

UDK 811.163.6'367.622

Tamara Mikolič Južnič

Filozofska fakulteta Univerze v Ljubljani

tamara.mikolic@guest.arnes.si

PRIKAZ NOMINALIZACIJE V RAZPOLOŽLJIVIH KORPUSIH SLOVENŠČINE

Prispevek je poskus širšega pregleda pojavljanja nominalizacije v slovenskem jeziku na podlagi korpusno osnovane primerjave izvornih slovenskih besedil različnih žanrov in prevodov iz izbranih tujih jezikov (italijanščine, angleščine, nemščine, francoščine). Izkaže se, da je nominalizacija izrazito žanrsko in registrsko pogojen fenomen, ki svoje mesto najde v praktično vseh vrstah diskurza.

Ključne besede: nominalizacija, slovenščina, sistemsko funkcijska slovnica, slovnična metafora, korpusna analiza

The article is an attempt to provide an overview of the occurrence of nominalization in the Slovene language. A corpus-based methodology is used to compare original Slovene texts of different genres and translations from selected foreign languages (Italian, English, German, French). It appears that nominalization is highly dependent on register, but present in virtually all types of discourse.

Keywords: nominalization, Slovene, systemic functional grammar, grammatical metaphor, corpus analysis

0 Uvod

Nominalizacija velja za pogosto značilnost (zlasti pisnega) izražanja odraslih (prim. Halliday in Matthiessen 1999, 2004; Halliday in Martin 1993) in je bila v preteklosti predmet številnih, raznolikih raziskav v sklopu posameznih jezikov (npr. Chomsky 1968, Bateman 1990, Colombi 2006, Stålhammar 2006), prav tako pa se je povečal interes s kontrastivnega in prevajalskega vidika (npr. Alexiadou 2001, Lapata 2002, Nordrum 2007, Gradečak-Erdeljić 2009), saj se njeno pojavljanje v skladu z navedenimi raziskavami v nekaterih primerih izrazito razlikuje tako od jezika do jezika kot tudi znotraj posameznega jezika z vidika jezikovne zvrstnosti in besedilnih tipov. Tako se nominalizacija postavlja kot potencialno problematičen element vsaj z dveh vidikov: po eni strani se postavlja vprašanje, ki je bilo večkrat izpostavljeno v literaturi (prim. Halliday in Matthiessen 2004: 636), kateri so tisti registri in besedilni tipi, kjer je prisotna v večji ali manjši meri. Tudi v slovenskem prostoru je nastalo nekaj študij, ki so osredotočene bodisi na posamezne jezikovne ali besedilne vrste (npr. Žele 1996, 1997; Mikolič Južnič 2012a) bodisi na primerjave z drugimi jeziki (Plemenitaš 2005, 2007; Mikolič Južnič 2007, 2012b, Košak 2007, Kuster 2014, Zajc 2014), pa

tudi na druge aspekte, povezane z rabo nominalizacije in t. i. samostalniškimi slogom (prim. Gnamuš 1979) ter z razmerjem med slovnično metaforo in leksikalno metaforo (Snoj 2009). Po drugi strani pri prevajalskem procesu, zlasti v primeru, ko prihaja do bistvenih razlik v pogostnosti rabe v izvornem oz. ciljnim jeziku, na kar nas bi lahko opozorile kontrastivne študije.

Pričujoči članek je na korpusni metodologiji osnovan poskus prikaza, kakšno je dejansko pojavljanje nominalizacije: s kvantitativno metodo želimo pokazati, v kolikšni meri je v skladu z opisi, ki jih v literaturi zasledimo za druge jezike, namreč da je nominalizacija značilna zlasti za besedila s področja znanosti, uradniškega sveta in prava (Halliday in Matthiessen 2004: 636) ter da praktično ni besedilnih tipov (tvorjenih za odrasle govorce), kjer bi bila nominalizacija popolnoma odsotna.

1 Nominalizacija kot slovnična metafora

1.1 Definicija

Nominalizacijo v pričujoči raziskavi razumemo v skladu s sistemsko funkcijskim jezikoslovnim modelom: v skladu s tem modelom je jezik sistem pomenov in določenih oblik, s katerimi lahko te pomene udejanji (Halliday 1994: xiv). Z vidika funkcij naravnega jezika vsaka izjava združuje tri različne ravni pomena: medosebno, besedilno in predstavnostno metafunkcijo (prim. Halliday 1994, Halliday in Matthiessen 1999, 2004). Pri nominalizaciji, ki je predmet te raziskave, gre primarno za premik na predstavnostni ravnini, čeprav velja spomniti, da sprememba na ravni ene metafunkcije nujno prinese tudi spremembe na drugih dveh ravneh.¹

Predstavnostna metafunkcija je povezana s konceptom stavka kot predstave in ubeseditvijo prehodnostnih struktur, ki izražajo predstavnostni pomen, kar načeloma vključuje glagolski dogodek in njegove udeležence (in okoliščine). Eden od ključnih konceptov sistemsko funkcijske slovnice, kot poudarja Ravelli (1988: 133), je pomenski potencial: veliko je elementov, ki lahko nosijo pomen, in govorce/pisce lahko izbiramo med temi pomenskimi potenciali, ko želimo komunicirati. Odnos med pomenskim potencialom in izbrano ubeseditvijo ni naključen: za vsak pomen, ki ga želimo izraziti, imamo na voljo naravno možnost, z izbiro katere pridemo do 'skladnih' ubeseditiv.² Halliday (1994: 343) pravi, da govorce jezika določene vzorce ubeseditve prepoznavajo kot tipične, in te poimenuje 'skladne'. Vendar pa jezik, nadaljuje avtor (prav tam), ponuja tudi možnost alternativnih ubeseditiv, kjer niso uporabljeni tipični vzorci, in v teh primerih govorimo o slovnični metafori.

¹ Na besedilni ravni se najpomembnejša sprememba odraža na tematski strukturi stavka in njegovi argumentacijski moči (prim. Halliday in Matthiessen 1999: 239), na medosebni ravni pa gre za ubeseditve predloga (*proposition*) ali ponudbe (*proposal*) kot udeleženca s posledično spremembo informacijske strukture (prav tam: 241–42).

² Glede uporabljene terminologije s področja sistemsko funkcijske slovnice prim. angleško-slovenski dvosmerni glosar v Mikolič Južnič (2007: 232–39).

Konkretno pri nominalizaciji gre za vrsto predstavnostne slovnične metafore, pri kateri glagolski dogodek,³ ki ga sicer skladno izrazimo z glagolom, z različnimi priponskimi (obrazilnimi) morfemi⁴ in drugimi sredstvi metaforično ubesedimo s samostalnikom. Na primer skladno ubeseditvev *Presenetilo me je, da je odšel* lahko s slovnično metaforo oz. nominalizacijo ubesedimo kot *Presenetil me je njegov odhod* ali celo, če oba dela povedi nominaliziramo in vključimo v novo izjavo, *Zaradi presenečenja nad njegovim odhodom sem onemela*.⁵ V navedenem primeru nazorno vidimo, kako nominalizacija povzroči spremembe v konfiguraciji udeležencev glagolskega dogodka (vršilec iz odvisnega stavka je v tem primeru prekategoriziran v svojilno določilo), kar tvorcu besedila dopušča, da kompleksne sekvence besedila 'stisne' tako, da tvorijo element v drugi semantični konfiguraciji (Halliday in Martin 1993: 15). Tudi sicer skladna in metaforična ubeseditvev nista enakovredni ali izmenljivi, saj nimata popolnoma identičnega pomena in se razlikujeta z vidika vseh treh metafunkcij, besedilne, predstavne in medosebne, kot je bilo omenjeno zgoraj.

Nominalizacija je torej tako ubeseditvev glagolskega dogodka v obliki samostalnika kot tudi primer take ubeseditve (npr. *odhod* zgoraj). Najpogosteje je ta samostalnik izpeljan iz glagola, s katerim bi izrazili glagolski dogodek v skladnem stavku (npr. *odhod* < *oditi*). Halliday in Matthiessen (2004: 636) ugotavljata, da je nominalizacija glagolskega dogodka bolj sofisticirana ubeseditvev kot skladna različica in ni del vsakdanjega spontanega govora. Tipična je za besedila iz področja izobraževanja, znanosti, birokracije in prava. Avtorja (prav tam) prav tako navajata, da je moč zaslediti velike razlike med različnimi registri glede tipov in pogostnosti nominalizacije, vendar v nobenem registru ni popolnoma odsotna (z izjemo otroških knjig in pesmic). V pričujočem prispevku želimo preveriti, ali take trditve veljajo tudi za slovenski jezik oziroma v kolikšni meri je raba nominalizacije v slovenščini pogojena z registrom oz. besedilnim tipom.

1.2 Kratek oris zgodovinskega razvoja nominalizacije

Nominalizacija je po besedah Hallidaya in Matthiessna (1999: 242–44) prisotna v vseh naravnih jezikih v kaki obliki: gre za naravni proces, ki podobno kot leksikalna metafora bogati potencial pomena. Spremenjene zgodovinske okoliščine (pred-

³ Nominalizacija lahko zadeva tudi druge vrste elementov, ne samo glagolske dogodke (npr. lastnosti, okoliščine, povezovalci so lahko metaforično ubesedeni na različne načine; prim. Mikolič Južnič 2007: 54).

⁴ V zvezi z besedotvorno terminologijo prim. podrobni prikaz Vidovič Muha (1988, 2011, 2013). Velja pripomniti, da se pričujoči prispevek besedotvorja dotika le obrobno in posredno: nominalizacijske besedotvorne pripone so uporabljene kot sredstvo identifikacije nominalizacij; uporabljeni vzorci so lahko posplošeni in prilagojeni možnostim uporabe iskalnih pogojev v obravnavanih korpusih (prim razdelek 3.2).

⁵ Na tak način lahko še naprej gradimo niz nominalizacij, povezanih v nove samostalniške besedne zveze, ki jih lahko uporabimo v različnih funkcijah in jim lahko dodajamo različne elemente kot kateremu koli drugemu samostalniku, npr. *Svojo onemelost zaradi presenečenja nad njegovim odhodom sem skušala zakriti s smehom*. A tovrstno nizanje bistveno zvišuje leksikalno gostoto povedi in besedila in nosi za sabo potrebo po večjem kognitivnem naporu s strani poslušalca/bralca.

vsem razvoj znanosti) so povzročile, da je pojav v številnih jezikih postal dominantna lastnost semantičnega sistema, toda zametke slovnične metafore lahko prepoznavamo že v najzgodnejšem tvorjenju znanstvenih terminov pri starogrških mislecih. Znanstveniki pa so izkoristili še druge pomembne možnosti slovnice, in sicer možnost razširjanja nominalne skupine s predložnimi frazami in vrinjenimi stavki ter ponovljivost, oz. možnost, da lahko nominalizacije poljubno kopičimo v stavkih in besedilu. Iz stare grščine se je ta potencial, kot pravita Halliday in Martin (1993: 12–13), po principu jezikovnega kalka prenesel v latinščino in arabščino, iz srednjeveške in poznejše latinščine pa je prešel tudi v številne novodobne evropske jezike, kot so angleščina, španščina, nemščina, italijanščina, grščina, češčina, norveščina itd.

Razvoj jezikov in njihovih izraznih sredstev ni bil povsod enak. Čeprav je bil v času, ko se je v italijanščini in v angleščini, na primer, vštric z razvojem moderne znanosti uveljavljal z nominalizacijami bogat slog, značilen za Galileja, Newtona ter njune sodobnike (prim. Mikolič Južnič 2007: 67–69), in sta ta dva jezika imela za sabo že večstoletno leposlovno tradicijo, slovenski pisni jezik praktično še v povojih in sta prvi slovenski knjigi, Trubarjeva *Abečedarij* in *Katekizem*, izšli leta 1550, le 50 let pred objavo prvih pomembnejših Galilejevih del, se je slovenščina v naslednjih stoletjih skokovito razvila in z različnih vidikov dohitela omenjena jezika. Prisotnost slovenskega jezika v pisnih besedilih je bila sicer v 17. stoletju v primerjavi z zgoraj omenjenima jezikoma skromnejša in znanstveno pisanje se je na Slovenskem uveljavilo pozneje (v zvezi s slovenskim znanstvenim jezikom prim. Vidovič Muha 1986 ter druge prispevke v Vidovič Muha (ur.) 1986 in Pisanski Peterlin 2005), zato ni presenetljivo, da se je razvoj slovnične metafore v naši prozi začel nekoliko pozneje, vendar pa je bila nominalizacija po mnenju Voduška (1933) že značilnost nekaterih strokovnih jezikov.⁶ Jasno pa je tudi, da tako kot je angleščina verjetno začela uporabljati »nominalni slog« pod vplivom italijanskih besedil (prim. Halliday in Matthiessen 1999: 240), tudi znanstvena slovenščina teži k prevzemanju vzorca, ki se je izkazal za zelo produktivnega in učinkovitega zlasti pri znanstvenem argumentiranju. Pri tem ne gre samo za transfer tujega vzorca pri prevajanju, temveč za rešitev, ki se marsikateremu avtorju zdi funkcionalno boljša in jo zato sprejema kot svojo.

1.3 Nominalizacija v sodobni slovenščini in primerjave z drugimi jeziki

Študije, posvečene nominalizaciji v sodobni slovenščini bodisi samostojno bodisi v primerjavi z drugimi jeziki, so relativno maloštevilne. Slovenščini je posebej posvečen prispevek Andreje Žele (1996), ki prikazuje diahroni pregled nominalizacije v publicističnem jeziku med letoma 1946 in 1995, iz katerega izhaja, da se število nominalizacij nenehno povečuje, pa tudi Žele (1997), ki prikazuje spremembe pomena

⁶ V svojem delu Vodušek (1933: 72–73) opozarja na »arhaizacijske težnje, ki so imele svoj izvor v panslavistični usmerjenosti«, v skladu s katerimi bi bila moč slovenskega jezika v njegovem verbalnem stilu (v nasprotju z nominalnim). Sam zagovarja stališče, da v takratni slovenščini delavcev, meščanov, tehnikov in znanstvenikov jezik, zaznamovan z nominalnim stilom, »že v resnici nastaja in je samo znak arhaizacijskih in rustikalističnih teženj, ki prevevajo vse naše jezikoslovce in slovnice, da skušajo ta naravni razvoj zaustaviti.« (Prim. tudi Vidovič Muha 2013a: 277–89, zlasti 283–87.)

glagolnikov. V zadnjih letih je nastalo tudi nekaj kontrastivnih oz. prevodoslovnih študij, v katerih avtorji opazujejo pojavljanje nominalizacije v odnosu do njenega pojavljanja v drugih jezikih. Tako že Klinar (1996: 149) ugotavlja, da je v primerjavi z angleščino, ki je bolj samostalniška, slovenščina bolj nagnjena k uporabi glagolov. Plemenitaš (2007) po drugi strani pri analizi vesti in kritik v angleškem in slovenskem jeziku opaza, da se pojavljajo »nekatero podobne težnje v obeh jezikih v pogostnosti in vrstnosti posamostaljenja glede na besedilno vrsto« ter da se v določenih okoliščinah nominalizacije pojavljajo celo pogosteje v slovenskem kot v angleškem jeziku.⁷ Udovič, Žigon in Zlatnar Moe (2011: 275) v diplomatskih besedilih prav tako zasledijo razlike med obravnavanimi jeziki pri rabi različnih izraznih sredstev, med katerimi je tudi nominalizacija. Kuster (2014), ki primerja sistemsko in funkcijsko pogojenost samostalniškega sloga (torej takega, ki ga zaznamuje pogosta raba nominalizacije) v slovenskem in nemškem jeziku, prepoznava veliko mero podobnosti med obema sistemoma. S primerjavo italijanščine in slovenščine se ukvarja tudi Mikolič Južnič (npr. 2007). Obsežnejših primerjav z drugimi jeziki zaenkrat še ni, zanimivo pa je, da so si rezultati omenjenih študij do neke mere nasprotujoči, zato se zdi toliko bolj pomembno temeljiteje preveriti stanje in ponuditi osnovo za prihodnje kontrastivne ali prevodoslovne študije.

2 Korpusi in metoda

2.1 Korpusi

V zadnjih letih so se korpusna orodja za slovenščino izredno hitro razvila in danes imamo na voljo več zanimivih korpusov različnih velikosti, pa tudi orodja za gradnjo lastnih korpusov (gl. npr. Vintar 2013, Fišer in Erjavec 2011, Verdonik in Zwitter Vitez 2011, Vintar, Jerko in Kulovec 2012, Logar Berginc idr. 2012). Tako bogata korpusna infrastruktura nam nudi odlična orodja za poskus razumevanja specifičnih pojavov na podlagi kvantitativnih podatkov: prav na primeru nominalizacije se izkaže, da nam lahko podatki, nabrani iz korpusov, pokažejo bistvene razlike v rabi, kot bomo videli v nadaljevanju.

V naši raziskavi bomo torej podatke, pridobljene z obsežno raziskavo, opravljeno s korpusoma Gigafida in Kres, primerjali z rezultati dveh prejšnjih raziskav (Mikolič Južnič 2011 in 2012b), ki se osredotočata na nekatere znanstvene in strokovne besedilne tipe ter na govor, ter s prevodno slovenščino, kot se izraža v korpusu Spook.

⁷ Čeprav tovrstne prevajalske in kontrastivne težave z nominalizacijo še niso bile obravnavane za angleščino, bi lahko na osnovi primerjave z italijanščino (gl. Mikolič Južnič 2010 in 2013) in na osnovi študij o prevajanju angleških neosebniških glagolskih oblik v slovenščino (Kocijančič Pokorn in Šuštaršič 1999, 2001) sklepali, da je razlog za tako visoko pogostnost nominalizacij verjetno dejstvo, da se z nominalizacijami pogosto nemara prevaja tudi izvorno glagolske strukture, v katerih so uporabljene neosebne glagolske oblike, ki v slovenščino niso neposredno prenosljive.

2.1.1 Korpusi izvirnih slovenskih besedil

Korpus Gigafida je, kot je zapisano na portalu www.slovenscina.eu,⁸ »obsežna zbirka slovenskih besedil najrazličnejših zvrsti, od dnevnih časopisov, revij do knjižnih publikacij vseh vrst (leposlovje, učbeniki, stvarna literatura), spletnih besedil, prepisov parlamentarnih govorov in podobno«, ki vsebuje 1.187.002.502 pojavnic. Namenjena je raziskovanju sodobnega slovenskega jezika. Podrobne informacije o korpusu so na voljo tako na navedeni spletni strani, na straneh samega korpusa (<http://www.gigafida.net>) ter zlasti v Logar Berginc idr. (2012). Za potrebe pričujoče raziskave so najbolj bistveni podatki o Gigafidi zbrani v tabelah 1 in 2.

	Delež	Absolutno št. pojavnic
Celota	100	1.187.002.026
Časopisi	55,91	663.664.965
Drugo	0,67	7.951.450
Internet	15,65	185.757.991
Leposlovje	2,02	23.969.196
Revije	21,51	255.271.089
Stvarna besedila	4,24	50.387.335

Tabela 1: Statistični podatki o pojavnicah v Gigafidi in njenih podkorpusih.

Vendar pa, kot pravita Erjavec in Logar Berginc (2012: 67), »[z]a korpuse, ki predstavljajo celovito podobo nekega jezika, je ključno, da so veliki in besedilnozvrstno pestri. Gigafida je tak referenčni korpus«, vsekakor pa to ni uravnotežen korpus, saj večina besed (77 %) v njem izhaja iz periodičnega tiska. Tako je nastal iz Gigafide vzorčeni uravnoteženi korpus Kres, ki obsega 99.831.145 pojavnic. Tabela 3 prikazuje delež in število pojavnic v posameznih podkorpusih Kresa.

Da bi bila dobljena slika o pojavljanju nominalizacije v slovenskem jeziku čim bolj popolna, bomo podatke, pridobljene s pomočjo zgoraj omenjenih korpusov, primerjali tudi z raziskavami, opravljenimi s korpusom Gos (Mikolič Južnič 2012b) in ad hoc korpusom znanstvenih in strokovnih besedil (Mikolič Južnič 2011). Korpus Gos je prvi in edini referenčni korpus govornega slovenščine; nastal je leta 2008 in je prosto dostopen na spletu kot del portala <http://www.slovenscina.eu>. Gre za uravnoteženi referenčni korpus govornega slovenskega jezika, zgrajen v skladu s tipološkimi in demografskimi kriteriji, ki vsebuje približno milijon besed (prim. Verdonik in Zwitter Vitez 2011).

⁸ Gre za portal projekta *Sporazumevanje v slovenskem jeziku*, ki sta ga v letih 2008–2013 financirala Evropski socialni sklad in Ministrstvo za izobraževanje, znanost in šport.

	Celota	Časopisi	Revije	Internet	Stvarna besedila	Leposlovje	Drugo
2011	44.107.619	0	0	44.107.619	0	0	0
2010	145.793.066	2.890.520	883.222	141.629.373	18.328	371.623	0
2009	84.356.305	55.199.270	23.892.233	0	3.437.957	1.810.283	16.562
2008	121.702.007	90.621.631	24.939.909	0	4.657.606	1.040.440	442.421
2007	108.140.924	81.503.571	22.835.793	0	2.876.657	659.201	265.702
2006	80.777.613	24.055.000	49.764.053	0	4.877.262	1.507.468	573.830
2005	56.890.951	26.105.244	20.651.135	0	5.278.481	4.685.482	170.609
2004	73.875.353	41.012.793	26.148.005	0	3.393.719	3.181.356	139.480
2003	84.597.588	58.565.985	20.770.816	0	2.578.800	2.614.758	67.229
2002	73.962.122	54.634.179	15.239.584	0	2.544.808	1.491.669	51.882
2001	75.976.476	64.320.835	10.133.159	0	808.090	671.846	42.546
2000	66.665.565	53.180.292	9.249.095	3.950	2.906.047	1.088.340	237.841
1999	57.928.325	42.950.926	6.222.997	881	5.463.649	2.091.333	1.198.539
1998	41.824.419	30.502.032	5.216.507	369	3.976.591	1.163.858	965.062
1997	31.007.015	21.083.610	4.279.681	15.799	3.809.647	1.132.863	685.415
1996	16.678.140	10.183.500	4.015.244	0	1.705.716	205.691	567.989
1995	8.303.878	3.833.733	2.931.861	0	663.829	196.605	677.850
1994	7.516.925	3.015.453	2.457.912	0	941.586	51.202	1.050.772
1993	3.137.625	6.391	2.378.872	0	176.003	5.178	571.181
1992	2.244.699	0	2.013.310	0	90.358	0	141.031
1991	1.488.078	0	1.247.701	0	155.317	0	85.060
1990	27.333	0	0	0	26.884	0	449

Tabela 2: Število pojavnic v Gigafidi in njenih podkorpusih po letih.

Delež	Delež	Absolutno št. pojavnic
Celota	100%	1.032.941
Časopisi	20%	19.919.327
Drugo	5%	5.014.206
Internet	20%	20.001.001
Leposlovje	17%	17.030.038
Revije	20%	19.807.912
Stvarna besedila	18%	18.058.661

Tabela 3: Statistični podatki o pojavnicah v Kresu in njegovih podkorpusedih.

Ad hoc korpus KoSinZ šteje približno milijon pojavnic. Besedila, ki so združena v korpusu, sodijo na štiri področja: znanstvena besedila s področja humanistike in družboslovja ter s področja naravoslovja, strokovna pravna besedila ter leposlovna besedila (podrobnejši podatki so na voljo v Mikolič Južnič 2011: 323–24).

2.1.2 Korpusi prevodne slovenščine

Rezultate za izvorna slovenska besedila želimo primerjati s pojavljanjem nominalizacije v prevedenih besedilih. Korpus Spook, ki je bil uporabljen za to raziskavo, vsebuje vzporedne podkorpuse izvornih leposlovnih del in njihovih prevodov za pare angleščina-slovenščina, nemščina-slovenščina, francoščina-slovenščina in italijanščina-slovenščina. Podrobni podatki so v tabeli 4.

	Absolutno št. pojavnic
Izvirna slovenska besedila	1.644.067
Prevodi iz angleščine	918.498
Prevodi iz francoščine	600.121
Prevodi iz italijanščine	389.355
Prevodi iz nemščine	442.773
Vsi prevodi	2.350.747

Tabela 4: Statistični podatki o pojavnicah v izbranih podkorpusedih korpusa Spook.

Da bi zajeli čim večjo žanrsko raznolikost, smo poleg tega uporabili tudi italijansko-slovenski vzporedni korpus⁹ neleposlovnih besedil s 621.036 pojavnicami, ki je podkorpus vzporednega korpusa ISPAC (prim. Mikolič Južnič 2007) in je po vsebini primerljiv s korpusom Stvarnih besedil v Gigafidi oz. Kresu.

Glede na to, da podkorpus leposlovnih besedil v korpusih Gigafida in Kres vključuje tako izvorna kot prevedena dela v slovenščini in med njimi ne ločuje, smo pre-

⁹ Vzporedni korpus je zbirka izvornih besedil v enem jeziku ter njihovih prevodov v drugem jeziku.

verili tudi pojavljanje nominalizacije v manjših (pod)korpusih izvirnih slovenskih besedil, in sicer v podkorpusu izvirnih slovenskih literarnih besedil korpusa Spook (prim. Mikolič Južnič 2015) ter v podkorpusu literarnih besedil omenjenega korpusa KoSinZ.

2.2 Metoda

Raziskava je potekala v več fazah: na eni strani smo čim bolj celostno preverili pojavljanje nominalizacij v izvirni slovenščini, na drugi pa podatke primerjali tudi s pojavljanjem v prevodni slovenščini.

Najprej smo omejili raziskavo na tri priponska obrazila, *-anje*,¹⁰ *-enje* in *-tev*, saj predhodne raziskave (zlasti Mikolič Južnič 2007)¹¹ kažejo, da je približno 85 % vseh priponskih nominalizacij, ki se pojavljajo v Slovarju slovenskega knjižnega jezika (SSKJ 1994), tvorjenih s samo temi tremi priponskimi obrazili.¹² Iz korpusov Gigafida in Kres smo s pomočjo spletnih vmesnikov izluščili vse pojavitve izglagolskih nominalizacij z navedenimi priponskimi obrazili. Čeprav gre za vzorec, ki ne vključuje vseh nominalizacij, menimo, da so podatki dovolj posplošljivi, da lahko na podlagi dobljenih rezultatov z razumno mero gotovosti ocenimo pojavljanje nominalizacije v slovenskem jeziku nasploh.

Iskanje nominalizacij je bilo tako v korpusu Gigafida kot v korpusu Kres izvedeno z možnostjo iskanja po seznamih: vmesnik avtomatsko poišče vse oblike, povezane z iskano lemo (osnovno obliko), zato so bili iskalni pogoji relativno preprosti: z uporabo nadomestnega znaka (*), določanjem besedne vrste *samostalnik* v dodatnih možnostih ter izbiro podrobnosti, kot je na primer iskanje izključno neživih občnih imen srednjega spola, smo za izbrana tri priponska obrazila v korpusu pridobili skupno število pojavitev nominalizacij.

Ker tako Gigafida kot Kres omogočata filtriranje podatkov glede na različne kriterije, povezane s taksonomijo besedil, smo izluščili podatke o pojavljanju nomina-

¹⁰ Zaradi narave korpusnih metod, ki so bile uporabljene pri raziskavi, so vidski morfemi pridruženi obrazilnim, npr. *-a-nje*, *-ova-nje*.

¹¹ V omenjeni raziskavi so bile med drugim pregledane vse nominalizacije v Slovarju slovenskega knjižnega jezika (SSKJ), tvorjene s priponskimi obrazili, ki jih navaja Toporišič (1991: 127–28), vključno s pripono *-s*. Iz te analize (Mikolič Južnič 2007: 132–35) izhaja, da je med vsemi nominalizacijami v SSKJ 55 % tvorjenih s priponskimi obrazili *-anje*, 17 % s pripono *-enje*, 12 % pa s pripono *-tev*. Preostalih 16 % zajema vsa druga priponska obrazila.

¹² Čeprav so v zgoraj omenjene raziskave poleg nominalizacij, tvorjenih s temi tremi priponskimi obrazili, vključene tudi tiste, ki so tvorjene s priponskimi obrazili *-cija* (gre za poenostavljeni prikaz iskalnega pogoja za *-cija*, *-acija*, *-ancija* in *-icija*, omenjene med izglagolskimi obrazili v Voršič 2013; *-cija* v Toporišič, 1991: 127–28) in *-tje*, to v tem primeru ni bilo izvedljivo, saj je bila količina šuma (tj. nepravilnih zadetkov), ki se je pojavljala z zadnjima dvema, tako velika, da je bila zaradi na splošno ogromnega števila primerov iz Gigafide neobvladljiva. Pri preverjanju ustreznosti rezultatov med prvimi 500 zadetki, pridobljenimi z ustreznimi iskalnimi pogoji v Gigafidi, se je izkazalo, da je pri priponskih obrazilih *-anje*, *-enje* in *-tev* šuma malo (do 4 %), medtem ko ga je pri besedotvornih priponah *-cija* in *-tje* neprimerno več (med 27 in 41 %).

lizacij z izbranimi priponskimi obrazili v različnih besedilnih tipih (prim. razdelek 2.1). Poleg tega smo s filtri leta publikacije preverili tudi, ali in v kakšni meri se je spreminjalo pojavljanje nominalizacij longitudinalno, v obdobju zadnjih dvajsetih let.

V naslednji fazi smo pridobljene podatke iz korpusov Gigafida in Kres primerjali s predhodnimi raziskavami, opravljenimi na korpusih Gos in KoSinZ ter s podatki, ki smo jih s primerljivo metodo pridobili za korpuse prevodne slovenščine. Ker gre za korpuse različnih velikosti, smo vse rezultate preračunali na 100.000 pojavnic. Na ta način smo poskusili čim bolj celostno oceniti pojavljanje nominalizacije v slovenskem jeziku. Rezultati izvedenih analiz so opisani v naslednjem razdelku.

Končno smo preverili tudi pojavljanje istih priponskih obrazil v korpusu Spook: poiskali smo vse pojavitve v vseh štirih podkorpusih slovenskih prevodov (iz angleščine, nemščine, francoščine in italijanščine), poleg tega pa smo podatke primerjali tudi s pojavljanjem nominalizacije v prevodih neliterarnih besedil iz italijanščine. Tako pridobljene podatke smo strnili v nadaljevanju.

3 Rezultati in diskusija

3.1 Nominalizacija v izvirnih slovenskih besedilih

Kvantitativna raziskava korpusov izvirnih slovenskih besedil je pokazala precejšnje razlike v pojavljanju nominalizacije. Tabela 5 povzema podatke o pogostnosti nominalizacij v korpusih Gigafida, Kres, Gos, KoSinZ in Spook, preračunane na 100.000 pojavnic.

Ne glede na različno zastopanost besedilnih tipov izkazuje korpusta Gigafida in Kres v celoti zelo podobno pogostnost nominalizacije (1.830 oz. 1.832 primerov).¹³ Podrobnejši pregled pa razkrije, da število izrazito niha pri različnih podkorpusih. V korpusu Gigafida so nominalizacije najpogostejše v podkorpusu *Stvarnih besedil* (2.444) in je na drugem mestu podkorpus *Drugo* (2.310), v Kresu pa so najpogostejše v podkorpusu *Drugo*¹⁴ (2.552), na drugem mestu pa je podkorpus *Stvarnih besedil* (2.445). Pogosteje kot v povprečju celote se nominalizacije pojavljajo še v podkorpusu *Internet* (1.945 v Gigafidi, 2.341 v Kresu), približno enako pogosto se pojavljajo v podkorpusih *Časopisi* (1.821 v Gigafidi, 1.767 v Kresu) in *Revije* (1.729 v Gigafidi, 1.727 v Kresu), bistveno redkeje pa se pojavljajo v podkorpusu *Leposlovje*, in sicer tako v Gigafidi (668) kot v Kresu (806), čeprav so v slednjem nominalizacije v leposlovnem podkorpusu nekoliko pogostejše kot v prvem.

¹³ Podatki so v vseh primerih navedeni preračunano na 100.000 pojavnic.

¹⁴ Kot navajajo Arhar Holdt, Kosem in Logar Berginc (2012: 18), kategorija »Drugo« vsebuje »podnapise filmov, nadaljevanj in dokumentarnih oddaj ter t. i. postproduksijska besedila, ki smo jih vse dobili na RTV Slovenija, iz obeh predhodnih korpusov pa smo sem umestili še zapise sej Državnega zbora RS in besedila z več manjkajočimi bibliografskimi podatki, pri katerih nismo mogli določiti druge kategorije, tudi t. i. besedilni drobiž«.

Korpus		Nominalizacij na 100.000 pojavníc
GIGAFIDA	celota	1.830
	časopisi	1.821
	revije	1.729
	internet	1.945
	stvarna besedila	2.444
	leposlovje	668
	drugo	2.310
KRES	celota	1.832
	časopisi	1.767
	revije	1.727
	internet	2.341
	stvarna besedila	2.445
	leposlovje	806
	drugo	2.552
GOS		546
KOSINZ	celota	2.267
	humanistika in družboslovje	2.386
	naravoslovje	2.664
	pravna	3.138
	Izvirno slovensko leposlovje	366
SPOOK	Izvirna slovensko leposlovje	586

Tabela 5: Pojavitve nominalizacije v različnih korpusih, preračunane na 100.000 pojavníc.

Podatki, zbrani s pomočjo ostalih analiziranih korpusov, se ujemajo z zgoraj omejenimi podatki: izvorno (sodobno) slovensko leposlovje izkazuje nizko prisotnost nominalizacije – še nižjo kot v korpusih Gigafida in Kres: tako je v leposlovnem podkorpusu korpusa KoSinZ le 366 nominalizacij na 100.000 pojavníc, v podkorpusu izvirnih slovenskih leposlovnih besedil korpusa Spook pa 586. Prav tako izrazito malo nominalizacije izkazuje govorni korpus Gos (546). Po drugi strani pa je nominalizacija zelo pogosta v preostalih podkorpusih korpusa KoSinZ, v znanstvenih besedilih – tako humanističnih in družboslovnih (2.386) kot naravoslovnih (2.664) – in v strokovnih, pravnih besedilih, kjer je nominalizacije daleč največ (3.138).

Ob upoštevanju besedil, ki so bila vključena v korpusa Gigafida in Kres (prim. Logar Berginc in Šuster 2009), na podlagi zgoraj navedenih podatkov lahko povza-

memo, da se na splošno obravnavane tipologije nominalizacij pojavljajo izrazito bolj pogosto v neumetnostnih besedilih, pri čemer izstopajo strokovna pravna besedila, znanstvena besedila (zlasti s področja naravoslovja), ter t. i. stvarna besedila, ki prav tako vključujejo strokovna besedila, pa tudi nestrokovna neumetnostna besedila. Izrazito pogosta je nominalizacija tudi v kategoriji *Drugo* korpusov Gigafida in Kres. Zanimiva je razlika med pojavnostjo nominalizacije v podkorpusu *Internet* med Gigafido in Kresom, saj se v slednjem pojavlja za dobrih 25 % večkrat kot v prvem, kar je verjetno pogojeno z izbiro tipa in vsebine spletnih strani, ki so bile vključene v manjši, referenčni korpus. Podkorpusa *Časopisi* in *Revije* sta po vsebnosti nominalizacije zelo podobna povprečni vrednosti za oba korpusa, kar je razložljivo z dejstvom, da so tu zbrana besedila splošne narave, v katerih niso pogosto uporabljena izrazna sredstva, tipična za strokovna/znanstvena besedila. Prav tako pa ta besedila niso po slogu z vidika nominalizacije podobna leposlovnim besedilom, kjer je pojavnost opazovane strukture neprimerno manjša; pri tem gre izpostaviti tudi dejstvo, da je nominalizacija v skladu z našimi podatki redkejša v izvirnem slovenskem leposlovju, kot ga predstavlja analizirani podkorpus Spooka (586 primerov), kot v leposlovnih podkorpusih Gigafide in Kresa, ki vsebujejo tudi prevedena besedila, kot smo že omenili. V govoru se nominalizacija pojavlja s podobno pogostnostjo kot v leposlovju, kar pomeni, da se v primerjavi z ostalimi analiziranimi besedilnimi tipi pojavlja razmeroma redko.¹⁵

3.1.1 Pregled pojavljanja nominalizacije v Gigafidi med letoma 1990 in 2011

Izhajajoč iz ugotovitev, ki jih navaja Žele (1996), smo želeli preveriti, ali se je trend naraščanja nominalizacij nadaljeval tudi med leti 1990 in 2011. Analiza je bila opravljena na korpusu Gigafida, ki ga zaznamuje velik delež publicističnega jezika (žanra, obravnavanega v omenjenem članku). Toda kot vidimo v tabeli 6, se je izkazalo, da na podlagi analiziranih podatkov ni opaznega trenda k povečevanju števila nominalizacij v obravnavanem obdobju: količina sicer nekoliko niha,¹⁶ a le v primeru podkorpusa *Stvarnih besedil* opazimo splošni trend povečevanja (če zaradi slabe reprezentativnosti besedil v korpusu izključimo obdobje 1990–1992), čeprav tudi v tem primeru prihaja do znatnih nihanj. Prav pri publicističnem jeziku (podkorpusa *Časopisi* in *Revije*) opažamo najbolj konstantne vrednosti nominalizacije v analiziranem obdobju; v podkorpusu *Drugo* pa opažamo celo obratni trend, kar je gotovo pogojeno s tipologijo besedil, ki so bila vključena v ta podkorpus (prim. opombo 14).

3.2 Nominalizacija v prevodni slovenščini

Končno smo želeli preveriti tudi, ali se prevodna slovenščina razlikuje od izvirne slovenščine. Pri tem smo se osredotočili zlasti na leposlovje, kjer je nominalizacija

¹⁵ Tudi v govoru se pojavljajo izrazite razlike glede na govorno situacijo in tip diskurza: v bolj spontanem, nenačrtovanem govoru se nominalizacije pojavijo približno 50 % redkeje kot v načrtovanem govoru (ki je navadno vnaprej pripravljen v pisni obliki); prim. Mikolič Južnič (2012b).

¹⁶ V prvih letih (zlasti 1990–1992), ko je bilo zbranega zelo malo korpusnega gradiva, je zelo težko govoriti o trendu, saj so vzorci zaradi omejenosti preveč zaznamovani s slogom posameznih avtorjev.

	Časopisi	Drugo	Internet	Leposlovje	Revije	Stvarna besedila
1990	0	6.013	0	0	0	2.351
1991	0	4.074	0	0	1.624	3.038
1992	0	3.621	0	0	1.684	3.311
1993	1.033	2.865	0	966	1.673	1.433
1994	1.886	2.778	0	570	1.909	1.935
1995	1.816	2.230	0	625	1.933	1.667
1996	1.811	2.391	0	737	1.868	2.588
1997	1.783	2.691	3.526	899	1.980	1.813
1998	1.841	2.176	2.981	825	1.951	2.202
1999	1.852	3.081	0	876	1.872	2.042
2000	1.824	1.657	1.570	769	1.736	2.458
2001	1.870	1.208	0	771	1.812	2.571
2002	1.835	1.249	0	879	1.774	2.336
2003	1.810	1.123	0	720	1.718	2.304
2004	1.813	1.020	0	718	1.675	2.732
2005	1.664	1.178	0	826	1.748	2.985
2006	1.691	1.446	0	765	1.667	2.515
2007	1.836	994	0	677	1.741	2.888
2008	1.866	922	0	1.007	1.687	2.720
2009	1.747	1.274	0	790	1.691	2.467
2010	2.359	0	2.000	1.125	1.885	3.328
2011	0	0	1.768	0	0	0

Tabela 6: Pojavitve nominalizacij v Gigafidi po letih objave, preračunane na 100.000.

v izvorni slovenščini izrazito redkejša kot v drugih besedilnih tipih in kjer bi morda zaradi vpliva tujih jezikov lahko zasledili odstopanja od slovenske norme. Tako v tabeli 7 navajamo podatke, zbrane s korpusom Spook, za prevode iz štirih jezikov ter skupno povprečje v podkorpusu prevodne slovenščine, ki je podobno povprečju leposlovnemu delu korpusa Kres (prim. Tabela 5). Vendar pa v korpusu Spook prihaja do precejšnjih razlik med prevodi iz različnih jezikov: v prevodih iz francoščine je na primer skoraj dvakrat toliko nominalizacij kot v prevodih iz angleščine, v prevodih iz italijanščine jih je prav tako razmeroma malo, v prevodih iz nemščine pa spet nekoliko več. Kljub temu je v vseh literarnih prevodih nominalizacij bistveno manj kot v korpusih drugih besedilnih tipov, navedenih v tabeli 5. Da bi razložili razlike, ki se pojavljajo med posameznimi podkorpusi Spooka, bi gotovo morali preučiti vpliv pojavnosti nominalizacij v izvornih besedilih v različnih jezikih, vendar to presega cilje pričujoče študije.

	Podkorpus	Nominalizacij na 100.000 pojavníc
Leposlovna besedila (Spook)	Prevodi iz nemščine	825
	Prevodi iz angleščine	548
	Prevodi iz francoščine	1.008
	Prevodi iz italijanščine	617
	Skupno povprečje	725
Neleposlovna besedila (ISPAC)	Prevodi iz italijanščine	1.936

Tabela 7: Pojavitve nominalizacije v leposlovnih prevodih iz različnih jezikov v korpusu Spook, ter v prevodih neliterarnih italijanskih besedil iz korpusa ISPAC, preračunane na 100.000 pojavníc.

Ob tem je za primerjavo naveden tudi podatek za prevode neleposlovnih del iz italijanščine, kjer v primerjavi s podkorpusoma *Stvarna besedila* Gigafide in Kresa najdemo bistveno manj nominalizacij: kot smo videli v tabeli 5, je v slednjih 2.444 oz. 2.445 pojavitev na 100.000 pojavníc, kar je za četrtno več kot v prevodih iz italijanščine. Analizirani vzorec prevedenih besedil relativno majhen in omejen na prevode iz enega jezika, vendar bi lahko nakazoval na določeno mero hiperkorektnosti s strani prevajalcev: predhodne študije nakazujejo, da naj bi italijanščina bila na splošno bolj nagnjena k nominalizaciji od slovenščine (prim. Mikolič Južnič 2007); zaradi bojazni, da bi v slovenščino prenesli (pre)veliko število izvernih nominalizacij, se zdi, da jih prevajalci dejansko uporabijo manj, kot je značilno za tovrstna besedila v slovenščini.

4 Sklepi

Na podlagi opravljenih analiz lahko torej potrdimo hipoteze, izpostavljene v uvodu: nominalizacija se pojavlja v vseh besedilnih tipih, ki so bili obravnavani. Tako kot nakazujejo raziskave za druge jezike, pa se tudi v slovenščini pojavljajo izrazite variacije med različnimi besedilnimi tipi. Najmanj nominalizacije je moč zaslediti v izvernih slovenskih leposlovnih besedilih in v govoru, čeprav se tudi v teh primerih pojavljajo znatne razlike. Povprečno število nominalizacij na 100.000 pojavníc v korpusih Gigafida in Kres (pribl. 1.830) lahko z razmeroma visoko mero gotovosti jemljemo kot povprečno vrednost pojavljanja obravnavanih nominalizacijskih priponskih obrazil – in morda nominalizacije na splošno – v pisnem jeziku odraslih govorcev slovenščine ter kot primerjalno vrednost za določanje stopnje nominaliziranosti v različnih žanrih in besedilnih tipih na eni ter za kontrastivne in prevodoslovne raziskave na drugi strani.

V obravnavanih leposlovnih (pod)korpusih je nominalizacij le za dobro četrtno do polovico zgornjega števila (prim. tabelo 5). V govornem korpusu Gos je pogostnost nominalizacije prav tako pričakovano nizka (le tretjino povprečja Gigafide oz.

Kresa), izpostaviti pa velja dejstvo, da je višja kot v nekaterih (pod)korpusih leposlovnih besedil, kar bi lahko bilo presenetljivo, saj načeloma v manj načrtovanih oblikah jezika, kot je govor, najdemo manj kompleksnejših struktur tipa slovnične metafore. Vzrok za to lahko iščemo v dejstvu, da Gos vključuje veliko televizijskih in radijskih oddaj, kjer sicer gre za govor, a so besedila pogosto vnaprej pripravljena v pisni obliki, kar očitno zaznamuje njihove jezikovne značilnosti.

Tudi korpusi leposlovnih besedil, kot rečeno, izkazujejo določene razlike: največ nominalizacij je prisotnih v leposlovnem podkorpusu Kresa, nekoliko manj v primerljivem podkorpusu Gigafide (prim. tabelo 5). Ker oba omenjena podkorpusa vsebujeta tako izvirno kot prevodno leposlovje v slovenščini, in ker je bilo večkrat poudarjeno, da se izvirna in prevedena besedila v istem jeziku lahko razlikujejo z vidika specifičnih slovničnih, pragmatičnih, retoričnih in drugih značilnosti (npr. Baroni in Bernardini 2006), smo preverili pojavljanje nominalizacije tudi v dveh manjših korpusih izključno izvirnega slovenskega leposlovja, in sicer leposlovnem delu korpusa KoSinZ ter podkorpusu izvirnih slovenskih besedil korpusa Spook. Podobno kot prej je tudi v tem primeru opazna razlika v pogostnosti, vendar je v obeh bistveno manj nominalizacije v primerjavi s povprečnimi vrednostmi v Gigafidi in Kresu, kar bi lahko bilo posledica dejstva, da se avtorji v izvirnih slovenskih leposlovnih besedilih nominalizacije poslužujejo manj pogosto kot v prevodni slovenščini. To nam potrjuje tudi tabela 5, v kateri so zbrani podatki o pojavljanju nominalizacije v prevodih iz različnih jezikov, saj je z izjemo angleščine, kjer je nominalizacija nekoliko redkejša kot v ostalih podkorpusih, to strukturo mogoče najti veliko bolj pogosto v prevodni slovenščini (prim. Tabela 7). Za globlji vpogled v razlike, ki se pojavljajo v prevodih iz posameznih jezikov, bi bilo nujno preveriti obseg vpliva izvirnih jezikov, kar je bilo do sedaj narejeno le deloma (prim. odsek 1.3).

Med ostalimi besedilnimi tipi največ nominalizacije po pričakovanjih izkazujejo znanstvena, še bolj pa pravna besedila. V teh dveh skupinah se nominalizacija pojavlja bistveno pogosteje od povprečja, ki ga izkazujeta tako Gigafida kot Kres: v humanističnih in družboslovnih besedilih je 30 odstotkov več nominalizacij od povprečja, v naravoslovnih kar 45 odstotkov več, v pravnih besedilih pa celo 71 odstotkov več. Dejansko se zdi, da je pojavljanje nominalizacije v teh besedilnih tipih prevladujoča slogovna in sintaktična značilnost, za katero ne bi mogli trditi, da izhaja neposredno iz vpliva tujih jezikov, temveč je postala del ubesedovanja in specifično argumentacije v slovenščini na podoben način, kot se je to zgodilo drugod.

Publicistični jezik, kamor prištevamo tako podkorpuse časopisov kot revij, se zelo približuje povprečju celotnih korpusov Gigafida in Kres. Če bi lahko pri Gigafidi trdili, da je v vsakem primeru izrazito zaznamovana s publicističnim jezikom, saj je delež tovrstnih besedil skupno 77 % celote, pa to zagotovo ne velja za korpus Kres, kjer publicistika zavzema skupno zgolj 40 % vseh besedil. Prav zato je še bolj pomenljivo, da sta povprečji pojavljanja nominalizacije v teh dveh korpusih v skladu s to analizo praktično enaki: z razumno mero gotovosti bi lahko kot rečeno to povprečje imeli za kazalnik pojavljanja nominalizacije v slovenščini nasploh. Vendar pa bi lahko

ta podatek s popolno zanesljivostjo potrdili, če bi v raziskavo dejansko vključili vse nominalizacije, kar zaradi trenutnega načina avtomatskega označevanja korpusov (še) ni izvedljivo, predvsem zaradi prevelike količine šuma.

S kvantitativno raziskavo pojavljanja nominalizacije v slovenščini smo želeli ob obstoječih kvalitativnih analizah ponuditi tudi številčne podatke, ki bi lahko v prihodnosti koristili tako kontrastivnim raziskavam med slovenščino in drugimi jeziki kot tudi prevodoslovnim analizam prevajanja nominalizacije iz različnih jezikov v slovenščino in obratno. Nominalizacija je namreč slovnična struktura, ki zaradi svojega vpliva na ubeseditev udeleženskih vlog korenito vpliva tudi na leksikalno gostoto in posledično na razumevanje besedila. V kakšni meri kateri izmed jezikov lahko prenese zgoščevanje informacij, ki ga povzroča nominalizacija, pa je predmet prihodnjih raziskav.

VIRI IN LITERATURA

- Artemis ALEXIADOU, 2001: *Functional structure in nominals: Nominalization and ergativity*. Amsterdam, Filadelfija: John Benjamins.
- Špela ARHAR HOLDT, Iztok KOSEM, Nataša LOGAR BERGINC, 2012: Izdelava korpusa Gigafida in njegovega spletnega vmesnika. *Zbornik Osme konference Jezikovne tehnologije, 8.–12. oktober 2012*. Ur. T. Erjavec, J. Žgank Gros. Ljubljana: Institut Jožef Stefan. 16–21.
- Marco BARONI in Silvia BERNARDINI, 2006: A new approach to the study of translationese: Machine-learning the difference between original and translated text. *Literary and Linguistic Computing* 21/3. 259–74.
- John A. BATEMAN, 1990: Finding translation equivalents: An application of grammatical metaphor. *Proceedings of the 13th conference on computational linguistics: Vol. 2*. Ur. H. Karlgren. Stroudsburg, PA: Association for Computational Linguistics. 13–18.
- Noam CHOMSKY, 1968: Remarks on nominalization. *Readings in transformational grammar*. Ur. A. JACOBS, P. ROSENBAUM. Waltham, MA: Ginn & Co. Publishing. 184–221.
- M. Cecilia COLOMBI, 2006: Grammatical metaphor: Academic language development in Latino students in Spanish. *Advanced Language Learning: The Contribution of Halliday and Vygotsky*. Ur. H. BYMES. London: Continuum. 147–63.
- Tomaž ERJAVEC in Nataša LOGAR BERGINC, 2012: Referenčni korpusi slovenskega jezika (cc)Gigafida in (cc)KRES. *Zbornik osme konference Jezikovne tehnologije: 8.–12. oktober 2012*. Ur. T. Erjavec, J. Žganec Gros. Ljubljana: Institut Jožef Stefan. 57–62.
- Darja FIŠER in Tomaž ERJAVEC, 2011: *sloWNet: slovenski semantični leksikon*. Ljubljana: FF.
- Olga GNAMUŠ, 1979: Izvor in funkcija samostalniškega izražanja. *Jezik in slovstvo* 25/6. 160–64.

- Tanja GRADEČAK-ERDELJIĆ, 2009: *Analitičke glagolsko-imenske konstrukcije u engleskom, hrvatskom i njemačkom jeziku - kognitivni pristup: Doktorska disertacija*. Zagreb: Univerza v Zagrebu.
- Michael Alexander Kirkwood HALLIDAY, 1994: *An Introduction to Functional Grammar*. London: Arnold.
- Michael Alexander Kirkwood HALLIDAY in Christian M. I. M. MATTHIESSEN, 1999: *Construing Experience through Meaning: A Language-Based Approach to Meaning*. London, New York: Continuum.
- , 2004: *An Introduction to Functional Grammar*. 3. izd. London: Arnold.
- Michael Alexander Kirkwood HALLIDAY in James R. MARTIN, 1993: *Writing Science: Literacy and Discursive Power*. Pittsburgh: University of Pittsburgh Press.
- Stanko KLINAR, 1996: Samostalniškost angleščine v primeri s slovenščino. *K tehniki prevajanja iz slovenščine v angleščino*. Ur. S. KLINAR. Radovljica: Didakta. 149–93.
- Nike KOCIJANČIČ POKORN, Rastislav ŠUŠTARŠIČ, 1999: Slovensko-angleška protistavna analiza nedoločnika v vlogi premege predmeta. *Vestnik* 33/1–2. 267–81.
- , 2001: Slovensko-angleška protistavna analiza nedoločnika v vlogi osebkovega oz. povedkovega določila. *Slovenski jezik/Slovene Linguistic Studies* 3. 37–41.
- Tina KOŠAK, 2007: *Posebnosti verbalnega in nominalnega sloga v poslovni nemščini in slovenščini. Magistrsko delo*. Ljubljana: FF.
- Helena KUSTER, 2014: Posamostaljanje med jezikovnim sistemom in besedilom v slovenščini in nemščini. *Slavistična revija* 62/1. 39–54.
- Marija LAPATA, 2002: The disambiguation of nominalizations. *Computational Linguistics* 28/3. 357–88.
- Nataša LOGAR BERGINČ, in Simon ŠUSTER, 2009: Gradnja novega korpusa slovenščine. *Jezik in slovstvo* 54/3–4. 57–68.
- Nataša LOGAR BERGINČ, Miha GRČAR, Marko BRAKUS, Tomaž ERJAVEC, Špela ARHAR HOLDT, Simon KREK, 2012: *Korpusi slovenskega jezika Gigafida, KRES, ccGigafida in ccKRES : gradnja, vsebina, uporaba*. Ljubljana: Trojina, zavod za uporabno slovenistiko, FDV.
- Tamara MIKOLIČ JUŽNIČ, 2007: *Nominalne strukture v italijanščini in slovenščini: Pogostnost, tipi, in prevodne ustreznice: Doktorska disertacija*. Ljubljana: UL.
- , 2010: Translation of Italian nominalizations into Slovene: A corpus-based study. *RITT: Rivista Internazionale di Tecnica della Traduzione: International Journal of Translation* 12. 145–58.
- , 2011: Vpliv besedilnih tipov na pojavljanje nominalizacije v slovenščini: Korpusna raziskava. *Meddisciplinarnost v slovenistiki*: Obdobja 30. Ur. S. Kranjc. Ljubljana: ZIFF. 321–27.
- , 2012a: A contrastive study of nominalization in the systemic functional framework. *Languages in Contrast* 12/2. 251–75.
- , 2012b: La nominalizzazione come indicatore del grado di formalità in alcuni tipi testuali della lingua parlata. *Linguistica* 52. 283–95.

- , 2013: Bridging a grammar gap with explication: A case study of the nominalized infinitive. *Across Languages and Cultures* 14/1. 75–98.
- , 2015: Nominalization in literary texts: A corpus based study of contrastive and translational aspects. *Contrastive Analysis in Discourse Studies and Translation = Analyse contrastive de discours et traduction = Kontrastive Diskursanalyse und Translition*. Ur. M. Schlamberger Brezar, D. J. Limon, A. Gruntar Jermol. Ljubljana: ZIFF. 108–30.
- Lene NORDRUM, 2007: *English Lexical Nominalizations in a Norwegian-Swedish Contrastive Perspective*: Diss. Gothenburg: University of Gothenburg.
- Agnes PISANSKI PETERLIN, 2005: Text-organising metatext in research articles: An English-Slovene Contrastive Analysis. *English for Specific Purposes* 24/3. 307–19.
- Katja PLEMENITAŠ, 2005: Discourse function of nominalization: A case study of English and Slovene newspaper articles. *Acta neophilologica* 38/1–2. 153–66.
- , 2007: *Posamostaljenja v angleščini in slovenščini: Primer časopisnih vesti in kritik*. Maribor: SD.
- Louise RAVELLI, 1988: *Grammatical metaphor: An initial analysis. Pragmatics, Discourse and Text. Some Systemically-inspired Approaches*. Ur. E. H. Steiner, R. Veltman. Norwood, NJ: Ablex. 133–48.
- Jerica SNOJ, 2009: *Metafora v slovenskem slovarju: Doktorska disertacija*. Ljubljana: UL.
- Slovar slovenskega knjižnega jezika*, 1994 (enozvezkovni). Ljubljana: DZS.
- Mall STÅLHAMMAR, 2006: Grammatical metaphor/metonymy in the treaty establishing a constitution for Europe: A comparison between the English and Swedish versions. *NJES (Nordic Journal of English Studies)* 5/1. 99–117.
- Jože TOPORIŠIČ, 1991: *Slovenska slovnica*. Maribor: Obzorja.
- Boštjan UDOVIČ, Tanja ŽIGON, Marija ZLATNAR MOE, 2011: Posebnosti prevajanja strokovnega jezika diplomacije: Primer Dunajske konvencije o diplomatskih odnosih. *Slavistična revija* 59/3. 269–91.
- Darinka VERDONIK in Ana ZWITTER VITEZ, 2011: *Slovenski govorni korpus*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- Ada VIDOVIČ MUHA, 1986: Tipološki pregled nekaterih vplivov na slovenski znanstveni jezik. *Slovenski jezik v znanosti 1*. Ur. Ada Vidovič Muha. Ljubljana: ZIFF.
- Ada VIDOVIČ MUHA (ur.), 1986: *Slovenski jezik v znanosti 1. Zbornik prispevkov*. Ljubljana: ZIFF.
- , 2011: *Slovensko skladenjsko besedotvorje*. Ljubljana: ZIFF.
- , 2013: *Moč in nemoč knjižnega jezika*. Ljubljana: ZIFF.
- Špela VINTAR (ur.), 2013: *Slovenski prevodi skozi korpusno prizmo*. Ljubljana: ZIFF.
- Špela VINTAR, Boštjan JERKO, Marjetka KULOVEC, 2012: Korpus slovenskega znakovnega jezika. *Zbornik Osme konference Jezikovne tehnologije, 8.–12. oktober 2012: Zbornik 15. mednarodne multikonference Informacijska družba - IS 212, zvezek C*. Ur. T. Erjavec, J. Žganec Gros. Ljubljana: Institut Jožef Stefan, 2012. 203–06.
- Božo VODUŠEK, 1933: Za preureditev nazora o jeziku. *Krog: Zbornik umetnosti in*

- razprav.* Ur. R. Ložar. Ljubljana: Sotrudniki, 1933. 66–76.
- Ines VORŠIČ, 2013: *Sistemska in nesistemska leksikalna tvorjav novejšem besedju slovenskega jezika: Doktorska disertacija.* Maribor: UM.
- Martina ZAJC, 2014: Nominalizacija kot slovnična metafora v predlogih raziskovalnih projektov v slovenščini in angleščini: Kontrastivna in prevodoslovna analiza. *Prevodoslovno usmerjene kontrastivne študije.* Ur. A. Pisanski Peterlin, M. Schlamberger Brezar. Ljubljana: ZIFF. 208–28.
- Andreja ŽELE, 1996: Razvoj posamostaljenja v slovenskem publicističnem jeziku med 1946 in 1995. *Jezik in čas.* Ur. A. Vidovič Muha. Ljubljana: ZIFF. 191–200.
- , 1997: Slovenski razvoj besedotvornih pomenov pri izglagolskih samostalnkih, posebno pri glagolniku. *Slovenski jezik/Slovene Linguistic Studies* 1. 69–90.

SUMMARY

The article is an attempt to provide an overview of the occurrence of nominalization in the Slovene language. Quantitative corpus research is used to verify the extent to which nominalization has permeated various genres of both original and translated adult written Slovene language. Nominalization is defined within Halliday's framework of Systemic Functional Grammar.

In the Systemic Functional Grammar framework, nominalization is explained as a grammatical metaphor of the ideational type: a process, which would be congruently realized by a verb, is metaphorically realized by a noun. This change entails a rearrangement of all the metafunctions and participants in the process, thus resulting in a structure that is more sophisticated and more condensed, i.e., has a higher lexical density. According to the literature, this type of discourse device is used in all adult language, but its presence is highly dependent on register and text type. In this article, the aim is to verify whether the use of nominalization in Slovene is in line with the findings for other languages, i.e., how much its use varies with genre and register. Furthermore, it has been claimed that not all languages cope with high levels of lexical density equally well, and Slovene is believed to be among such languages that do not (sometimes called verbal, as opposed to nominal ones).

A corpus-based methodology is used to compare original Slovene texts of different genres and translations from selected foreign languages. Several corpora have been used in order to evaluate the frequency of nominalization in different text types: the Gigafida and Kres corpora (comprising newspapers, magazines, internet, literature, nonfiction, and other texts); the Gos corpus of spoken Slovene (non-public—private discourse, public—entertainment discourse, public—informative and educational discourse, and non-public—non-private discourse); an ad hoc corpus of scientific, legal, and literary texts; and the Spook parallel multilingual corpus (Slovene translations from Italian, English, German, French, as well as original Slovene texts). Instances of nominalizations formed with the three most frequently used deverbal nominal suffixes (*-anje*, *-enje*, *-tev*) were extracted from all the corpora to compare the frequency in the various genres.

The results seem to confirm that the nominalizations analyzed are indeed present in all the genres and that their frequency is highly dependent on genre, with the frequency ranging from as little as 366 instances per 100,000 words in literary texts to as much as 3,138 instances per 100,000 in legal documents. The mean frequency displayed by the large corpora (Gigafida, Kres) is in line with the frequency displayed by the newspaper and magazines subcorpora, therefore it could be inferred that this is the average rate of nominalization in written Slovene, although to prove this, further analysis would be needed; that analysis would include a larger number of deverbal suffixes and consequently a more accurate sample of analysis.

According to the gathered data, within the analyzed period (1990–2011), despite slight variations, there seems to be no noticeable increase in the frequency of occurrence of nominalizations: virtually in all the analyzed subcorpora the numbers remain approximately at the same level in the two decades analyzed (with the exception of the non-literary subcorpus of the Gigafida and Kres corpora). It would seem that despite the increase in frequency in the second half of the 20th century reported by previous studies on the language of newspapers, the trend has subsided and the numbers have become more stable.

Furthermore, the study found evidence that nominalization seems to be more frequent in translated literary texts than in original ones and that there are differences in frequency between translations from different languages, but more detailed studies are needed to confirm this and to explain the reasons behind the differences among translations from different languages.