

# Sequenze N+pN (nome comune + nome proprio): descrizione linguistica da un corpus dell'italiano.

Felicia Logozzo

Università di Roma 'Tor Vergata'

via Columbia, 1 – 00133 Roma

logozzo@lettere.uniroma2.it

## Abstract

**English.** This paper describes the most important N+pN (noun + proper noun) structures in Italian from the corpus of La Repubblica 2002-2005

**Italiano.** *Il contributo descrive le strutture N+pN più significative estratte dal corpus de La Repubblica 2002-2005.*

## 1 Introduzione

I dati oggetto dell'analisi sono stati ricavati dal corpus de *La Repubblica* 2002-2005 (circa 20 milioni di token), mediante T2K<sup>1</sup> (*Text to Knowledge*), sistema automatico per l'estrazione e l'organizzazione della conoscenza da testi.

Al software è stato richiesto di selezionare le sequenze N+N (nome comune + nome comune) e N+pN (nome comune + nome proprio) con frequenza pari o superiore a 18 occorrenze.

Dei 557 risultati estratti (24.013 occorrenze complessive) ne sono stati considerati validi 437 (78%<sup>2</sup> circa); i restanti 120 sono prevalentemente sequenze N+A (nome + aggettivo) o A+N tra cui composti nominali con aggettivo<sup>3</sup>, sia con testa a sinistra (*maglia rosa, giudice istruttore, gruppo dirigente, piano regolatore* ecc.), sia con testa a destra (*sostituto procuratore* ecc.); tra questi ultimi, anche *anni Sessanta, anni Novanta, anni Trenta, anni Venti, anni Quaranta* - classificati

dal software come sequenze con nomi propri, per l'uso convenzionale della maiuscola.

Il gruppo delle 437 sequenze potenzialmente utili ai fini dell'analisi N+N/N+pN si riduce ulteriormente, se si eliminano le 14 sequenze non complete (*calo rispetto, tv via, pistola calibro, via G, via de'*), le 2 locuzioni avverbiali (*passo passo, fin fine*) e i 41 prestiti stranieri ed espressioni latine (*talk show, pole position, week end, par condicio, alter ego, call center, best seller, tour operator* ecc.).

Escludendo infine un appellativo - *signor Presidente* - e la parte di polirematica - *parola fine* -, l'insieme delle 380 sequenze rimanenti è costituito da 147 composti N+N e da 231 sequenze N+pN, oggetto, queste ultime, di analisi specifica nel prossimo paragrafo.

Quanto al primo gruppo, sono rappresentati nel corpus composti N+N<sup>4</sup> di tutte le categorie: coordinati endocentrici (*decreto legge*) ed esocentrici (*centro sinistra*), subordinati (*ufficio stampa*). Tra questi, ben 53 sequenze su 147 sono rappresentate da espressioni di spazio o di tempo, formate con pochi e diffusissimi elementi base (*centro, metà, mattina, pomeriggio, sera, inizio, fine, punti cardinali*<sup>5</sup>).

Sebbene l'italiano non esprima tradizionalmente rapporti di subordinazione mediante giustapposizione, ma prediliga sintagmi preposizionali e

<sup>1</sup>(Dell'Orletta et al., 2014).

<sup>2</sup>Sul corpus in questione, l'accuratezza del *Part of Speech Tagger* impiegato risulta essere complessivamente pari al 96,34% (cfr. Dell'Orletta, 2009).

<sup>3</sup>Per la classificazione dei composti, là dove non diversamente specificato, si fa riferimento a Bisetto (2004: 33-50).

<sup>4</sup>Per i composti N+N dell'italiano, cfr. i lavori di Jan Radimský, in particolare Radimský (2013), sul corpus de *La Repubblica*, e Radimský (2015).

<sup>5</sup>Sono esempi prototipici di basi di composti grounding: 'Le teste dei composti grounding in italiano esprimono concetti intrinsecamente relazionali, che richiedono una specificazione ulteriore per convogliare un contenuto efficacemente informativo; esempi tipici di teste grounding sono i portatori/contenitori di informazione, organizzazioni, luoghi, aggregatori, puntatori nel tempo e nello spazio, proprietà misurabili' (Baroni, Guevara e Pirrelli, 2009)

aggettivi relazionali, la composizione N+N è, com'è noto, molto più diffusa di quanto appaia dalle grammatiche descrittive, che non prevedono, ad esempio, la categoria del nome attributivo, impiegata nella descrizione della lingua inglese, là dove convivono aggettivi denominali e nomi giustapposti corradicali con medesima funzione (*wool swater / woolen sweater*)<sup>6</sup>.

Di contro, in alcuni dizionari<sup>7</sup>, per superare le difficoltà descrittive delle strutture N+N, basi nominali di composti, prevalentemente ma non esclusivamente coordinati, vengono registrati come aggettivi invariabili (*parola chiave, punti chiave, cifra record, tempo record, motore diesel, rabbino capo, ruolo chiave, città campione, uomo chiave, campo base, livelli record*).

## 2 Denominazione e composizione N+pN

Il rapporto prevalente per le sequenze N+pN è senza dubbio quello identificativo/appositivo (129 su 231 sequenze, 5451 occorrenze su 9430). La lingua italiana prevede infatti che ad un nome proprio possa/debba essere affiancata la propria categoria di appartenenza mediante apposizione.

*N+pN*

I casi prototipici di identificazione-apposizione sono rappresentati da sequenze titolo+pN (*presidente Bush, presidente Ciampi, ministro Tremonti, sindaco Veltroni, ministro Castelli* ecc.) o nomi societari (*società Autostrade*) in cui

*pN è un N*

come confermano usi quali: *'Autostrade è una società privatizzata negli anni '90 che opera in un regime di concessione statale'* (il giornale.it).

A differenza degli esempi di cui sopra, in molti casi l'apposizione è indispensabile e fa parte integrante dell'espressione 'nome proprio': è il caso, ad esempio, degli odonimi<sup>8</sup> che si compongono obbligatoriamente con struttura N+pN (*via Cristoforo Colombo ≠ Cristoforo Colombo*). Se col tempo una specifica strada riesce ad acquisire una sufficiente notorietà, potrà essere identificata anche soltanto col nome proprio preceduto dall'articolo, segno – quest'ultimo - della pre-

senza silenziosa dell'omesso nome comune (*la Colombo*).

Gli odonimi infatti possono essere inclusi nella più generale categoria per cui vale l'espressione:

*N intitolato a pN*

*premio Nobel, via Nazionale, palazzo Chigi, viale Mazzini, piazza Vittorio, piazza Navona, via Roma, corso Vittorio, teatro Politeama, palazzo Koch.*

In sequenze di questo tipo, i nomi propri sono, per così dire, nomi di riuso; la relazione N+pN continua ad essere una relazione identificativo-appositiva, ma inizia ad acquisire aspetti più complessi che la proiettano verso la subordinazione.

Le vicende specifiche di ognuna di queste sequenze fa sì che sia sentita accettabile o meno l'equivalenza *pN è un N*. Tutti d'accordo, per esempio, che *il Nobel* sia un premio e che *il Politeama* sia un teatro; più difficile che *via Nazionale* o *via Roma* viaggino senza nome comune:

*via Nazionale è una via / Piazza Venezia è una piazza / palazzo Chigi è un palazzo*

*\*/\* la Nazionale è una via / \*(la) Venezia è una piazza / \*il Chigi è un palazzo*

In questi casi, la motivazione che lega N a pN, sfugge ad un'analisi esclusivamente linguistica. Se si considera ad esempio *palazzo Chigi*, in rapporto ad altre sequenze dalla medesima struttura, come *palazzo Benetton* e *palazzo Fuksas*, emerge che solo conoscenze extralinguistiche permettono di sapere che i Chigi comprarono e abitarono l'omonimo palazzo nel 600; *palazzo Benetton* è stato progettato da Fuksas su commissione appunto di Benetton, e poi rivenduto a H&M; *palazzo Fuksas* è stato progettato anch'esso da Fuksas, a Pescara, e ospita la sede di una importante multinazionale.

Si tratta sostanzialmente di strutture analizzabili solo a livello pragmatico<sup>9</sup>: col passare del tempo, la relazione di motivazione che lega N+pN (il fatto di aver comprato, commissionato o progettato il palazzo) svanisce nella coscienza dei parlanti e l'intera sequenza N+pN diventa denominazione. Il processo è pressoché concluso nel caso di *palazzo Chigi*, sentita quasi solo come denominazione; appena iniziato per *palazzo Fuksas*.

A dispetto di quanto dichiarato dalle grammatiche, è cristallizzata nella lingua dell'uso la giustapposizione N+pN per l'espressione di relazio-

<sup>6</sup> Un'alternanza parzialmente ma non completamente sovrapponibile – non si tratta infatti di corradicali in senso stretto - si verifica in italiano in presenza di Scritture Brevi come in *tv/televisivo - programma tv / programma televisivo* (già oggetto di analisi in Logozzo, in stampa).

<sup>7</sup> Sabatini-Coletti, per esempio.

<sup>8</sup> Per denominazione e composizione con toponimi e marchionimi cfr. Logozzo (2015).

<sup>9</sup> Cfr. Bartning (2001: 150 ss).

ni morfosintattiche diverse da quella identificativa-appositiva.

Il corpus ci restituisce in tal senso 102 sequenze che coprono il 44% del totale delle sequenze N+pN.

Rendere conto accuratamente delle strutture soggiacenti a ciascuna sequenza è questione complessa; per questo, le sequenze estratte dal corpus verranno classificate mediante criteri volutamente inclusivi, volti ad individuare analogie più che a differenziare, al solo fine dell'analisi del ruolo dei nomi propri.

Il primo e più numeroso gruppo identificato (46 sequenze per 1475 occorrenze) è quello delle sequenze che rimandano in qualche modo ad un rapporto di appartenenza in senso lato, morale, sociale o figurato<sup>10</sup>:

*N appartiene a pN*

*commissario Ue, soldati Usa, coppa Italia, amministrazione Usa, ispettori Onu, segretario Ds, vertice Rai, truppe Usa, segretario ds, dati Istat, paesi Ue, canone Rai, economia Usa, coppa Uefa, comando Usa, presidente Ue, reti Mediaset, leader Cgil, deputato Ds, ambasciata Usa, capogruppo Ds, militari Usa, commissione Ue, esercito Usa, leader Ds, senatore Ds, presidenza Ue, ex dc, reti Rai, forze Usa, stabilimento Fiat, governo Usa, norme Ue, giornalista Rai, agenti Cia, presidente Rai, vertice Ue, operai Fiat, consiglio Rai, consigliere Rai, base Usa, vertici Fiat, segretario Cgil, classifica Fifa, presidente ds, segretario Udc.*

Andando ad osservare il rapporto semantico e sintattico che intercorre tra N e pN, appare evidente che si tratta di relazioni di subordinazione che potrebbero essere espresse in forma analitica mediante l'impiego di parole grammaticali – la preposizione 'di' primariamente – o di perifrasi verbali, riconducibili ai rapporti cosiddetti genitivi.

La mancanza della preposizione 'di' a collegare N e pN, fa sì che le strutture formalmente appaiano come composti. Nello specifico, si tratterà di composti determinativi che 'indicano una sottoclasse degli oggetti individuati dalla testa e nei

quali il nome di destra si pone come una specificazione restrittiva del nome che lo precede'.<sup>11</sup>

Sulla base dei dati emersi dal corpus, le strutture che consentono questo tipo di costruzioni sembrano essere piuttosto uniformemente costituite.

La composizione N+pN per l'espressione di relazioni di appartenenza concrete o figurate, è dunque ammessa in italiano seppure con restrizioni riguardanti, e la forma, e la sostanza del determinante: quanto alla sostanza, entrano in composizione nomi propri di enti, associazioni, istituzioni o marchionimi; quanto alla forma, sono rappresentati determinanti di un'unica parola (?/\* *deputato Unione di Centro*), meglio se in forma di sigla (*Scritture Brevi*)<sup>12</sup>.

Alla base della grande produttività della composizione con *Scritture Brevi* vi è una motivazione semantica (il valore referenziale delle sigle coinvolte che, in quanto nomi propri dotati di riferimento, riducono le possibili ambiguità interpretative della relazione N+pN) e una motivazione formale: i composti sono strutturalmente formati da due elementi; la sigla permette di ridurre un nome proprio complesso, eventualmente costituito da un'espressione con coordinazioni e subordinazioni al suo interno (*Figc = Federazione Italiana Giuoco calcio*), ad elemento unico, evitando così lunghe successioni asindetichiche di elementi (?/\* *Presidente Federazione Italiana Giuoco Calcio*)<sup>13</sup>.

Non sembrano ammesse giustapposizioni di nomi propri che abbiano generato nella lingua aggettivi relazionali denominativi - si pensi agli etnici derivati da nomi di stato/nazione - tranne nel caso in cui essi siano siglati; si ricrea in tal modo l'alternanza già vista per *tv/televisivo* (cfr. nota 6): *soldati USA / soldati statunitensi* ma non \**soldati Stati Uniti*.

La composizione N+pN del tipo appena analizzata può essere considerata produttiva perché si presta alla composizione occasionale, pratica originariamente non diffusa nella lingua italiana<sup>14</sup>, a differenza dell'inglese<sup>15</sup> che conia costantemente strutture pN+N.

E' importante, a questo proposito, distinguere l'imposizione di un nome proprio o di un'etichetta individuante (*Coppa Italia, via Marchesi*) dalla composizione occasionale (*presidente USA*). Se infatti nel primo caso le regole della

---

<sup>10</sup>What is normally called possession is the linguistic expression of the relation between two entities, a Possessor and a Possessum, such that one, the Possessor, is seen as being in some way related to the other, the Possessum, as having it near or controlling it. The kind of relation between the two can be of various sorts' (Herslund and Baron, 2001: 2).

---

<sup>11</sup>(Bisetto, 2004: 40).

<sup>12</sup>Cfr. Chiusaroli (2012).

<sup>13</sup> Cfr. Logozzo (in stampa).

<sup>14</sup>Dardano (1978: 143) parla di 'creazioni effimere'.

<sup>15</sup>Cfr. Bauer (1983: 202 ss).

grammatica si possono prescindere senza troppi imbarazzi e l'approccio di analisi può essere solo descrittivo, a posteriori, nel caso della composizione occasionale è possibile azzardare ipotesi predittive, assumendo che certe strutture saranno più probabilmente scelte, rispetto ad altre, sulla base dell'uso.

Merita un'attenzione particolare l'impiego degli antroponimi giustapposti in sequenze N+pN, per i quali il corpus ci restituisce i seguenti esempi: *governo Berlusconi, amministrazione Bush, legge Gasparri, legge Cirami, riforma Moratti, governo Sharon, lodo Schifani, legge Biagi, legge Bossi-Fini, governo Prodi, governo Blair, lodo Maccanico, governo Zapatero, riforma Dini, riforma Biagi, riforma Tremonti, governo Cuffaro, legge Boato*

Sulla base dei dati estratti, si evince che non sono ammessi antroponimi giustapposti ad esprimere vere e proprie relazioni di possesso<sup>16</sup> (\*auto Antonio, \*giacca Marcello).

Quando essi compaiono in sequenze N+pN, esprimono relazioni che oscillano dall'identificativo-appositivo al determinativo-specificativo e richiedono di spostare l'analisi dalla semantica degli elementi coinvolti, alla pragmatica del discorso.

E' infatti invalso l'uso di attribuire una denominazione a leggi e riforme, impiegando antroponimi come nomi propri di riuso. La riforma della scuola del 2003 è dunque 'intitolata' al Ministro (Moratti) che si è fatto promotore della stessa, così come si danno, per esempio, nomi ai governi (*Berlusconi 1, Berlusconi 4, Prodi 2*) in espressioni in cui la relazione N+pN può essere al contempo di subordinazione (*governo guidato da, nominato da, riforma ideata da, legge proposta da*) o di denominazione (*Il governo Berlusconi 3 è durato dal 23 aprile 2005 al 17 maggio 2006* – Wikipedia.it).

Come già detto a proposito degli odonimi di tipo *N intitolato a pN*, gli antroponimi non entrano bene in predicazione, nemmeno in presenza di un articolo che sottintenda l'omissione di un nome comune, per le evidenti possibili ambiguità: *la Bossi-Fini* è infatti una legge, ma *la Moratti* non è una riforma. Un elemento determinativo ulteriore di pN, che escluda il riferimento diretto ad un essere umano, consente il superamento del problema, come, ad esempio, un altro antroponimo in *Bossi-Fini*, un cardinale in *Berlusconi 3*. Al limite tra relazione identificativo-appositiva e specificativo-determinativa anche quella tra N e

pN in *gruppo Fiat, gruppo Parmalat, gruppo Telecom, gruppo Mediaset*, dal momento che, per esempio, Mediaset è una holding proprietaria di un gruppo di aziende, ma è anche l'etichetta dell'intero gruppo che prende il nome della capofila.

Denominate sulla base delle tematiche di cui si occupano, come testimonia la maiuscola dei determinanti che le pone nel gruppo N+pN e non in quello N+N, le *commissioni Giustizia, Lavoro, Bilancio e Affari (Costituzionali)* in cui però si mantiene un rapporto di subordinazione tra N e pN: \*(*La Giustizia è una commissione / Quella della giustizia è una commissione.*

Altre sequenze del corpus si possono raccogliere infine in una categoria aspecifica:

*si è verificato un N e riguarda pN*  
*processo Sme, caso Parmalat, crisi Fiat, caso Sofri, caso Fiat, caso Cirio, crac Parmalat, caso Moro, crac Cirio, caso Fazio, caso Catania, processo Imi-Sir, caso Sme, affare Sme*

Rientra in questo gruppo il tipo *caso+pN*, ispirato forse al modello francese dell'*affaire Dreyfus*<sup>17</sup>, tipo estremamente produttivo in cui è possibile inserire anche il sottotipo *crisi Fiat, crac Parmalat*. Sebbene rappresentato nel corpus da un'unica occorrenza (*omicidio Biagi*) è bene citare in questo gruppo anche i genitivi oggettivi con antroponimi, che sembrano entrare in composizione con specifiche teste (es: *omicidio, rapimento*) che permettono uno scioglimento del tipo: *si è verificato un omicidio e ha riguardato Biagi*, probabilmente nel momento in cui un certo episodio di cronaca è talmente di dominio pubblico da diventare un 'caso' quasi identificabile con il solo pN.

### 3 Conclusioni

La presenza dei nomi propri nelle sequenze N+pN fa sì che si abbia a che fare, nella maggior parte dei casi, con relazioni di denominazione con motivazione più o meno trasparente, analizzabili, ora ricorrendo alla semantica degli elementi coinvolti, ora, più frequentemente, ricorrendo alla conoscenza del mondo extralinguistico. Completamente estranee alla denominazione, sono invece le sequenze del tipo *commissario Ue, soldati USA* che rappresentano per l'italiano esempi di composizione N+pN, alternativa all'impiego di aggettivi denominali o sintagmi preposizionali.

<sup>16</sup>Cfr. Heine (2006).

<sup>17</sup>Per l'influsso di modelli alloglotti sulle costruzioni N+pN cfr. Klajn (1972: 180 ss).

## Bibliografia

- Baroni M., Guevara E. e Pirrelli V. 2009. *Sulla tipologia dei composti N+N in italiano: principi categoriali ed evidenza distribuzionale a confronto*, in Ferrari G., Mosca M. e Benatti R. (a cura di) *Linguistica e modelli tecnologici di ricerca*. Bulzoni, Roma, 73-95.
- Bartning I. 2001. *Towards a typology of French NP de NP structures or how much possession is there in complex noun phrases with de in French?*, in Baron I., Herslund M. and Sørensen F. (a cura di) *Dimensions of possessions*. Benjamins, Amsterdam – Philadelphia, 147-167.
- Bauer L. 1983. *English word formation*. Cambridge University Press, Cambridge.
- Bisetto A. 2004. *Composizione con elementi italiani*, in Grossmann M. e Reiner F. (a cura di) *La formazione delle parole in italiano*. Max Niemeyer, Tübingen, 33-50.
- Chiusaroli F. 2012. *Scritture brevi oggi: tra convenzione e sistema*, in Chiusaroli F. e Zanzotto F.M. (a cura di) *Scritture brevi di oggi*, Quaderni di Linguistica Zero, Napoli, 4-44.
- Dardano M. 1978. *La formazione delle parole nell'italiano di oggi*. Bulzoni, Roma.
- Dell'Orletta F., Venturi G., Cimino A. and Montemagni S. 2014. *T2K<sup>2</sup>: a System for Automatically Extracting and Organizing Knowledge from Texts* in *Proceedings of 9th Edition of International Conference on Language Resources and Evaluation (LREC 2014)*, 26-31 May, Reykjavik, Iceland, 2062-2070 [http://www.lrec-conf.org/proceedings/lrec2014/pdf/590\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2014/pdf/590_Paper.pdf)
- Dell'Orletta F. 2009. *Ensemble system for Part-of-Speech tagging*, in *Evaluation of NLP and Speech Tools for Italian*. Reggio Emilia. [http://www.evalita.it/sites/evalita.fbkeu/files/proceedings2009/PosTagging/POS\\_ILC.pdf](http://www.evalita.it/sites/evalita.fbkeu/files/proceedings2009/PosTagging/POS_ILC.pdf)
- Heine B. 2006. *Possession*. Cambridge University Press, Cambridge 2006.
- Herslund M. and Baron I. 2001. *Introduction: Dimensions of possession*, in Baron I. Herslund M., Sørensen F. (a cura di), *Dimensions of possessions*. Benjamins, Amsterdam – Philadelphia, 1-25.
- Klajn I. 1972. *Influssi inglesi nella lingua italiana*, Olschki, Firenze.
- Logozzo F. 2015. *Sull'apparente omissione della preposizione 'di' davanti a nomi propri, con particolare riguardo a toponimi e marchionimi*. in *Rivista italiana di onomastica* 21 (1): 93-116.
- Logozzo F. (in stampa). *Scritture brevi tra sintagmi nominali, preposizionali e aggettivali*. Atti del convegno 'Scritture brevi e varietà dialettale', Procida 26-28 giugno 2014.
- Radimský J. 2013. *Tight N-N compounds in the Italian la Repubblica corpus*, in Baptista, J., Monteleone, M. (a cura di) *Actes du 32ème Colloque international sur le lexique et la grammaire (10-14 septembre 2013, Faro, Portugal)*, 41-47.
- Radimský J. 2015. *Les composés de coordination en italien: esquisse d'une typologie*, in *Studia Romanica Posnaniensia* 42 (1): 97-111.