

## LA NOMINALIZZAZIONE COME INDICATORE DEL GRADO DI FORMALITÀ IN ALCUNI TIPI TESTUALI DELLA LINGUA PARLATA

### 1 INTRODUZIONE

Nonostante la naturale primarietà del parlato rispetto allo scritto, tradizionalmente, la lingua parlata è stata trascurata o addirittura evitata nelle analisi linguistiche, data la sua natura acustica e inafferrabile, nonché l'informalità delle situazioni alle quali essa viene tipicamente associata. Solo con l'avvento della tecnologia e con la conseguente possibilità di registrare e, in seguito, digitalizzare il suono, si è resa possibile un'analisi sistematica della lingua parlata. Negli ultimi anni, l'ingigantimento delle capacità di memoria dei computer e lo sviluppo di nuovi software capaci di sincronizzare con relativa semplicità le trascrizioni del parlato alle registrazioni, hanno reso possibili progetti di ricerca in passato impensabili. Sono nati così i corpora della lingua parlata, tra i quali anche il corpus referenziale dello sloveno parlato GOS.

Il corpus linguistico è definito come una raccolta di testi che sia rappresentativa di una data lingua e che sia disponibile in una forma utilizzabile per analisi linguistiche (Tognini-Bonelli 2001: 2), dove il formato in questione è oggi naturalmente quello elettronico e dove si presuppone anche che i testi raccolti siano di produzione naturale e scelti con criteri chiaramente definiti. Un corpus parlato, oggi, solitamente offre la trascrizione (ortografica e/o altro) delle registrazioni di eventi comunicativi diversi (da conversazioni spontanee a registrazioni di trasmissioni radiofoniche ecc.), la quale include, tra l'altro, tutti gli elementi lessicali e non della produzione e a cui si aggiunge la segmentazione e l'annotazione dei testi. L'uso di convenzioni e annotazioni standardizzate permette al ricercatore di svolgere analisi precise e di alta qualità. In tal modo diventa infatti possibile osservare innumerevoli fenomeni linguistici quali aspetti fonetici, prosodici, morfologici e sintattici attraverso i generi testuali del parlato inclusi nel corpus analizzato.

Questo contributo presenta una ricerca in cui si vuole verificare se e in quali modi differiscono tra loro vari generi di testi parlati sloveni in relazione a una caratteristica specifica, solitamente associata piuttosto alla lingua scritta, ovvero la nominalizzazione. Il fenomeno, definito con più precisione nella sezione 3, è spesso associato ai registri formali della lingua scritta, pertanto l'ipotesi che si propone nel presente scritto è che la nominalizzazione avrà un numero di occorrenze molto basso nel cor-

---

\* *Indirizzo dell'autrice:* Filozofska fakulteta, Oddelek za prevajalstvo, Aškerčeva 2, 1000 Ljubljana, Slovenia. E-mail: tamara.mikolic@guest.arnes.si

pus analizzato: sarà pressoché assente nei generi parlati meno sorvegliati quali la comunicazione telefonica privata, mentre sarà probabilmente un po' più presente in testi letti, preparati per iscritto in anticipo, come ad esempio numerose trasmissioni televisive o radiofoniche. A tale scopo ci avvarremo delle possibilità di analisi separate di determinate parti del corpus referenziale del parlato GOS.

## 2 LO STUDIO DELLA LINGUA PARLATA SLOVENA

Sebbene gli inizi dello studio moderno del discorso parlato sloveno si possano datare agli anni Sessanta (p. es. Pogorelec 1965), è durante l'ultimo decennio che i fenomeni del parlato sono stati studiati con nuova profondità (cf. Zemljarič Miklavčič 2008: 15–16). Lasciando da parte i numerosi studi dialettologici, le ricerche si sono occupate di aspetti pragmatici del parlato spontaneo (Verdonik 2007) e dei suoi vari generi (Smolej 2012), di generi specifici quali i dialoghi giornalistici radiofonici (Verovnik 2010), di prosodia (Zwitter Vitez 2005) ecc. Per ragioni di spazio non è possibile entrare in maggiori dettagli, né per altro offrire un elenco esaustivo delle opere dedicate al parlato sloveno. Le sopraccitate opere monografiche vengono menzionate con il solo scopo di dare un'idea sul crescente interesse e sull'attenzione dedicati a progetti legati alla lingua parlata slovena.

Per quanto riguarda la nominalizzazione, infine, non sembrano essere reperibili studi sulla sua presenza nella lingua parlata slovena, mentre si riscontrano invece numerosi studi di altre lingue, alcuni dei quali sono menzionati nella sezione seguente.

## 3 LA NOMINALIZZAZIONE E LA LINGUA PARLATA

Il termine nominalizzazione viene usato per denominare quel tipo di metafora grammaticale per cui, riassumendo Halliday/Matthiessen (2004: 636–658), un processo, che congruentemente verrebbe codificato con un verbo, viene espresso con un nome. L'esempio (1a)<sup>1</sup> presenta una frase slovena (e la relativa traduzione in italiano) tratta dal corpus GOS, in cui *izjava* “dichiarazione” è la nominalizzazione che potrebbe essere espressa congruentemente con il verbo *izjaviti* “dichiarare”, come si vede in (1b).

- (1) a. *mal sem bil razočaran ker je nekako dal eno tako izjavo zlo podobno tej*  
“ero un po' deluso perché ha fatto una dichiarazione così simile a questa”  
b. *mal sem bil razočaran ker je nekako izjavil nekaj zelo podobnega tej*  
“ero un po' deluso perché ha dichiarato qualcosa di molto simile a questa”

Secondo Halliday/Matthiessen (ibid.: 654), il fattore più determinante per la quantità di metafore nella grammatica di un testo è se si tratta un testo scritto o parlato. I due modi comunicativi implicano tipi diversi di complessità – la lingua scritta è più complessa in termini di densità lessicale,<sup>2</sup> mentre quella parlata lo è per la sua intri-

<sup>1</sup> L'esempio (1a) è tratto dal corpus GOS, mentre la variante (1b) è opera dell'autrice dell'articolo.

<sup>2</sup> La densità lessicale è definita come la relazione tra il numero di parole lessicali e grammaticali in una proposizione (cf. Halliday/Martin 1993: 76–77).

cata grammatica proposizionale (ibid.: 655) – ed è il linguaggio scritto (ibid.: 656) quello che implica la presenza della metafora grammaticale.

Anche Chafe (1982: 46) collega l'uso della nominalizzazione allo scritto formale e al distacco (*detachment*, ibid.) che caratterizza lo scritto, benché ammetta (ibid.: 48) che esso sia riscontrabile anche in alcuni generi del linguaggio parlato.

Biber et al. (1998) presentano uno studio sulla frequenza della nominalizzazione<sup>3</sup> in alcuni generi testuali inglesi e appurano che essa è di gran lunga più frequente negli scritti accademici, sia se comparata con opere letterarie sia al confronto con il parlato. La sua distribuzione è invece molto simile nelle opere letterarie e nel corpus parlato (con una leggera maggioranza in quest'ultimo). In conclusione, gli autori deducono (ibid.: 65) che la nominalizzazione come gruppo, come anche i suoi singoli suffissi derivazionali, presentano forti legami di associazione con diversi registri, riflettendo le funzioni comunicative primarie dei registri.

Considerando quanto esposto sopra, si è voluto verificare se anche nella lingua parlata slovena si riscontrano le stesse tendenze evidenziate per altre lingue, in particolare l'inglese. Oltre alla generale presenza della nominalizzazione nel parlato sloveno, però, lo studio è orientato al paragone tra diversi tipi di discorso parlato: si ipotizza che i generi parlati possano essere caratterizzati in qualità di entità distinte anche in relazione alla frequenza della nominalizzazione in essi riscontrata, poiché diverse situazioni comunicative e la conseguente differenziazione del grado di formalità possono dar luogo a una presenza maggiore o minore di nominalizzazioni.

In sloveno la nominalizzazione, comunque, non è stata ancora oggetto di uno studio generale che ne determinasse la presenza (né nello scritto né tanto meno nel parlato), sebbene esistano studi su tipi testuali (scritti) particolari, quali Žele (1996) o Mikolič Južnič (2011), o paragoni con altre lingue come Plemenitaš (2007), Mikolič Južnič (2007) o Klinar (1996). I risultati potranno così colmare una lacuna sia dal punto di vista dell'analisi del discorso parlato che da quello sintattico-grammaticale.

#### 4 IL CORPUS GOS

Il primo (e unico) corpus referenziale del parlato sloveno, il corpus GOS, è nato nel 2008. Il corpus, disponibile gratuitamente in rete, è stato concepito come parte del progetto *Sporazumevanje v slovenščini*<sup>4</sup> (“Comunicazione in Sloveno”).

Si tratta, come già detto, di un corpus referenziale della lingua parlata slovena bilanciato secondo criteri tipologici e demografici, contenente circa un milione di parole. L'intero processo di compilazione del corpus, assieme ad alcuni esempi del suo uso, è descritto nell'opera monografica di Verdonik/Zwitter Vitez (2011). La Tabella 1 rias-

---

<sup>3</sup>In Biber et al. (1998: 59–65) la nominalizzazione è intesa in un senso leggermente più ampio di quello usato nel presente studio. Gli autori, infatti, comprendono anche le nominalizzazioni di aggettivi (o meglio, attributi), non solo quelle di verbi (ovvero processi). Va detto che in entrambi i casi si tratta comunque dello stesso tipo di metafora grammaticale, ossia la reificazione di elementi primariamente (o congruentemente) non nominali.

<sup>4</sup><http://www.slovenscina.eu/vsebine/en/domov/domov.aspx>

sume gli aspetti strutturali salienti del corpus ai fini della presente ricerca. Secondo il tipo di discorso, il corpus è suddiviso in due categorie principali, ovvero il discorso pubblico e il discorso non pubblico, le quali sono poi suddivise ulteriormente in relazione al contenuto del discorso (il discorso pubblico comprende il genere informativo-educativo e quello d'intrattenimento, mentre il discorso non pubblico comprende il genere non privato e quello privato). La tipologia si ramifica anche rispetto al canale usato per la comunicazione (radio, televisione, contatto personale, telefono) e ad altri fattori pragmatici che però non sono stati presi in considerazione in questa ricerca.

Tipo di discorso		Canale	No. parole	%
pubblico	informativo-educativo	radio	94.536	9,16
		televisione	102.263	9,90
		contatto personale	162.750	15,76
	di intrattenimento	radio	123.152	11,93
		televisione	105.613	10,22
non pubblico	non privato	contatto personale	119.987	11,62
		telefono	33.484	3,24
	privato	contatto personale	222.907	21,58
		telefono	68.083	6,59
<b>Totale</b>			<b>1.032.775</b>	<b>100</b>

Tabella 1: Dati statistici del corpus GOS (adattato da Verdonik/Zwitter Vitez 2011: 109)

## 5 METODO

La ricerca presentata è articolata in tre fasi principali. Inizialmente sono state selezionate, in base ad analisi precedenti (Mikolič Južnič 2007), le desinenze deverbali nominali slovene più produttive;<sup>5</sup> quindi se n'è verificata la presenza nel corpus GOS (e, più dettagliatamente, nei suoi vari subcorpora) con l'aiuto del motore di ricerca disponibile al sito internet. Infine, i dati così estratti sono stati analizzati ed interpretati in relazione alle ipotesi esposte nella sezione 1.

Le desinenze nominali che sono state verificate sono le seguenti: *-anje* (esempio 2), *-enje* (3), *-tev* (4), *-cija* (5), *-tje* (6), *-ava* (7) e *-ek* (8).<sup>6</sup>

(2) *ki jih večinoma lahko porabimo kr tako brez predhodnega kuhanja na voljo so nam tud zeleni pripravljene s špinačo in*

“che possiamo per lo più usare così senza previa cottura sono a disposizione anche quelli verdi preparati con spinaci e”

<sup>5</sup> Con produttività si intende in realtà la frequenza con la quale le desinenze sono usate nel grande dizionario sloveno Slovar slovenskega knjižnega jezika o SSKJ (1994), non le reali occorrenze in un qualsiasi corpus. Le sette desinenze studiate sono usate in più del 90% di tutte le nominalizzazioni derivate presenti nel dizionario.

<sup>6</sup> Tutti gli esempi sono tratti dal corpus GOS.

- (3) *in si tako z štirimi srečnimi bomo rekli vrtenji kolesa sreče prisluzila deset tisoč*  
 “e con quattro fortunati chiamiamoli così giri della ruota della fortuna guadagnare dieci mila”
- (4) *te okrepijo eee to kredibilnost eem Obama veliko govori o krepitvi mednarodnega prava eee dejmo ga spodbudit da vstopi v kekšen*  
 “queste rafforzano eee questa credibilità eee Obama parla molto di rafforzamento del diritto internazionale eee incoraggiamolo a entrare in qualche”
- (5) *je imel čas torej par mescev za zbiranje podpisov za organizacijo referendum in na podlagi tega da da prepreči*  
 “ha avuto quindi un paio di mesi per la raccolta di firme per l’organizzazione del referendum e in base a questo per per impedire”
- (6) *ne vem a bom zadela te tone in tko pa petje v živo je res nekej kr neskončno občudujem in tukej*  
 “non so se beccherò quelle note e così e il canto dal vivo è veramente qualcosa che adoro immensamente e qui”
- (7) *v postmoderno družbo pa ni mogoče da bi prišlo do izboljšav samo po sebi zaradi tega ker osnova tega prehoda je*  
 “nella società postmoderna ma non è possibile che si abbiano miglioramenti da sé a causa del fatto che la base di questo passaggio è”
- (8) *stroko je zmotu in se mi zdi da tud ta poudarka ki sta bla tukej dana nnn nista prov vsebinsko pravilna*  
 “ha disturbato gli esperti e mi sembra che anche gli accenti posti qui non siano proprio giusti dal punto di vista del contenuto”

Il motore di ricerca permette di controllare vari parametri contemporaneamente, alcuni dei quali vanno selezionati nella finestra iniziale, come il tipo di ricerca (semplice o avanzata) o la ricerca per trascrizione standardizzata o per versione parlata, altri invece possono essere determinati in seguito, dopo aver ottenuto la lista di risultati, come ad esempio il tipo di discorso, il canale, la regione di provenienza dei parlanti ecc. Nella presente ricerca abbiamo scelto l’opzione della ricerca avanzata, in cui si può determinare la categoria grammaticale della parola ricercata. Ma poiché le nominalizzazioni ovviamente non sono marcate diversamente da tutti gli altri nomi presenti nel corpus, ci siamo serviti della “wild card” ricercando tutte le parole che terminano con le desinenze elencate. Ciò ha comportato naturalmente che insieme agli esempi desiderati, contenenti nominalizzazioni, sono stati estrapolati anche numerosi esempi con altri tipi di nomi (ad esempio *stanovanje* “appartamento” con la desinenza *-anje*), che sono stati in seguito eliminati manualmente. Inoltre, dopo la selezione iniziale, i risultati ottenuti sono stati suddivisi in relazione ai parametri di tipologia testuale e di canale esposti nella Tabella 1. I risultati dell’analisi sono presentati nella sezione a seguire.

## 6 RISULTATI E DISCUSSIONE

Nel corpus GOS sono state dunque trovate, in base alla metodologia esposta, un numero complessivo di 7.013 nominalizzazioni. Rispetto al numero di tutti i nomi presenti nel corpus (153.148), le nominalizzazioni ammontano al 4,58%.<sup>7</sup> Il numero indubbiamente non è altissimo, ma non può di certo essere ignorato. Esso testimonia il fatto che la nominalizzazione, attestata precedentemente in alcuni generi scritti dello sloveno, com'è stato menzionato nella sezione 3, è penetrata anche nel parlato, nonostante sia stato più volte affermato che, in sloveno, la struttura tenda a risultare innaturale e di difficile comprensione, specialmente se usata di frequente (cf. Klinar 1996, Žele 1996).

### 6.1 Distribuzione delle singole desinenze

La Tabella 2 mostra dettagliatamente la distribuzione delle nominalizzazioni formate con le singole desinenze analizzate nei vari generi testuali del parlato.

Tipo di discorso	Desinenze (occorrenze assolute)						
	-anje	-enje	-tev	-cija	-tje	-ava	-ek
<b>Pubblico informativo educativo</b>	<b>1.683</b>	<b>636</b>	<b>468</b>	<b>588</b>	<b>74</b>	<b>338</b>	<b>452</b>
Televisione	600	180	211	215	16	138	156
Radio	546	204	143	158	20	60	136
Contatto personale	537	252	114	215	38	140	160
<b>Pubblico d'intrattenimento</b>	<b>501</b>	<b>222</b>	<b>105</b>	<b>108</b>	<b>34</b>	<b>97</b>	<b>276</b>
Radio	332	127	57	50	22	50	104
Televisione	169	95	48	58	14	47	172
<b>Non pubblico non privato</b>	<b>318</b>	<b>94</b>	<b>97</b>	<b>124</b>	<b>8</b>	<b>89</b>	<b>118</b>
Contatto personale	208	70	72	77	2	74	92
Telefono	110	24	25	47	6	15	26
<b>Non pubblico privato</b>	<b>173</b>	<b>117</b>	<b>34</b>	<b>104</b>	<b>8</b>	<b>56</b>	<b>91</b>
Contatto personale	142	100	30	75	8	38	86
Telefono	31	17	4	29	0	18	5
<b>Totale</b>	<b>2.675</b>	<b>1.069</b>	<b>704</b>	<b>924</b>	<b>124</b>	<b>580</b>	<b>937</b>

Tabella 2: Statistica delle nominalizzazioni riscontrate nel corpus GOS.

Dalla Tabella 2 si può dedurre che la desinenza più usata nel parlato sloveno in generale (come rappresentato nel corpus GOS) è *-anje*, con 2.675 occorrenze, il che è in linea con la frequenza preponderante di questa desinenza riscontrata in alcuni generi testuali sloveni scritti (cf. Mikolič Južnič 2011) e con la sua grande produttività

<sup>7</sup> La percentuale è stata calcolata in base ad una ricerca con la sola "wild card" (\*, ovvero zero o più caratteri) per la categoria grammaticale del nome.

riscontrata nell'analisi del dizionario SSKJ (cf. Mikolič Južnič 2007). Sono molto meno frequenti le altre desinenze, tra cui primeggia *-enje* (con 1.069 occorrenze, cioè meno della metà di quelle con *-anje*), quindi *-ek* (937), *-cija* (924), *-tev* (704), *-ava* (580) e *-tje* (124). L'immagine 1 presenta questi dati sotto forma di percentuali del numero complessivo di tutte le nominalizzazioni riscontrate.

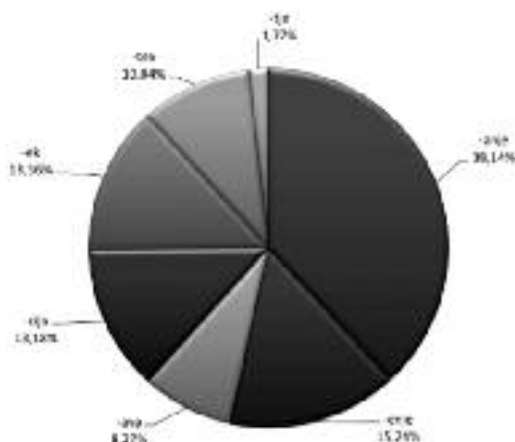


Immagine 1: Percentuali delle singole desinenze delle nominalizzazioni riscontrate nel corpus GOS.

## 6.2 Tipi testuali del parlato e nominalizzazione

Molto più significativo, comunque, della statistica sulla presenza delle varie desinenze delle nominalizzazioni analizzate, è osservare come si differenzia il parlato in relazione ai tipi di discorso e ai canali usati. La Tabella 3 mostra i dati relativi ai due tipi di discorso generali (discorso pubblico e non pubblico), mentre nella Tabella 4 troviamo i relativi sottotipi (discorso informativo-educativo, d'intrattenimento, non privato e privato). Infine, la Tabella 5 offre uno sguardo sulle differenze riscontrabili nei discorsi appartenenti ai vari canali usati (televisione, radio, contatto personale, telefono). Al fine di dare un termine di paragone più completo per la frequenza delle nominalizzazioni, le tre tabelle indicano anche i dati sul numero di parole grafiche nel singolo subcorpus nonché il numero di tutte le parole categorizzate come nomi (sostantivi) presenti in esso.

	Discorso pubblico	Discorso non pubblico
<b>Numero parole grafiche</b>	588.314	444.461
<b>Numero nomi</b>	102.490	50.658
<b>% nomi</b>	17,42	11,40
<b>Numero nominalizzazioni</b>	5.582	1.431
<b>% nominalizzazioni</b>	5,45	2,82

Tabella 3: Occorrenze delle nominalizzazioni e dei nomi nei due tipi testuali principali.

	Discorso pubblico		Discorso non pubblico	
	informativo educativo	d'intrattenimento	non privato	privato
<b>Numero parole grafiche</b>	359.549	228.765	153.471	290.990
<b>Numero nomi</b>	64.924	37.566	18.827	31.831
<b>% nomi</b>	18,06	16,42	12,27	10,94
<b>Numero nominalizzazioni</b>	4.239	1.343	848	583
<b>% nominalizzazioni</b>	6,53	3,58	4,50	1,83

Tabella 4: Occorrenze delle nominalizzazioni e dei nomi nei quattro sottotipi testuali.

	Canale			
	televisione	radio	contatto personale	telefono
<b>Numero parole grafiche</b>	207.876	217.688	505.644	101.567
<b>Numero nomi</b>	36.720	38.837	65.850	11.741
<b>% nomi</b>	17,66	17,84	13,02	11,56
<b>Numero nominalizzazioni</b>	2.530	2.009	2.117	357
<b>% nominalizzazioni</b>	6,89	5,17	3,21	3,04

Tabella 5: Occorrenze delle nominalizzazioni e dei nomi nei quattro canali analizzati.

Osservando le tre tabelle notiamo che il discorso parlato analizzato comprende un numero di nominalizzazioni relativamente basso (ma è tutt'altro che privo di esse); tuttavia è comparando i vari generi che si notano differenze importanti. In primo luogo, appare significativo il dato che la percentuale di nomi tra tutte le parole grafiche nel discorso pubblico ammonta al 17,42%, mentre nel discorso non pubblico ve ne sono solo l'11,40%, il che potrebbe indicare un legame tra il tipo di discorso e il grado di pianificazione e controllo presente nel singolo genere di discorso, considerando che un linguaggio ricco di nomi viene associato a un grado più elevato di formalità data la maggiore densità lessicale che comporta, come abbiamo visto nella sezione 3. Il discorso pubblico, per sua natura, richiede più preparazione e spesso un più alto grado di formalità rispetto al discorso non pubblico. Ciò sembra riflettersi sia nel più alto tasso di nomi sia nel tasso nettamente più alto di nominalizzazioni<sup>8</sup> (il 5,45% nel discorso pubblico contro il 2,82% in quello non pubblico), come indica la Tabella 3, da cui si può dedurre che siano i discorsi pubblici quelli caratterizzati dalla nominalizzazione, considerando che ne contengono quasi il doppio dei discorsi non pubblici.

<sup>8</sup> Naturalmente si parla delle nominalizzazioni derivate con i sette suffissi analizzati; è chiaro che tra i nomi presenti nei vari subcorpora sono presenti anche altre nominalizzazioni, che tuttavia in base alle ricerche statistiche esposte sopra (citando Mikolič Južnič 2007) non dovrebbero raggiungere numeri abbastanza alti da poter modificare sensibilmente i risultati ottenuti.



La situazione, tuttavia, non è così semplice se si considerano i dati nella Tabella 4. Se le percentuali di nomi nei due tipi di discorso pubblico (informativo educativo e d'intrattenimento) da una parte e in quelli del discorso non pubblico (non privato e privato) dall'altra sono piuttosto vicine (18,06% vs. 16,42% nel primo, 12,27% vs. 10,94% nel secondo caso), la differenza è più eloquente nella percentuale di nominalizzazioni tra questi nomi. È il discorso pubblico informativo educativo quello con la percentuale più alta (6,53%), mentre nel discorso pubblico d'intrattenimento le nominalizzazioni sono presenti con una frequenza pari a poco più della metà (3,58%). Si conferma anche qui l'uso più esteso della nominalizzazione nel genere dove il grado di formalità della lingua richiesto è più elevato. Tra i due generi di discorso non pubblico, poi, la discrepanza è ancora più marcata: nel discorso non privato le nominalizzazioni rappresentano il 4,50% di tutti i nomi (addirittura più che nel discorso pubblico d'intrattenimento), mentre in quello privato ve ne sono un mero 1,83%. Considerando che tra gli eventi comunicativi del discorso non pubblico non privato, Verdonik/Zwitter Vitez (2011: 109) enumerano riunioni formali e informali, consultazioni, servizi, colloqui formali ecc., mentre tra quelli del discorso privato troviamo dialoghi in famiglia o tra amici o conoscenti, si conferma di nuovo il legame tra un uso più frequente della nominalizzazione e la formalità del discorso.

Similmente, anche nella Tabella 5, la nostra ipotesi viene confermata: i canali con un'impostazione più sorvegliata (e formale) mostrano una presenza più estesa di nominalizzazioni (nel discorso televisivo ammontano al 6,89%, in quello radiofonico al 5,17%). Nei discorsi dove spesso il grado di formalità si abbassa, come il contatto personale (3,21%) e il telefono (3,04%), invece, se ne fa un uso minore. Coincide con gli altri dati sopra esposti e con la nostra ipotesi generale anche la percentuale di nomi rispetto al numero di tutte le parole presenti nei subcorpora: gli eventi comunicativi che sono avvenuti tramite contatto personale o telefono sono caratterizzati da una minore presenza di nomi in generale, non solo dal minor numero di nominalizzazioni tra questi nomi.

Andando oltre l'aspetto quantitativo e accennando di passaggio a quello qualitativo, si nota nel discorso parlato analizzato, anche solo osservando gli esempi citati nella sezione 5, la tipica complessità sintattica del parlato, caratterizzata da una densità lessicale relativamente bassa e da una sintassi con numerose proposizioni coordinate o subordinate. Nell'esempio (9), che ripetiamo per comodità aggiungendovi una versione congruente della nominalizzazione, ci sono tre proposizioni con 20 parole, tra cui 13 parole lessicali; dunque la densità lessicale è pari a 4,33 parole lessicali per proposizione. Poiché si tratta di una frase che comprende una nominalizzazione (*kuhanje*), la densità della proposizione in questione avrebbe potuto essere ancora più bassa se invece della nominalizzazione fosse stata usata una forma congruente come quella espressa in (9b), dove la densità lessicale scende addirittura a 3,50, vista l'introduzione di una proposizione in più.

- (9) a. *ki jih večinoma lahko porabimo kr tako brez predhodnega kuhanja na voljo so nam tud zeleni pripravljene s špinačo in*  
 “che possiamo per lo più usare così senza previa cottura sono a disposizione anche

quelli verdi preparati con spinaci e”

b. *ki jih večinoma lahko porabimo kr tako ne da bi jih predhodno kuhali [...]*

“che possiamo per lo più usare così senza che le cucinassimo prima [...]”

La situazione sembra essere molto simile anche negli altri esempi: la nominalizzazione condensa leggermente il messaggio, ma trattandosi in genere di casi isolati, di nominalizzazioni relativamente lontane l’una dall’altra, e non di un uso frequente, la densità lessicale aumenta, ma di poco e quindi non sembra avere un grande impatto sui testi in generale. Comunque sia, si tratta per lo più di ipotesi che andrebbero confermate con una ricerca approfondita.

## 7 CONCLUSIONI

La lingua parlata si articola in diversi generi e sottogeneri, ognuno caratterizzato da un proprio corredo di regole, convenzioni e caratteristiche. Molte di queste si differenziano da quelle tipiche della lingua scritta e sembra sensato giudicarle piuttosto in base a esse stesse, non solamente in relazione allo scritto. Benché nel presente contributo si sia partiti da analisi della presenza delle nominalizzazioni nella lingua scritta, si è cercato di esplorare i vari generi del parlato senza paragoni esterni.

I dati raccolti con il corpus GOS confermano la nostra ipotesi iniziale solo in parte: la nominalizzazione è nettamente più frequente nei generi testuali parlati più formali rispetto a quelli meno formali, sia se si considerano diverse tipologie testuali, sia guardando il tipo di canale usato nell’evento comunicativo. Si è però dimostrato che il fenomeno non è del tutto marginale nella lingua slovena parlata in generale: nel corpus GOS sono state trovate circa 7.000 nominalizzazioni su un milione di parole; tenendo conto che nella ricerca non sono state incluse tutte le nominalizzazioni, bensì solo quelle con una delle sette desinenze analizzate, il numero reale sarebbe senza dubbio leggermente più alto. La percentuale, come abbiamo visto, in alcuni generi testuali sale considerevolmente, in altri si abbassa anche fino a poche centinaia di occorrenze, ma la nominalizzazione non è mai del tutto assente. Si può supporre che i tipi testuali in cui la nominalizzazione è più ricorrente siano anche quelli più influenzati dalle regole e dalle usanze dello scritto, poiché spesso si tratta di testi parlati almeno in parte preparati in anticipo in forma scritta (per esempio le trasmissioni televisive o radiofoniche, ma anche alcuni tipi di contatti personali come le riunioni formali e simili), tuttavia per una conferma conclusiva si renderebbe necessaria una ricerca mirata e approfondita. Indipendentemente da ciò, però, non sembra possibile sostenere che l’influenza dello scritto si estenda a tutte le nominalizzazioni riscontrate, poiché se ne trovano anche nei tipi di discorso più spontanei e meno controllati, come le conversazioni private.

Molte sono le questioni che rimangono aperte: prima fra tutte, come si è accennato nella sezione precedente, rimarrebbe da svolgere un’analisi qualitativa approfondita della densità lessicale nei vari generi del parlato, la quale potrebbe essere un criterio pertinente nella determinazione delle caratteristiche peculiari dei singoli tipi di discorso parlato.

## Fonte primaria

GOS – Referenčni govorni korpus slovenskega jezika. 30.06.2012. <http://www.korpus-gos.net/>.

## Fonti secondarie

- BIBER, Douglas/Susan CONRAD/Randi REPPEN (1998) *Corpus Linguistics. Investigating Language Structure and Use*. Cambridge: Cambridge University Press.
- CHAFE, Wallace (1982) "Integration and Involvement in Speaking, Writing and Oral Literature." In: D. Tannen (a cura di), 35–53.
- HALLIDAY, Michael A. K./James R. MARTIN (1993) *Writing Science. Literacy and Discursive Power*. Pittsburgh: University of Pittsburgh Press.
- HALLIDAY, Michael A. K./Christian M. I. M. MATTHIESSEN (2004) *An Introduction to Functional Grammar. Third Edition*. London: Arnold.
- KLINAR, Stanko (1996) "Samostalnikost angleščine v primeri s slovenščino." In: S. Klinar (a cura di), 149–193.
- KLINAR, Stanko (a cura di) (1996) *K tehniki prevajanja iz slovenščine v angleščino*. Radovljica: Didakta.
- KRANJC, Simona (a cura di) (2011) *Meddisciplinarnost v slovenistiki. Obdobja 30*. Ljubljana: Znanstvena založba Filozofske fakultete.
- LEECH, Geoffrey (a cura di) (1995) *Spoken English on Computer: Transcription, Markup and Application*. New York: Longman Publishing.
- MIKOLIČ JUŽNIČ, Tamara (2007) *Nominalne strukture v italijanščini in slovenščini: pogostnost, tipi, in prevodne ustreznice*. Tesi di dottorato di ricerca. Ljubljana: Università di Ljubljana.
- MIKOLIČ JUŽNIČ, Tamara (2011) "Vpliv besedilnih tipov na pojavljanje nominalizacije v slovenščini: korpusna raziskava." In: S. Kranjc (a cura di), 321–327.
- PLEMENITAŠ, Katja (2007) *Posamostaljenja v angleščini in slovenščini: primer časopisnih vesti in kritik*. Maribor: Slavistično društvo.
- POGORELEC, Breda (1965) "Vprašanja govornega jezika." In: F. Vurnik (a cura di), 132–182.
- SINCLAIR, John (1995) "From Theory to Practice." In: G. Leech (a cura di), 99–112.
- Slovar slovenskega knjižnega jezika* (1994) Ljubljana: DZS.
- SMOLEJ, Mojca (2012) *Besedilne vrste v spontanem govoru*. Ljubljana: Znanstvena založba Filozofske fakultete.
- STABEJ, Marko (a cura di) (2005) *Večkulturnost v slovenskem jeziku, literaturi in kulturi: zbornik predavanj/41. seminar slovenskega jezika, literature in kulture*. Ljubljana: Center za slovenščino kot drugi/tuji jezik pri Oddelku za slovenistiko Filozofske fakultete.
- TANNEH, Deborah (a cura di) (1982) *Spoken and Written Language: Exploring Orality and Literacy*. Norwood: Ablex.
- TOGNINI-BONELLI, Elena (2001) *Corpus Linguistics at Work*. Amsterdam/Philadelphia: John Benjamins.

- VERDONIK, Darinka (2007) *Jezikovni elementi spontanosti v pogovoru: diskurzni označevalci in popravljanja*. Maribor: Slavistično društvo.
- VERDONIK, Darinka/Ana ZWITTER VITEZ (2011) *Slovenski govorni korpus*. Ljubljana: Trojina, zavod za uporabno slovenistiko.
- VEROVNIK, Tina (2010) *Radijski novinarski dvogovorni žanri kot okvir jezikovnih izbir novinarjev*. Ljubljana: Università di Ljubljana.
- VIDOVIČ MUHA, Ada (a cura di) (1996) *Jezik in čas*. Ljubljana: Znanstveni inštitut Filozofske fakultete.
- VURNIK, France (a cura di) (1965) *Jezikovni pogovori*. Ljubljana: Cankarjeva založba.
- ZEMLJARIČ MIKLAVČIČ, Jana (2008) *Govorni korpusi*. Ljubljana: Znanstvena založba Filozofske Fakultete.
- ZWITTER VITEZ, Ana (2005) "Prozodija v spontanem govoru." In: M. Stabej (a cura di), 160–169.
- ŽELE, Andreja (1996) "Razvoj posamostaljenja v slovenskem publicističnem jeziku med 1946 in 1995." In: A. Vidovič Muha (a cura di), 191–200.

#### Riassunto

#### LA NOMINALIZZAZIONE COME INDICATORE DEL GRADO DI FORMALITÀ IN ALCUNI TIPI TESTUALI DELLA LINGUA PARLATA

Il presente contributo si pone come oggetto una ricerca su alcune caratteristiche del discorso parlato legate alla complessità sintattica e lessicale determinata dall'uso della nominalizzazione. Partendo dalla considerazione, più volte espressa, su come la lingua parlata differisca da quella scritta anche dal punto di vista della complessità sintattica e lessicale, dalla quale può dipendere il grado di formalità, e basandoci su studi antecedenti sulla presenza di un fenomeno sintattico-lessicale relativamente complesso, quale è la nominalizzazione, in alcuni generi della lingua slovena scritta, si intende verificare se e in quali aspetti i vari tipi di testi della lingua parlata differiscano tra loro sotto questo punto di vista. La nominalizzazione, qui intesa come metafora grammaticale nel senso hallidayano, secondo vari studiosi è, infatti, generalmente un elemento caratteristico della lingua formale scritta, la quale trova minore spazio nei generi parlati, sebbene non sia completamente assente in alcuni suoi generi formali. Per la raccolta dei dati necessari ai fini della ricerca viene usato il corpus della lingua parlata slovena GOS, in cui i testi compresi sono divisi in vari subcorpora secondo criteri predefiniti riguardanti la tipologia testuale, il canale dell'evento comunicativo, il tipo di evento comunicativo, la regione di provenienza dei parlanti ecc. La metodologia usata è quella della linguistica dei corpora, integrata, in parte, con ipotesi basate soprattutto sul quadro teorico della linguistica sistemico funzionale.

**Parole chiave:** lingua parlata, linguistica dei corpora, metafora grammaticale, nominalizzazione, registro.

Povzetek  
NOMINALIZACIJA KOT POKAZATELJ STOPNJE FORMALNOSTI  
V NEKATERIH BESEDILNIH TIPIH GOVORA

Prispevek je osredotočen na raziskavo o nekaterih značilnostih govornega diskurza, povezanih s sintaktično-leksikalno kompleksnostjo, ki se pojavlja zaradi uporabe nominalizacij. Če izhajamo iz večkrat izražene trditve, da se govorni diskurz razlikuje od pisnega (tudi) s stališča sintaktične in leksikalne kompleksnosti, od katere je lahko odvisna stopnja formalnosti, in upoštevamo predhodne študije o prisotnosti nominalizacije kot relativno kompleksnega sintaktično-leksikalnega fenomena v nekaterih besedilnih vrstah slovenskega pisnega jezika, želimo preveriti, če in v katerih pogledih se razni besedilni tipi govornega jezika s tega vidika razlikujejo med sabo. Nominalizacija, ki jo razumemo kot slovnično metaforo v hallidayevskem smislu, je namreč po mnenju številnih znanstvenikov večinoma prvina, tipična za formalni pisni jezik, v govornih vrstah diskurza pa je manj prisotna, čeprav jo je mogoče zaslediti v nekaterih relativno formalnih besedilnih vrstah. Potrebne podatke smo pridobili s pomočjo korpusa slovenskega govornega jezika GOS, v katerem so besedila razdeljena v različne podkorpuse v skladu s predhodno določenimi kriteriji o besedilnih tipih, kanalih govornega dogodka, tipu govornega dogodka, regiji pripadnosti govorcev itd. Uporabljena metodologija izhaja iz korpusnega jezikoslovja, delno pa je podkrepljena tudi s hipotezami, ki slonijo zlasti na teoretičnem okviru sistemsko funkcijske slovnice.

**Ključne besede:** govorni jezik, korpusno jezikoslovje, nominalizacija, slovnična metafora, register.