pogostnost nominalizacije zelo razlikuje tako od jezika do jezika (npr. Mikolič Južnič 2007 za primerjavo slovenščine in italijanščine in Plemenitaš 2004 za primerjavo angleščine in slovenščine), pa tudi glede na besedilne tipe, s katerimi imamo opravka (Halliday, Martin 1993). Za slovenščino velja prepričanje, da je bolj glagolski jezik in da ne prenese pogoste uporabe nominalizacije (Klinar 1996; Žele 1996), čeprav obstajajo tako zgodovinska (Vodušek 1933) kot sodobnejša nasprotna mnenja (Plemenitaš 2004).

### 2.1 Nominalizacija v slovenščini

Vprašanje nominalizacije je bilo deležno omejene pozornosti v našem prostoru s stališča slovenščine kot take in tudi v primerjavi z različnimi drugimi jeziki. Razprave se navadno osredotočajo na opazovanje nominalizacije v jeziku oz. na njeno dejansko prisotnost v besedilih ter tudi na njeno upravičenost, potrebnost, odvisnost od zunanjih

dejavnikov (transfera iz drugih jezikov) in podobno.

Med študijami, ki se ukvarjajo le s slovenskim aspektom, omenimo najprej zanimiv pogled Boža Voduška (1933), ki je opazoval konflikt med dejanskim stanjem in predpisanimi pravili v zvezi z nominalizacijo: trdil je, da je zmotno pripisovati slovenščini glagolsko naravo (tako naravo torej, ki je bolj ekspresivna z glagoli kot s samostalniki) v nasprotju z »razvitejšimi«<sup>1</sup> zahodnoevropskimi družbami in njihovimi jeziki, ki naj bi bili pretežno samostalniški. Po Voduškovem mnenju gre za jezikovne razlike, pogojene s sociološkimi in vsebinskimi oz. žanrskimi dejavniki: »Zaradi abstraktnosti predmeta je na primer stil vsake filozofije izrazito nominalen, dočim je stil literature v primeri z njim pretežno bolj verbalen«<sup>2</sup> (Vodušek 1933: 72). Kljub posplošenosti številnih Voduškovih trditev je zanimiva izrazita podobnost nekaterih njegovih tez (npr. razumevanja razvoja jezikov od verbalnih k nominalnim) z mnogo poznejšimi in z obsežnimi raziskavami podkrepljenimi ugotovitvami sistemsko-funkcijskih slovnicearjev s Hallidayem na čelu (Halliday, Martin 1993).

Andreja Žele (1996) je raziskovala pojavljanje nominalizacije v slovenskem publicističnem jeziku med leti 1946 in 1995. Avtorica opaža porast pojavljanja nominalizacij v obravnavanem besedilnem tipu skozi desetletja, ki jo zlasti v zvezi s sedemdesetimi in osemdesetimi leti komentira kot »čezmerno kopičenje izglagolskih tvorjenk, ki zabrišejo jasnost sporočila in vodijo celo v napačno razumevanje in razlage.«<sup>3</sup>

Stanko Klinar (1996) omenja večjo glagolskost slovenščine v primerjavi s samostalniškostjo angleščine, po drugi strani pa Katja Plemenitaš (2004) pri primerjavi časopisnih vesti in kritik v angleščini in slovenščini ugotavlja, da v zvezi z rabo nominalizacije med angleščino in slovenščino ne prihaja do večjih razlik ali pa je njena pogostnost celo večja v slovenščini.

Na tem mestu želimo preveriti prisotnost nominalizacije v nekaterih slovenskih tipih besedil z relativnega, medsebojno primerjalnega stališča, zlasti glede na standardni, nespecializirani jezik.

### 3 Korpus in metoda

Raziskava je bila opravljena na dveh različnih korpusih: korpusu JOS 100k (Erjavec idr. 2010), ki zajema 100.000 pojavnic in temelji na vzorcih iz referenčnega korpusa FidaPLUS ter predstavlja splošni slovenski jezik (uporabljen je bil kot kontrolni vzorec),

in namensko zbranim specializiranim korpusu različnih področij in besedilnih tipov, ki skupno šteje približno milijon pojavnic. Viri besedil in področja, ki so bila izbrana za ta specializirani korpus, so navedeni v Tabeli 1.

Kot je razvidno iz Tabele 1, je specializirani korpus razdeljen na tri vrste besedil: znanstvena, strokovna in umetnostna besedila. Znanstvena besedila, ki zajemajo tako članke s humanističnih kot naravoslovnih področij, so bila izbrana na podlagi reprezentativnosti za slovenski prostor in proste dostopnosti na spletu. V tabeli so specificirane

Tabela 1: Sestava specializiranega korpusa in število pojavnic

Vrsta besedila	Izbrana področja (in viri)	Število pojavnic
Znanstvena besedila	Humanistika	415.000
	<i>Slavistična revija</i>	81.399
	<i>Sodobna pedagogika</i>	74.355
	<i>Revija za kriminalistiko in kriminologijo</i>	109.788
	<i>Geografski vestnik</i>	53.611
	<i>Etnolog</i>	95.847
	Naravoslovje	304.744
	<i>Zobozdravstveni vestnik</i>	55.456
	<i>Medicinski razgledi</i>	72.178
	<i>Acta biologica slovenica</i>	52.780
	<i>Geodetski vestnik</i>	55.364
	<i>Farmacevtski vestnik</i>	68.966
Strokovna besedila	Slovenska zakonodaja	154.088
	Stanovanjski zakon SZ-1	
	Zakon o društvih ZDRu-1	
	Zakon o verski svobodi ZVS	
	Zakon o volilni in referendumski kampanji ZVRK	
	Zakon o javnih zbiranjih ZJZ	
	Zakon o splošnem upravnem postopku ZUP	
	Zakon o upravnih taksah ZUT	
	Kazenski zakonik KZ-1	
Umetnostna besedila	Literatura	160.543
	Ivan Sivec, <i>Triglavski kralj: Črtice iz življenja župnika Jakoba Aljaža</i>	
	Marijan Pušavec, <i>Zbiralec Nasmehov</i>	
	Marko Simčič, <i>Jantarina</i>	
	Mart Lenardič, <i>Iz mojega življenja</i>	
	<i>Kavč</i>	
	<i>Kmetavzar</i>	
	<i>Moja ženska</i>	
	Miha Remec, <i>Pomnik za Evridiko</i>	
	<i>Žar ptica</i>	
	Sebastijan Pregelj, <i>Pianino</i>	
	<i>Smrt Lužarjeve Zale</i>	
SKUPAJ POJAVNIC		1.034.375

tudi revije, iz katerih so bili izbrani članki. Število pojavnic v različnih podkorpusih se precej razlikuje, kar je posledica zelo različne dolžine znanstvenih člankov v teh revijah (izbranih je bilo po deset člankov na revijo, razen v primerih, ko so bili članki izjemno kratki in jih je bilo zato vključenih do 20, npr. pri zobozdravstvu in biologiji). Strokovni podkorpus je v celoti sestavljen iz besedil s področja prava, in sicer osnovnih besedil zakonov, navedenih v Tabeli 1.<sup>1</sup> Umetnostni podkorpus pa združuje nekaj črtic in kratkih zgodb sodobnih avtorjev, ki so prosto dostopne na spletu.

Izbrana besedila so bila nato označena s pomočjo spletnega servisa projekta JOS (Erjavec idr. 2010), opremljena tako z lemmami kot oblikoskladenjskimi oznakami, da so bila nato primerno pripravljena za pregledovanje s programom Sketch Engine (Krek, Kilgariff 2006), s pomočjo katerega so bile opravljene avtomatske analize.

Ocena relativne pogostnosti nominalizacije je bila izvedena na osnovi pogostnosti deverbálnih nominalizacijskih priponskih obrazil, ki se najpogosteje pojavljajo v *Slovarju slovenskega knjižnega jezika* (SSKJ; Mikolič Južnič 2007): preko 90 % vseh nominalizacij, ki jih navaja SSKJ, je tvorjenih s sedmimi obrazili, in sicer (po vrstnem redu pogostnosti) -anje, -enje,<sup>2</sup> -tev, -cija, -tje, -ava in -ek. Čeprav se pojavljajo še številna druga obrazila, pa tudi take nominalizacije, ki so

tvorjene brez priponskega obrazila (kot npr. v besedah *dvom, izbruh, izbris* itn.), vseeno glede na pogostnost zgoraj omenjenih obrazil menimo, da je ocena pogostnosti nominalizacije dovolj natančna in rezultati izkazujejo oceno realnega stanja.

S pomočjo programa Sketch Engine in iskanja lem so bili torej izluščeni vsi primeri, v katerih se pojavljajo navedena obrazila. Ročno so bile izločene nenominalizacijske pojavnice, nato pa so bili zaradi primerljivosti velikosti korpusov rezultati normalizirani na 100.000 pojavnic. V nadaljevanju so predstavljeni rezultati te analize skupaj s komentarjem.

#### 4 Rezultati in razprava

Analiza pojavljanja zgoraj naštetih nominalnih priponskih obrazil v omenjenih korpusih je dala zelo zanimive rezultate, ki so v absolutnih številkah strnjeni v Tabeli 2. Ker je velikost posameznih podkorpusov različna in ker smo rezultate želeli primerjati s korpusom JOS 100k, smo rezultate normalizirali na 100.000 pojavnic. Primerljivi rezultati so torej navedeni v Tabeli 3. Zatem so v Tabeli 4 v odstotkih prikazana izražena odstopanja posameznih podkorpusov od kontrolnega korpusa JOS 100k.

Preden se posvetimo interpretaciji zgornjih podatkov, naj poudarimo, da na tem mestu ne želimo govoriti o pogostnosti nominalizacije v slovenščini nasploh oz. o vprašanju,

Tabela 2: Absolutno število analiziranih nominalizacij v izbranih besedilih

	-NJE	-TEV	-CIJA	-TJE	-AVA	-EK	skupaj
Humanistična besedila	8.498	1.402	1.608	196	1.158	1.099	13.961
Naravoslovna besedila	6.594	1.528	1.805	195	1.288	1.373	12.783
Literarna besedila	500	88	28	59	49	96	820
Zakonska besedila	3.520	1.316	194	42	367	1.018	6.457
JOS 100k	1.278	338	49	242	297	225	2.429

<sup>1</sup> V humanističnem podkorpusu je del besedil prav tako s področja prava, vendar gre za drugačen tip besedil – znanstvene članke.

<sup>2</sup> Pri analizi in podajanju rezultatov sta bili priponi -anje in -enje združeni v splošnejšo obliko -nje.

Tabela 3: Pogostnost nominalizacije v izbranih besedilih na 100.000 besed

	-NJE	-TEV	-CIJA	-TJE	-AVA	-EK	skupaj
Humanistična besedila	2.048	368	387	47	279	265	3.364
Naravoslovna besedila	2.163	501	592	64	423	451	4.194
Literarna besedila	311	55	17	37	31	60	511
Zakonska besedila	2.695	1.008	149	32	281	779	4.944
JOS 100k	1.278	338	225	49	242	297	2.429

Tabela 4: Odstopanja od pogostnosti kontrolnih vrednosti v korpusu JOS 100k glede na vrsto besedila in posamezno obrazilo (izražena v odstotkih; podatki za JOS predstavljajo 100 %)

	-NJE	-TEV	-CIJA	-TJE	-AVA	-EK	skupaj
Humanistična besedila	160,3	108,9	172,0	95,9	115,3	89,2	138,5
Naravoslovna besedila	169,2	148,2	263,1	130,6	174,8	151,9	172,7
Literarna besedila	24,3	16,3	7,6	75,5	12,8	20,2	21,0
Zakonska besedila	210,9	298,2	66,2	65,3	116,1	262,3	203,5

ali je slovenščina dejansko bolj glagolsko orientiran jezik ali ne, saj znotraj jezika samega nimamo kriterija za oceno take splošne pogostnosti, marveč bi o tem bilo mogoče razpravljali zgolj v okviru primerjave z drugimi jeziki, kar je onkraj domene tega prispevka.

Kar zgornji tabeli jasno nakazujeta, je, da prihaja med različnimi tipi besedil znotraj slovenščine do bistvenih razlik. Če korpus JOS 100k upoštevamo kot merodajen z ozirom na dejstvo, da gre za uravnotežen vzorčni korpus slovenskega jezika (Erjavec idr. 2010), lahko predvidevamo, da odraža splošno pogostnost nominalizacije v pisnem slovenskem jeziku. Vrednosti, navedene za ta korpus, torej upoštevamo kot kontrolne vrednosti (tj. vrednosti, veljavne za splošni jezik), in sicer tako za skupno število nominalizacij kot za primere s posameznimi priponami. Izhajajoč iz te osnove, opazimo bistveni odklon med umetnostnimi besedili na eni ter znanstvenimi in strokovnimi na drugi strani.

Pri umetnostnih besedilih se analizirana obrazila pojavljajo skoraj petkrat manj pogosto kot v kontrolnem korpusu. Če se ozremo na posamezna obrazila, opazimo, da do največje razlike prihaja pri priponi -cija (npr.

*interpretacija*), pri kateri je razmerje med pogostnostjo v literarnem podkorpusu in korpusu JOS 100k 1 : 13, medtem ko je pogostnost pripone -tje (npr. *vpitje*) najbolj podobna frekvenci v korpusu JOS: razmerje je 1 : 1,3.

Na drugi strani se med seboj precej razlikujeta tudi oba tipa znanstvenih besedil: nominalizacija je v obeh primerih bistveno pogostejša kot v kontrolnem korpusu, vendar se v naravoslovnih besedilih pojavlja skoraj dvakrat pogosteje kot v humanističnih: v primerjavi s kontrolnim korpusom je nominalizacij skoraj 40 % več v humanističnih besedilih in preko 70 % več v naravoslovnih besedilih. Razlike se pojavljajo tudi pri pogostnosti posameznih obrazil. V nasprotju z umetnostnim podkorpusom je pripona -cija pri obeh tipih znanstvenih besedil najpogostejša: v znanstvenih besedilih se pojavlja kar preko 160 % več nominalizacij kot v korpusu JOS 100k, v humanističnih besedilih pa preko 70 % več. Medtem ko se v humanističnem korpusu priponi -tje in -ek (npr. *zaključek*) pojavljata celo manj pogosto kot v kontrolnem korpusu (čeprav le nekaj odstotkov manj), se v naravoslovnih besedilih vse pripone pojavljajo pogosteje, najmanjša razlika pa je razvidna pri priponi -tje (30 % več nominalizacij v naravoslovnih besedilih).

Strokovni podkorpus zakonskih besedil prav tako izkazuje večjo pogostnost nominalizacij v primerjavi s kontrolnim korpusom, celo preko 100 % več, toda razlike med pojavljanjem posameznih obrazil so še bolj izrazite. Priponi -cija in -tje sta relativno redki, pojavita se približno 35 % manjkrat kot v korpusu JOS 100k. Po drugi strani je pripona -tev (npr. *vzpostavitev*) izjemno pogosta (pojavi se skoraj 200 % večkrat kot v kontrolnem korpusu), zelo veliko pa je tudi primerov s priponama -ek in -nje (npr. *zbiranje*, *zdravljenje*).

Vzroke za take razlike pri rabi posameznih pripon lahko verjetno pripišemo različnim dejavnikom. V znanstvenih besedilih s področja naravoslovja je npr. zelo pogosta uporaba tujih besednih korenov, ob katerih se najpogosteje najde pripona -cija, ki ima ravno tako tuj izvor; take nominalizacije so npr. *vazodilatacija*, *stabilizacija* itn. Po drugi strani je v strokovnih besedilih (zakonih) izredno pogosta pripona -tev, značilna za dovršne glagole, ki izražajo zaključena dejanja (npr. *vzpostavitev*, *uveljavitev* itn.).

Glede na rezultate raziskave lahko torej sklepamo, da se nominalizacija v primerjavi s splošnim jezikom pojavlja v znanstvenih in strokovnih besedilih veliko pogosteje, še bolj izrazito pa v primerjavi z literarnimi besedili. Najpogosteje se v izbranih besedilih pojavlja v podkorpusu zakonov, nekoliko manjkrat pa v znanstvenih besedilih s področja naravoslovnih ved ter še manj pogosto v besedilih s področja humanistike.

## 5 Sklep

V prispevku smo skušali osvetliti razlike v pojavljanju nominalizacije v različnih besedilnih tipih v slovenskem jeziku. S pomočjo nominalizacijskih priponskih obrazil, ki se najpogosteje pojavljajo v slovenščini, in namensko sestavljenega korpusa znanstvenih, strokovnih in literarnih besedil ter kontrolnega korpusa JOS 100k smo primerjali pogostnost nominalizacije v omenjenih besedilnih tipih ter ugotovili, da se dejansko

pojavljajo velike razlike v pogostnosti nominalizacije: v primerjavi z besedili v kontrolnem korpusu se v literarnih besedilih nominalizacija pojavlja izrazito redko, v znanstvenih in strokovnih besedilih pa veliko pogosteje. Zlasti izstopajo strokovna (pravna) besedila, kjer je nominalizacij preko 100 % več kot v vzorcu splošnega jezika.

Ob tem se odpirajo številna vprašanja. Po eni strani je gotovo še veliko besedilnih tipov, ki na tem mestu niso bili obravnavani in bi se morda s stališča nominalizacije izkazali kot zanimivi. Po drugi strani nas v prihodnosti prav tako čaka še kaka primerjava pogostnosti nominalizacije v slovenskem jeziku in drugih, tujih jezikih, ki bo morda potrdila ali ovrgla velikokrat izrečeno prepričanje, da je slovenščina dejansko bolj glagolski jezik.

## Literatura

- BAJEC, Anton idr., 2000: *Slovar slovenskega knjižnega jezika*. Ljubljana: DZS.
- ERJAVEC, Tomaž, FIŠER, Darja, KREK, Simon, LEDINEK, Nina, 2010: The JOS Linguistically Tagged Corpus of Slovene. *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. Malta. <http://www.lrec-conf.org/lrec2010/>
- HALLIDAY, Michael A. K., MARTIN, James R., 1993: *Writing Science. Literacy and Discursive Power*. Pittsburgh: University of Pittsburgh Press.
- HALLIDAY, Michael A. K., MATTHIESSEN, Christian M. I. M., 1999: *Construing Experience through Meaning. A Language-Based Approach to Cognition*. London, New York: Continuum.
- HALLIDAY, Michael A. K., MATTHIESSEN, Christopher M. I. M., 2004: *An Introduction to Functional Grammar*. London: Arnold.
- KLINAR, Stanko, 1996: Samostalniškost angleščine v primeri s slovenščino. Stanko Klinar (ur.): *K tehniki prevajanja iz slovenščine v angleščino*. Radovljica: Didakta. 149–193.

- KREK, Simon, KILGARRIFF, Adam, 2006: Slovene Word Sketches. Tomaž Erjavec, Jerneja Žganec Gros (ur.): *Zbornik 5. slovenske in 1. mednarodne konference Jezikovne tehnologije 2006*. Ljubljana: Institut Jožef Štefan. <http://nl.ijs.si/is-ltc06/proc/>
- JOS 100k: <http://nl.ijs.si/jos/>
- MIKOLIČ JUŽNIČ, Tamara, 2007: *Nominalne strukture v italijanščini in slovenščini: pogostnost, tipi in prevodne ustreznice*. Doktorska disertacija. Ljubljana: Filozofska fakulteta.
- PLEMENITAŠ, Katja, 2004: *Posamostaljenja v angleščini in slovenščini na primeru dveh besedilnih vrst*. Doktorska disertacija. Ljubljana: Filozofska fakulteta.
- VODUŠEK, Božo, 1933: Za preureditev nazora o jeziku. *Krog. Zbornik umetnosti in razprav*. Ljubljana. 66–76.
- ŽELE, Andreja, 1996: Razvoj posamostaljenja v slovenskem publicističnem jeziku med 1946 in 1995. Ada Vidovič Muha (ur.): *Jezik in čas*. Ljubljana: Znanstveni inštitut Filozofske fakultete. 191–200.