

# L'analisi corpus-driven delle strutture *essere* + *Preposizione* in italiano. Costrutti grammaticali e variazione marcata del predicato

Alessandro Panunzi  
Università di Firenze  
Università di Torino  
alessandro.panunzi@unifi.it

## Abstract

Structures with shallow patterning "*essere* + Preposition" are used within a wide range of constructs with different meaning, which are often not well-distinguished within the lexicographic tradition. From the analysis on corpora which are representative of both written and spoken language, it emerges that several uses corresponding to the configuration "*essere* + Preposition" take place within a relevant distinction between periphrastic constructs, which carry grammatical values, and a variation of marked, secondary and idiosyncratic uses. Such a distinction is reflected by a different quantitative distribution of the two types, which is consistent across the written-spoken variation of language use, and therefore it can be considered as a proper feature of Italian. The *corpus driven* analysis of *essere* + Preposition shows several periphrastic constructs with grammatical values related to the expression of modal, temporal and aspectual traits.

## 1 Introduzione

Il lemma *essere* è la voce verbale a più alta frequenza dell'italiano. Nell'arco dei suoi usi si riscontrano variazioni molto forti di valore, in particolare tra tipi di uso *grammaticale* e variazioni nell'interpretazione lessicale del verbo, come è proprio dei verbi predicativi. Questo studio focalizza su una porzione significativa di variazione d'uso di *essere* di particolare rilievo lessicografico, ovvero le configurazioni *essere* + preposizione, il cui trattamento non è uniforme all'interno della pratica di redazione dei dizionari.

A partire da un'analisi su corpora rappresentativi dell'uso sia scritto che orale (corpus NUNC e corpus C-Oral-Rom italiano, cfr. i paragrafi successivi)<sup>1</sup> evidenzieremo, all'interno della configurazione sintagmatica *essere* + Preposizione, la presenza di una distinzione significativa a livello lessicografico tra costrutti perifrastici con valore grammaticale (propri della lingua italiana) e variazioni d'uso marcate, in quanto tali idiosincratice. Tale distinzione si riflette in una diversa distribuzione dei due tipi nei corpora sia scritti che orali. Gli usi

---

<sup>1</sup> Entrambi i corpora sono lemmatizzati ed etichettati sulle parti del discorso. Tale livello di annotazione rende possibile ricerche che riguardano le configurazioni lessico-grammaticali specifiche.

grammaticali e quelli lessicali marcati hanno frequenze quantitativamente simili (intorno al 10% del totale delle strutture **essere** + Preposizione), ma distribuite in modo diverso nei rispettivi tipi d'uso. Negli usi grammaticali si notano infatti addensamenti di frequenza su ogni singolo uso, che testimoniano la produttività della struttura, mentre gli usi lessicali marcati mancano di addensamenti quantitativi sui singoli usi.

L'analisi *corpus-driven* sugli usi perifrastici con valore grammaticale permettono inoltre di indurre generalizzazioni sul valore semantico dei costrutti individuati rispetto ai tratti grammaticali modificati.

### 1.1 Struttura della variazione: dizionari e corpus

I valori di uso di **essere** sono distinti, sia nella tradizione lessicografica che grammaticale, in (a) usi *verbali*, ulteriormente distinti al loro interno in usi *copulari* (base nominale della predicazione); (b) usi *predicativi* (soprattutto identità e localizzazione); e usi *ausiliari* (di tempo e di diatesi). Tale variazione è stata analizzata a partire dai dati del corpus di italiano parlato C-ORAL-ROM (300.000 tokens totali; Cresti, Moneglia 2005), in cui si evidenziano 13.995 occorrenze del verbo **essere**, distribuite in 10.819 enunciati (sui 40.791 del totale).<sup>2</sup> In Panunzi (2005) è stato evidenziato che solo nel 15% circa dei casi il verbo **essere** ha il valore di ausiliare, e quindi di marca grammaticale dei tratti di tempo o diatesi, legata all'economia del sistema flessivo del verbo italiano, a base analitica. In questo settore d'uso, il verbo **essere** ha una funzione puramente grammaticale.

**Essere** in funzione di *copula* occorre invece nel 40% dei casi totali del corpus. L'istanza tipica di tale uso prevede che il "complemento" verbale sia un aggettivo o un sintagma nominale (tipicamente non definito) che possa funzionare da predicato.

Tra gli usi del verbo **essere** che legano due sintagmi nominali c'è anche quello di *predicato di identità*, che si distingue dagli usi copulari per la struttura logica della predicazione, in quanto il complemento del verbo ha esso stesso un riferimento individuale, e non predicativo (*quello è Carlo*). Dallo spoglio dei dati emerge che tale variazione incide per circa il 12% della totale.

Una buona parte degli usi di **essere** (poco meno del 10% del totale) è costituita dalle istanze predicative con valore di locazione (spaziale o temporale), tipicamente in corrispondenza del riempimento del suo complemento da parte di un sintagma preposizionale.

Un altro settore di variazione d'uso particolarmente consistente è quello che riguarda il verbo **esserci**, il cosiddetto *predicato di esistenza*, costituisce da solo circa il 15% delle occorrenze di **essere** nel nostro corpus di riferimento. **Esserci** viene spesso considerato nelle grammatiche (cfr. Serianni 1988; Renzi 1988) e nei dizionari (Devot-Oli 1995) come variante dell'uso locativo di **essere** (cfr. l'alternativa proposta in Berruto 1986),<sup>3</sup> mentre in De

<sup>2</sup> Tali misure quantitative sottolineano la centralità dell'uso di **essere** nella lingua italiana da due punti di vista: (a) il verbo **essere** è il secondo lemma per rango d'uso nella lingua italiana parlata (preceduto unicamente dall'articolo determinativo 'il'); (b) il verbo **essere** è presente in 1 enunciato su 4 del parlato spontaneo.

<sup>3</sup> Berruto riconosce il valore presentativo di **esserci** come autonomo, in aggiunta a quello esistenziale, a riprova della specificità anche lessicale di tale costrutto.

Mauro (2000) viene considerato come un'entrata lessicale autonoma, e specificamente un lemma verbale procomplementare.

Escludendo quest'ultima porzione di variazione (che potrebbe essere ascritta alla variazione predicativa del lemma autonomo **esserci**), si può affermare che la configurazione sintagmatica **essere** + Preposizione è presente nel 9% dei casi in cui il lemma occorre.

### 1.2 La variazione di *essere* + *preposizione*

Gli usi verbali individuati dalla tradizione grammaticale e lessicografica come "predicati" sono collegati a contesti in cui il complemento del verbo **essere** è costituito da un sintagma preposizionale. Tale configurazione sintagmatica si lega prevalentemente agli usi di localizzazione spaziale-temporale come negli esempi:

- (1) (a) è in camera sua  
(b) è sul tavolo
- (2) (a) è a gennaio  
(b) è nel 1983

Tuttavia, il valore di **essere** nei suoi usi in correlazione una preposizione non è sempre di tipo predicativo-locativo. L'analisi condotta sul corpus ha individuato la presenza cospicua di **essere** + Preposizione in costruzioni lessicalizzate (metaforiche e polirematiche) e in grammaticalizzazioni con valori diversi.

Tali usi coinvolgono oltre il 20% delle occorrenze totali della configurazione **essere** + Preposizione, e risultano di particolare interesse lessicografico, in quanto parimenti distribuite sia nella lingua scritta che in quella parlata, e con valori simili: sono quindi strutture specifiche della lingua italiana d'uso.

## 2 Analisi del corpus

All'interno del corpus C-ORAL-ROM sono state individuate 1059 occorrenze del pattern **essere** + Preposizione. Tali occorrenze sono state analizzate singolarmente al fine di individuare il loro valore in contesto. A integrazione della base di dati, sono state individuate e classificate ulteriori 1086 occorrenze del verbo seguito da Preposizione estratte dal corpus scritto NUNC (costituito da testi di newsgroups), realizzato presso l'Università deli Studi di Torino e interrogabile in rete (<http://www.corpora.unito.it/>).

Nell'analisi delle occorrenze, sono stati identificati i casi in cui il verbo **essere** dà luogo a un costrutto, distinguendo

- forme complesse o polirematiche (Voghera 2004), tra le quali costrutti in cui **essere** funge da verbo supporto;
- perifrasi con funzione grammaticale.

### 2.1 Strutture lessicalizzate: *variazione marcata di essere* + *SP*

Si distinguono diverse tipologie di usi in cui il verbo **essere** seguito da Preposizione può essere considerato un lessema complesso o polirematico. In tali tipologie si ripropone la di-

stinzione tra usi copulari e usi predicativi di localizzazione.<sup>4</sup> Nel primo caso, è il sintagma preposizionale che si colloca alla destra di “essere” ad avere valore polirematico, come negli esempi:

- (2) (a) è a piedi  
(b) è a buon mercato

In questi contesti, il Sintagma Preposizionale (SP) assume valore aggettivale, mentre il verbo **essere** ricopre il ruolo di copula, e in quanto tale non fa strettamente parte del lessema complesso (es. *essere [alla mano]*; *essere [in gamba]*). La produttività di questa struttura da una parte, e la possibilità di sostituzione dell’SP con aggettivi monotematici dall’altra, dimostrano che non siamo di fronte ad usi particolari del verbo **essere**, quanto ad associazioni tra la copula e SP polirematici con valore di aggettivo. Tuttavia negli studi lessico-grammaticali (Vietri 1996) e nella pratica lessicografica questi usi non sono sistematicamente distinti dagli usi di **essere** all’interno di lessemi complessi (cfr. la relativa lista di polirematiche nel dizionario De Mauro 2000).

Nel secondo caso, invece, l’uso di **essere** + SP è propriamente quello di un Sintagma Verbale polirematico, in quanto rimanda a una localizzazione metaforica o a un uso stereotipo, in cui il senso dell’espressione è costruito a partire dal valore locativo del predicato, come negli esempi:

- (3) (a) è alla deriva  
(b) è in giro

Significativamente, in questo settore di variazione non vengono rilevati addensamenti d’uso. È questo il caso più specifico della variazione d’uso marcata (cfr. Panunzi, Moneglia 2004).

Un terzo gruppo di esempi, che si inseriscono invece nei processi sistematici di costruzione del lessico, ed è costituito dai casi nei quali il verbo **essere** si dimostra un verbo supporto (cfr. Gross 1998, Marello 2005), così come avviene per il lessico verbale generale a più alta frequenza cfr. **fare**). In questi casi la struttura analitica **essere** + SP risulta sostituibile con un verbo sintetico in corrispondenza semantica con il nome presente all’interno del sintagma preposizionale:

- (4) (a) è a conoscenza (sa, conosce)  
(b) è alla ricerca (cerca, ricerca)  
(c) è in evoluzione (si evolve)  
(d) in espansione (si espande)

La funzione di **essere** come verbo supporto è produttiva in quanto strategia di costruzione lessico-grammaticale, senza evidenziare addensamenti d’uso su selezioni nominali specifiche.

La tabella 1 evidenzia i dati di occorrenza delle tipologie descritte (valori assoluti e percentuali rispetto alle occorrenze di **essere** + Preposizione):

---

<sup>4</sup> Tale suddivisione non è sempre tenuta in considerazione nella redazione dei dizionari: si noti come tipologie d’uso simili siano descritte con difformità nel dizionario Devoto Oli (all’interno dell’inventario degli usi copulari) e in quello De Mauro (all’interno della descrizione degli usi predicativi).

	C-ORAL-ROM		NUNC	
ATTRIBUTIVO (SP aggettivale)	14	1,32%	24	2,21%
LOCATIVO (SP come luogo figurato)	48	4,53%	73	6,72%
VERBO SUPPORTO	20	1,89%	36	3,31%
ESSERE d'accordo <sup>5</sup>	30	2,83%	0	0,00%

Tabella 1. Occorrenze delle tipologie di variazione marcata *essere* + Preposizione

2.2 Strutture grammaticalizzate: perifrasi [ESSERE + Prep + infinito]

La tabella 2 riporta le costruzioni perifrastiche a struttura “*essere* + Preposizione + infinito” riscontrate nei corpora. Significativamente, per ognuno di questi usi diviene possibile identificare nelle occorrenze del corpus un valore semantico specifico (riportato nella colonna “valore” nella tabella).

tipo	valore	Esempi
PERIFRASI con VALORE MODALE	1. [ESSERE DA] + infinito valore modale “deontico”	<i>è da fare/discutere</i>
	sostituibile con “dovere” (+ infinito passivo), “bisogna”	<i>deve essere fatto/discusso</i> <i>bisogna fare/discutere</i>
	2. [ESSERE IN GRADO DI] + infinito valore modale “aletico”;	<i>è in grado di aiutare/salvare</i>
	sostituibile con “potere” (possibilità in termini fattuali)	<i>Poter aiutare/salvare</i>
PERIFRASI con valore TEMPORALE - ASPETTUALE	3. [ESSERE A] + infinito valore “durativo”	<i>è a giocare</i>
	simile alla costruzione con [STARE + gerundio]	<i>sta giocando</i>
	4. [ESSERE PER] + infinito 1 valore “incoattivo” di un evento imminente (evento con aspetto puntuale)	<i>era per cascare</i>
PERIFRASI CON DIVERSO VALORE	5. [ESSERE PER] + infinito 2 valore strumentale del soggetto rispetto all’evento denotato dall’infinito	<i>è per fare la cioccolata</i>
	sostituibile con [SERVE A + infinito]	<i>Serve a fare la cioccolata</i>

Tabella 2. Costruzioni perifrastiche *essere* + Preposizione + infinito.

Strutture di questo tipo non sono propriamente ascrivibili alla variazione d’uso verbale di *essere*, in quanto sono portatrici di un valore grammaticale distintivo e introducono tratti se-

<sup>5</sup> L’espressione complessa *essere d’accordo* è stata considerata autonomamente, in quanto formula dialogica specifica di parlato.

mantici modali o temporali-aspettuali che modificano il verbo. Nel dizionario Sabatini-Colletti 1997 alcune di queste costruzioni sono in effetti descritte separatamente dagli usi copulari e predicativi.

È da notare come, a differenza di quanto avviene nelle strutture a verbo supporto (strategia produttiva sul piano della creazione lessicale, che prevede selezione dell'argomento nominale), in questo caso è la sola preposizione a costituire espressione complessa insieme al verbo, e non l'intero sintagma preposizionale.

Dai dati della variazione d'uso nei corpora analizzati, riportati nella tabella 3, si nota che le strutture perifrastiche descritte sono produttive nell'uso linguistico sia scritto che parlato:

	C-ORAL-ROM		NUNC	
ESSERE DA + inf <sup>6</sup>	60	5,67%	48	4,42%
ESSERE IN GRADO DI + inf	23	2,17%	59	5,43%
ESSERE A + inf	29	2,74%	4	0,37%
ESSERE PER + inf (A) <sup>7</sup>	2	0,19%	0	0,00%
ESSERE PER + inf (B)	12	1,13%	2	0,18%

Tabella 3. Occorrenze delle strutture perifrastiche essere + Preposizione + infinito

Si può notare come, diversamente dai casi trattati nella variazione lessicale marcata, su ognuna delle strutture identificate da una preposizione si riscontrano addensamenti di frequenza che ne dimostrano la produttività.

### 3 Conclusioni

Dal lavoro parallelo su corpora scritti e orali è possibile apprezzare i dati di frequenza relativi alle occorrenze delle tipologie d'uso del verbo essere + Preposizione, distinguendo all'interno della variazione i costrutti con valore grammaticale propri dell'italiano e quelli con valore lessicale marcato. La tabella 4 riassume i dati provenienti dai due corpora.

	C-ORAL-ROM		NUNC		TOTALE	
TOTALE ESSERE + Prep	1059		1086		2145	
COSTRUTTI LESSICALI	82	7,74%	133	12,25%	215	10,02%
COSTRUTTI GRAMMATICALI	120	11,33%	113	10,41%	233	10,86%
ESSERE d'accordo	30	2,83%	0	0,00%	30	1,40%
TOTALE USI NON-PRIMARI	233	22,00%	246	22,65%	479	22,33%

Tabella 4

<sup>6</sup> Tale uso è possibile anche con "esserci": *c'è da + infinito*.

<sup>7</sup> Con la variante [essere lì lì per] + infinito: *ero lì lì per venderlo*.

La distribuzione dei costrutti **essere** + Preposizione presenta un numero di occorrenze molto simile tra costrutti lessicali marcati (215) e costrutti grammaticali (233), ma rispecchia due situazioni molto diverse. Nel caso delle costruzioni lessicalizzate, infatti, le occorrenze si distribuiscono in oltre 70 tipi di unità complesse: l'uso reale è distribuito, senza particolari addensamenti, in una variazione molto varia. Tale dato rispecchia la distribuzione riscontrata in Panunzi, Moneglia (2004) per i verbi generali a più alta frequenza nell'italiano parlato: all'interno della variazione d'uso marcata di un verbo gli usi metaforici e stereotipi sono dispersi.

Nel caso invece delle perifrasi con valore di grammaticalizzazione, un numero analogo di occorrenze totali si addensa intorno a 5 tipi (enumerati nella tabella 3). Ciò avviene a causa della produttività grammaticale dei costrutti, ognuno dei quali ha funzione semantico-grammaticale distinta, destinata all'espressione di specifici tratti modali e tempo-aspettuali.

## Bibliografia

### A. Dizionari

De Mauro, T. (2000), *Il dizionario della lingua italiana*. Milano, Paravia.

Devoto, G., Oli, G. C. (1995), *Dizionario della lingua italiana* (prima edizione 1967, seconda edizione 1982, terza edizione 1990). Firenze, Le Monnier.

Sabatini, F., Coletti, V. (1997). *DISC – Dizionario Italiano Sabatini Coletti*. Firenze, Giunti.

### B. Altri Testi

Corpora dell'Università di Torino in rete, <http://www.corpora.unito.it>

Gross, M. (1998), 'La fonction sémantique des verbes support'. *Travaux de linguistiques* 37, pp. 25-36.

Marello, C. (2005), 'Verbi supporto e insegnamento della lingua italiana a stranieri', in Korzen, J. (ed) *Atti dell'VIII convegno internazionale SILFI (2004) Lingua, cultura e intercultura: l'italiano e le altre lingue*, Copenhagen, Etudes Romanes, pp. 29-40.

Cresti, E., Moneglia, M. (eds) (2005), *C-ORAL-ROM. Integrated reference corpora for spoken romance languages*, vol. I° + DVD. Amsterdam, Benjamins.

Panunzi, A., Moneglia, M. (2004), 'La variazione primaria del verbo nel lessico dei corpora di parlato', in Albano Leoni, F., Cutugno, F., Pettorino, M., Savy, R. (eds) *Atti del convegno nazionale "Il Parlato Italiano"*, Napoli, D'Auria.

Panunzi, A. (2005), "'Essere" e "esserci" nella lingua italiana d'uso. Indagine su un corpus di parlato spontaneo e primi confronti interlinguistici nelle lingue romanze', in *Atti dell'VIII convegno internazionale SILFI (2004) Lingua, cultura e intercultura: l'italiano e le altre lingue*, Copenhagen, Etudes Romanes, pp. 255-266.

Vietri, S. (1996), 'The syntax of the Italian verb essere Prep', *Linguisticae Investigationes*, XX: 2.

Voghera, M. (2004), 'Polirematiche', in Grossmann, M., Rainer, F. (eds), *La Formazione delle Parole in Italiano*. Tübingen, Niemeyer, pp. 56-69.