



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

## Università degli Studi di Padova

Dipartimento di Studi Linguistici e Letterari

Corso di Laurea Magistrale in Linguistica  
Classe LM-39

Tesi di Laurea

# *I pattern collocazionali di “notizia”: ricognizione nei repertori e censimento di un corpus di riferimento*

Relatore  
Prof. Michele Cortelazzo

Laureanda  
Giovanna Fracasso  
n° matr. 1180197 / LMLIN

Anno Accademico 2022 / 2023



# Indice

<b>Introduzione</b> .....	3
<b>1 Panoramica degli approcci teorici alle collocazioni e concetti fondamentali</b> .....	7
1.1 Il contestualismo britannico.....	10
1.1.1 La svolta di Firth.....	10
1.1.2 I neo-firthiani: Halliday, Sinclair e Hoey.....	12
1.1.3 Sinclair e la Corpus Linguistics.....	15
1.2 Il formalismo continentale e le interpretazioni di collocazione più influenti in ambito fraseologico.....	18
1.2.1 Cowie e il continuum dell'idiomaticità.....	19
1.2.2 Hausmann e lo status dei costituenti in collocazione.....	21
1.2.3 Mel'čuk e le <i>Lexical Functions</i> .....	22
1.3 Apporti della Linguistica Cognitiva e della Psicolinguistica: le collocazioni come <i>pre-fabricated chunks</i> .....	25
1.3.1 Langacker e le <i>symbolic units</i> .....	26
1.3.2 Linguistica Cognitiva e la <i>teoria dei prototipi</i> .....	26
1.3.3 Psicolinguistica e il <i>lessico mentale</i> .....	27
1.4 Approccio adottato nel presente lavoro.....	30
<b>2 Definizione funzionale, strumenti e metodologia di ricerca</b> .....	33
2.1 Definizione funzionale di collocazione lessicale e proprietà delle collocazioni.....	33
2.2 Repertori e trattamento lessicografico delle collocazioni.....	37
2.3 Corpus utilizzato per il censimento dei pattern collocazionali.....	45
2.4 Metodologia di ricerca.....	46
<b>3 Descrizione dei pattern collocazionali censiti nel corpus e raffronto con la ricognizione condotta nei repertori</b> .....	49
3.1 Pattern Aggettivo in posizione libera.....	50
3.1.1 Descrizione del pattern Aggettivo in posizione libera.....	50
3.1.2 Descrizione di alcuni casi del pattern Aggettivo in posizione libera: <i>brutta, triste, curiosa, scarse</i> .....	52
3.1.3 Ampliamento del pattern Aggettivo in posizione libera: <i>priva di</i> .....	56

3.2	Pattern Sequenze Obbligate.....	57
3.2.1	Descrizione del pattern Sequenze Obbligate.....	58
3.2.2	Descrizione di alcuni casi del pattern Sequenze Obbligate: <i>ultime notizie e notizie riservate</i> .....	59
3.2.3	Ampliamento del pattern Sequenze Obbligate: <i>notizia scomoda</i> .....	61
3.3	Pattern Verbo + Complemento.....	61
3.3.1	Descrizione del pattern Verbo + Complemento.....	62
3.3.2	Descrizione di alcuni casi di pattern Verbo + Complemento: <i>verbi ad alternanza forma transitiva ~ pronominale</i> .....	64
3.3.3	Ampliamento del pattern Verbo + Complemento: <i>scambiarsi, reperire e notizie coperte</i> .....	65
3.4	Pattern Soggetto + Verbo.....	66
3.4.1	Descrizione del pattern Soggetto + Verbo.....	67
3.4.2	Descrizione di alcuni casi di pattern Soggetto + Verbo: <i>trapela e filtra</i> .....	68
3.4.3	Ampliamento del pattern Soggetto + Verbo: <i>compare, proviene, raggiunge, risale; riguardante e concernenti</i> .....	69
	<b>Conclusioni</b> .....	71
	<b>Appendice</b> .....	81
	<b>Bibliografia</b> .....	113

## Introduzione

*Serbare rancore, discutere animatamente, idea brillante, la paura attanaglia qualcuno*, sono alcuni esempi di collocazioni lessicali, che assieme alle polirematiche, modi di dire e proverbi, rientrano nel variegato panorama delle combinazioni di parole (*Multi Word Expressions*, MWE), un'area del linguaggio che si trova al di là della parola e al di qua della frase.

Le collocazioni lessicali, definibili in senso ampio come *co-occorrenze* di parole culturalmente motivate, che, in ragione di una solidarietà lessicale consolidata dall'uso, esibiscono la consuetudine di presentarsi insieme nello stesso contesto linguistico, nonostante la lingua metta a disposizione dei parlanti delle scelte alternative, ma meno efficaci nel veicolare i concetti, ad esempio *avere rancore, discutere in modo acceso, buona idea, la paura prende qualcuno*. Un buon livello di competenza linguistica è garantito quindi, dall'uso appropriato delle collocazioni oltre che dalle conoscenze grammaticali e sintattiche e se i parlanti madrelingua acquisiscono tale capacità in modo implicito, per gli apprendenti di una lingua straniera rappresentano un'inevitabile difficoltà, soprattutto nella produzione, per il mancato isomorfismo causato dall'essere culturalmente tipizzate; ad esempio, in italiano si dice *fare la doccia*, in inglese *to have a shower* (trad. letterale, *avere una doccia*), in francese *prendre une douche* (*prendere una doccia*), in tedesco *Duschen* (*docciarsi*). Considerate nella dimensione diafasica, le collocazioni e i *tecnicismi collaterali* contraddistinguono i linguaggi settoriali, oggetto di studio della Terminologia, dove assumono una connotazione estremamente specifica, ad es. in medicina *apprezzare una lesione* (nel senso di *riscontrare*), in giurisprudenza *comminare la pena* (nel senso di *assegnare*), in informatica *scaricare un file* (nel senso di *trasferire un file nella memoria di un computer*), o in campo amministrativo *liquidare gli stipendi* (nel senso di *calcolare e pagare*). Similmente, denotano anche l'appartenenza di un membro ad uno specifico gruppo sociale, ad es. *far scendere il morto e affrancamento del colore* nel gioco del Bridge. Le collocazioni sono dunque, un elemento costitutivo dell'organizzazione del discorso e i lessemi fra loro combinati stabiliscono relazioni sintagmatiche che innescano processi di varia natura, primo fra tutti la reciproca di influenza di significato, il *meaning by collocation* di firthiana memoria.

Le collocazioni sono un fenomeno linguistico tanto pervasivo quanto sfuggente, che per lungo tempo è passato inosservato, ma che a partire dalla seconda metà del ventesimo secolo ha catturato l'interesse dei linguisti specializzati nelle diverse scienze del linguaggio, i quali pur ricercando criteri unitari per descriverle, finora non sono pervenuti alla formulazione di una teoria unitaria e sotto questo aspetto rappresentano una sfida stimolante per la Linguistica teorica e applicata.

Gli studi condotti negli ultimi decenni si sono avvantaggiati del progresso tecnologico e la Linguistica dei Corpora (*Corpus Linguistics*) ha supportato i ricercatori con collezioni di testi di lingua d'uso orale e scritto di sempre più vaste dimensioni e con strumenti di estrazione automatica sempre più potenti; le descrizioni e le analisi linguistiche condotte sullo spoglio dei corpora si sono rivelate utili per essere applicate, sia a livello intralinguistico che interlinguistico, in ambiti quali la lessicologia, per approfondire la nozione unità lessicale e segno linguistico, e la lessicografia, che si interroga sulla modalità per registrare e inserire le collocazioni nei repertori al fine di illustrare il comportamento di un'entrata lessicale, siano essi dizionari generali della lingua d'uso o speciali di combinazioni di parole o bi- e pluri-lingue,.

Il presente lavoro origina dalle sollecitazioni indotte nel corso degli ultimi trent'anni da lessicologi e lessicografi stranieri e italiani.

I principi metodologici *accept the evidence* e *reflect the evidence* affermati da Sinclair, fondatore della Corpus Linguistics e del progetto COBUILD, da un lato privilegiano l'abilità di esaminare i corpora per la qualità delle prove che mostrano a scapito dei giudizi di grammaticalità e accettabilità emessi dall'intuizione umana e dall'altro invitano i linguisti ad attingere a citazioni di lingua d'uso, più genuine rispetto ad esempi plausibili, ma inventati *ad hoc* (Sinclair, 1991: 4-5).

Se l'intento dei lessicografi è di offrire la descrizione della lingua d'uso e non di una lingua ideale, i pattern collocazionali devono inevitabilmente farne parte, poiché non si possono ignorare ulteriormente le prove ricavabili dallo spoglio dei corpora (Moon, 2012: 315); la necessità di dedicare sempre più spazio ai rapporti sintagmatici più frequenti fra parole all'interno dell'articolo di un lemma è dimostrata anche dalle analisi condotte sui corpus e dagli studi cognitivi e psicolinguistici, secondo i quali la produzione e l'interpretazione dei messaggi avvengono combinando pacchetti di parole (*pre-fabricated-chunks*) (Marello, 1996: 43). Giacomini pone a fondamento della descrizione

lessicografica la rilevanza linguistica e il posizionamento stabile delle collocazioni all'interno di un testo, ottenibili per mezzo dell'estrazione e della selezione dei dati basate su criteri di significatività statistica, e auspica che la lessicografia italiana dedichi maggiore attenzione alla rappresentazione delle collocazioni nei dizionari (Giacomini, 2012: 287).

Il patrimonio collocazionale di una lingua è un sistema in continua evoluzione la cui rappresentazione nei dizionari non è semplice e incontra incoerenze o omissioni: per queste ragioni diventa imprescindibile il riferimento ai dati autentici ottenuti dai corpora sulla base della frequenza delle co-occorrenze e delle Misure di Associazione; questa modalità operativa garantisce l'inserimento nei repertori dei pattern collocazionali più rilevanti per una lingua, promuovendo al tempo stesso una descrizione dei vocaboli rispettosa degli obiettivi del dizionario e delle esigenze dei destinatari, i quali andrebbero guidati nel riconoscimento del carattere collocazionale delle combinazioni dei vocaboli evidenziandole graficamente oppure inserendole separatamente (Casadei, 2019: 180)

La tesi muove dall'ipotesi che la descrizione dell'italiano offerta nei glossari accusi un certo grado di soggettività e sia fondata più sull'intuizione e sull'introspezione del compilatore, piuttosto che sulla rigorosa osservazione dell'uso reale che i parlanti fanno della lingua; la ricerca instaura un raffronto tra l'uso effettivo nella lingua attestato in un corpus di riferimento e ciò che è recensito nei repertori, allo scopo di verificare e eventualmente ampliare i *pattern lessicali collocazionali* citati nei glossari stessi; un *pattern* (modello, schema ricorrente, struttura ripetitiva) è la configurazione regolare assunta da particolari sequenze di parole all'interno di un insieme di risultati sperimentali e in ambito linguistico è la struttura sintagmatica minima nella quale tutte le parole sono stabili nel senso, o meglio non ambigue (Ježek, 2016: 77). La parola-perno scelta a paradigma dei pattern collocazionali lessicali è *notizia*.

L'approccio adottato è quantitativo in parte mitigato da quello fraseologico e il metodo *corpus-driven* si intreccia con alcuni criteri (morfo)sintattici e semantici idonei a distinguere e descrivere i vari tipi di pattern collocazionali.

La tesi è articolata in quattro capitoli. Nel Cap. 1, dopo aver fornito una panoramica sullo *state of the art* dei diversi orientamenti allo studio delle collocazioni, viene individuato l'approccio adottato nel presente lavoro. Il Cap. 2 è dedicato agli strumenti e alla metodologia utilizzati per la ricerca; i primi tre paragrafi presentano la definizione

operazionale di collocazione, i repertori sui quali è stata effettuata la ricognizione dei pattern lessicali collocazionali e il corpus di riferimento su cui è stato condotto lo spoglio delle concordanze, mentre il quarto paragrafo precisa il metodo impiegato per stabilire il raffronto repertori e corpus. Il Cap. 3, sulla base della tipologia dei pattern collocazionali del lessema *notizia*, procede alla descrizione dei risultati censiti con lo spoglio del corpus e confrontati con i dati rintracciati nei dizionari. Nel Cap. 4, si compie una sintesi dei risultati ottenuti dall'analisi del corpus e dei dizionari e si delineano le prospettive per futuri lavori. Infine, in Appendice, sono inserite le tabelle contenenti i dati lessicografici (ordinati prima in ordine alfabetico e per frequenza assoluta) e i risultati ottenuti con l'interrogazione del corpus, questi ultimi distinti fra l'esposizione delle co-occorrenze e le classifiche dei collocati stilate sulla base delle Misure di Associazione.

Desidero ringraziare il Prof. M. Cortelazzo per la cortese disponibilità e fiducia con cui mi ha seguita durante il percorso. Ringrazio inoltre il Prof. J. Garzonio, già relatore della tesi di laurea triennale, e i docenti del corso di studi in Linguistica che con i loro insegnamenti hanno fornito le conoscenze adeguate alla realizzazione di questo elaborato.



# Capitolo 1

## Panoramica degli approcci teorici alle collocazioni e concetti fondamentali

Nel corso del XX secolo, i teorici della lingua affrontano il fenomeno delle espressioni di più parole (*Multi-Word Expressions*, MWE) e concentrano il loro interesse su lessemi complessi, espressioni idiomatiche e legami lessicali, pervenendo a definizioni e strumenti di analisi condizionati dall'orientamento di appartenenza.

I prodromi di questo fervore si intravedono a partire dal XIX secolo (Squillante, 2016: 12-3): Prendergast (1864) osservò che l'apprendimento linguistico dei bambini è basato su *chunks* di lingua riutilizzabili e non solo su parole; Paul (1880) notò come alcune MWE equivalgono a unità concettuali, in ragione del loro significato non compositazionale e del loro comportamento, che ammette parziali o nulle sostituzioni dei costituenti e limitate modificazioni sintattiche; Sweet (1891) individuò delle espressioni, *special sentences* o *idioms*, che pur essendo regolari nella forma, sono semanticamente irregolari, non compositazionali; Bréal in *Essai de Sémantique* (1897) intuì che le parole tendono a creare dei *groupes articulés*, combinazioni di parole presenti nel lessico mentale dei parlanti madrelingua come gruppi precostituiti, in virtù dell'uso reiterato nel tempo. Il primo ad approfondirne la tipologia fu Bally in *Traité de Stylistique française* (1909), il quale rintracciò fra le risorse disponibili in una determinata lingua, l'esistenza di *séries phraséologiques* (serie o gruppi usuali di lessemi), cioè delle espressioni intermedie fra le combinazioni libere e le *unités phraséologiques* (non scomponibili e dal significato non compositazionale), i cui elementi pur conservando la propria autonomia sono in qualche modo affini e quando compaiono in un testo producono l'effetto di *déjà vu* (Legallois-Tutin, 2013: 4-5).

A metà del XX secolo, i linguisti focalizzano l'attenzione sui legami lessicali tra parole e formulano la nozione di *collocazione*, che nell'accezione più ampia e neutra del termine è “an institutionalized word combination, corresponding to a conventionalized way of saying a certain thing” (Ježek, 2016: 199).

Gli approcci principali alla definizione e descrizione delle collocazioni sono il contestualismo britannico (1.1) e il formalismo continentale (1.2), entrambi di matrice strutturalista. Il contestualismo, presuppone che i processi di comunicazione siano

impossibili da formalizzare, perché complessi e mutevoli e considera le collocazioni un fenomeno psicolinguistico, identificabile con certezza solo per mezzo di metodi statistici, che misurano le co-occorrenze significative e regolari di parole, senza necessità di ricorrere alla competenza e all'intuizione dei parlanti; per il formalismo invece, si tratta di *phrasal units* con statuto lessicale, descrivibili e classificabili linguisticamente. I due approcci, pur divergenti nell'atteggiamento, descrittivo il primo e prescrittivo il secondo, si possono ritenere complementari e convergono nell'impegno lessicografico, poiché le riconducono a unità della *parole* e non della *langue*, in termini saussuriani.

A questi due modelli, si aggiungono la Linguistica Cognitiva e i contributi della Psicolinguistica (1.3), che al pari del contestualismo, attribuiscono massima importanza alla frequenza d'uso, tanto da indicarla quale parametro decisivo per spiegare la memorizzazione e il recupero delle collocazioni nel lessico mentale.

Gli approcci menzionati condividono una visione olistica del linguaggio (le funzioni cognitive formano un complesso apparato inscindibile, composto di sotto-elementi con diverse strutture e funzioni), mentre l'approccio generativista si distingue per la visione modulare dello stesso (le funzioni cognitive sono isolate e localizzate in modo separato nell'encefalo).

Il modello descrittivo della Grammatica Generativa (*Transformational-Generative Grammar*) incontra delle difficoltà e dei limiti nel trattamento delle MWE; la Grammatica Generativa teorizza il sistema linguistico articolato su due livelli separati: *grammar*, un insieme di regole che aggregano gli elementi linguistici e interessata alle caratteristiche della struttura ma che trascura la significazione, e *lexicon*, un deposito di irregolarità non composizionali che può essere solamente imparato a memoria e, in quanto appendice della grammatica, ha la finalità di procurare liste di parole o di espressioni atte a riempire i nodi terminali ( $X^0$ ) della struttura sintagmatica; questi *slot* prevedono *item* semplici, ma considerato che le MWE esibiscono un comportamento estremamente variabile in termini di operazioni sintattiche ammesse e non tutte spiegabili con fattori di *performance*, la teoria generativista dovrebbe postulare un *lexicon* contenente una lista delle manipolazioni concesse da una determinata unità; tuttavia, questa soluzione non sarebbe attrattiva per un approccio che evita la ridondanza (Gries, 2008: 11). La tradizione generativista rivolge il proprio interesse alle espressioni idiomatiche (per es. *il mattino ha l'oro in bocca*) e in particolare quelle verbali (per es. *mettere in moto*), considerati

strutture grammaticali ben formate, ma gravate da restrizioni sintattiche e anomalie semantiche; quest'ultimo aspetto è un effettivo e forte ostacolo all'inclusione delle MWE nella teoria standard di Chomsky, il quale però non manca di sottolineare la differenza fra *close constructions* (collocazioni grammaticali), ad es. *decide on a boat*, col significato di “choose (to buy) a boat” e *loose associations* (combinazioni libere), ad es. *decide on a boat*, in cui la combinazione Verbo+Preposizione ha il significato di “make a decision while on a boat”) (Benson-Benson-Ilson, 1986: ix).

La pervasività delle MWE impone una revisione critica al paradigma generativista dominante il '900 e alla visione modulare del linguaggio, innanzitutto riguardo al riconoscimento dell'idiomaticità come fenomeno centrale della produzione linguistica, non periferico né minoritario, e dell'esistenza di sottili interdipendenze fra i diversi livelli linguistici riconoscibili dall'analisi dei fraseologismi prodotti nella *performance*.

Pustejovsky in *Generative Lexicon* (1995), nota che il modello generativista trascura l'uso creativo delle parole in nuovi contesti e sostenendo che molta parte dell'informazione strutturale di una frase è meglio codificabile da una prospettiva lessicale, delinea i problemi più urgenti che la semantica lessicale necessita di affrontare: la natura polimorfica della lingua, la semanticità delle espressioni delle lingue naturali, l'uso creativo delle parole in nuovi contesti e lo sviluppo di una rappresentazione semantica co-composizionale.

Nonostante il notorio dualismo fra sintassi e semantica, gli studi condotti sulle MWE rappresentano un tentativo di superamento dello stesso, in cui non esiste conflitto fra lo studio dei livelli della grammatica e quello del comportamento del lessico, perché sono modi differenti che impiegano tecniche diverse per indagare lo stesso fenomeno: l'andamento *top-down* della grammatica ragiona per leggi generali, ma è alla costante ricerca di dettagliare le informazioni di cui dispone, mentre *bottom-up* del lessico osserva il comportamento dei singoli termini, ma è alla ricerca di somiglianze fra item che possano raggrupparli; quindi, se la prima scava in profondità nella lingua, il secondo porta in superficie.

Le collocazioni rappresentano a tutti gli effetti una sfida stimolante per la linguistica e le scienze coinvolte nello studio del linguaggio umano, che a oggi non sono pervenute a una teoria unitaria del fenomeno, contraddistinto dalla debolezza intrinseca al concetto stesso di collocazione (2.1).

## 1.1 Il contestualismo britannico

Che il significato di una parola risulti implicato dalle altre che si trovano nello stesso contesto linguistico è l'idea fondante del contestualismo, il quale si configura come una corrente influenzata dagli studi condotti da Firth a partire dagli anni '30, relativi al lessico dei linguaggi settoriali che, originandosi anche nella dimensione sintagmatica, sarebbe in grado di influenzare la semantica; la preminenza assegnata al contesto rappresenta il presupposto che fonda le ricerche di Halliday sulla coesione testuale, di Sinclair sul rintracciamento delle collocazioni ancorato all'indagine statistica di ampi corpora di lingua d'uso e di Hoey sul funzionamento di tali co-occorrenze.

L'approccio empirico-induttivo (detto anche statistico-quantitativo, distribuzionale o *frequency-oriented approach*) interpreta le collocazioni come co-occorrenze di due o più termini (combinazioni di parole) che compaiono a una determinata distanza (*span*), la cui frequenza d'uso è maggiore sia di quanto ci si potrebbe attendere e sia rispetto alla frequenza d'uso in isolamento degli stessi termini; nei suoi sviluppi, il metodo si basa sull'osservazione orientata *bottom-up* e *corpus driven*, e i criteri significativi sono l'*uso* e la *frequenza*.

Il lavoro svolto in lessicologia da Firth è alla base della *distributional hypothesis* sulla natura del significato, fondata sull'assunto che il significato di una parola correla con la sua distribuzione, cioè l'insieme dei contesti nei quali essa compare, costituiti dalle parole con le quali la parola stessa entra in una relazione di tipo grammaticale (per es. l'Aggettivo modificatore del Sostantivo o il ruolo di Complemento Oggetto di un Sostantivo in riferimento al Verbo); il significato è inerentemente differenziale e può essere stabilito comparando il set dei contesti, un'operazione da cui si evince che le parole aventi proprietà distribuzionali simili hanno significati simili (Ježek, 2016: 73-4).

### 1.1.1 La svolta di Firth

Nel 1944 J.R. Firth elaborò una teoria del significato basato sull'uso delle parole nel contesto linguistico, che venne divulgata negli articoli confluiti nelle raccolte tarde *Papers in Linguistics 1934-51* (Firth, 1957) e *Selected Papers of J.R. Firth 1952-59* (Palmer, 1960).

Quando L. Wittgenstein afferma "One cannot guess how a word functions. One has to look at its use, and learn from that." (Palmer, 1960: 138), riecheggia le parole di Firth:

“The use of the word ‘meaning’ is subject to the general rule that each word when used in a new context is a new word.” (Firth, 1957: 190)

L’idea di contesto è mutuata da *context of situation*, nozione impiegata per la prima volta dall’etnografo B. Malinowski, il quale durante le sue ricerche antropologiche incentrate sulle lingue delle popolazioni del Sud Pacifico, nota che il linguaggio umano esprime dei bisogni funzionali e che la comprensione di un enunciato non può prescindere dal contesto culturale e situazionale in cui viene formulato; in Firth, *context of situation* si configura in un costrutto schematico da applicare agli eventi comunicativi, costituiti sia da elementi linguistici che extra-linguistici (Palmer, 1968: 155); al tempo stesso, il contesto linguistico diventa funzionale a definire la significazione tramite collocazione (*meaning by collocation*).

In *Modes of Meaning* del 1951, Firth teorizza il lessico non come un semplice catalogo di parole, bensì come una serie di possibilità di collocazioni consolidata dall’uso (*collocability*) e sostiene che il significato di una parola si stabilisce su più livelli interdipendenti (ad es. sociale, sintattico, fonetico, etimologico) “rather like the dispersion of light of mixed wave-lengths into a spectrum” (Firth, 1957: 192); per cogliere appieno il significato di un termine è necessario quindi, focalizzarsi sull’osservazione delle relazioni sintagmatiche che si creano nel testo e una di queste è *meaning by collocation*: in questo caso il senso di ciascun costituente lessicale è specificato dalla relazione di collocazione nel contesto sintagmatico stesso, corroborato dalla routine delle comunicazioni, spesso stereotipate e strettamente condizionate dal tipo di cultura (ad es. in *nebbia fitta* i termini acquisiscono piena specializzazione per la ricorrente comparsa di *nebbia* nelle vicinanze di *fitta* e viceversa) (Firth, 1957: 196); infine, afferma che “Collocation states the habitual company a key-word keeps. Words must at some level or other be taken at their face value in their common and usual verbal environment” (Palmer, 1968: 113).

Il suo contributo segna una svolta rispetto la tradizione strutturalista; infatti, se gli strutturalisti attestano l’esistenza di unità intermedie fra parola e frase ed evidenziano la difficoltà di analizzarne le caratteristiche sintattiche e semantiche, l’approccio *strutturale* di Firth si prefigge di indagarne il funzionamento nel linguaggio; i primi segnali di questa reazione si colgono in *The Technique of Semantics* del 1935, quando riflettendo sulla

sistematicità dell'impiego di citazioni ed esempi a corollario della definizione di un lemma nei dizionari, osserva:

“...the complete meaning of a word is always contextual, and no study of meaning apart from a complete context can be taken seriously. [...] Throughout our review of the study of meaning we have seen how it has been split up as a relation or system of relations. [...] I propose to split up meaning or function into a series of component functions. Each function will be defined as the use of some language form or element in relation to some context.” (Firth, 1957: 7 e 19)

Firth applica la sua teoria allo studio dei linguaggi settoriali, alla traduzione, alla comprensione e all'insegnamento della lingua e auspica che i lessicografi la utilizzino per redigere nuovi tipi di dizionari appositamente dedicati alla terminologia, alla lingua d'uso e alle espressioni idiomatiche (Firth, 1957: 33); inoltre, già a partire dagli anni '30, Firth constata che per cogliere il reale funzionamento della lingua è tempo di utilizzare tecniche di tipo statistico nel trattamento dei testi, nella prospettiva di un approccio unitario empirico-teorico (Firth, 1957: 17). Nel panorama attuale, si fa riferimento alle collocazioni studiate da Firth con il termine *collocazioni empiriche*, intese come una semplice combinazione di parole particolarmente ricorrente o statisticamente rilevante, la cui frequenza è attestata in un corpus.

### 1.1.2 I neo-firthiani: Halliday, Sinclair e Hoey

Gli innovativi contenuti di Firth sulle relazioni di collocazione condizionano le ricerche dei suoi più illustri allievi: Halliday, focalizzato sulla grammatica, Sinclair, concentrato sulle evidenze statistiche e Hoey, attento ai processi di acquisizione linguistica. I tre epigoni condividono sia la visione delle collocazioni come la regolare e statisticamente discernibile co-occorrenza di parole esprimibile quantitativamente, sia il *principio di co-selezione* (simultanea selezione da lessico e grammatica), che vanifica la distinzione tra grammatica e lessico, *competence* e *performance*, *langue* e *parole* (Sinclair, 1991: 103-4); tra i seguaci dell'approccio statistico, a fronte dei punti contatto, le definizioni di collocazione presentano importanti punti discordanti, ad esempio nell'individuazione dell'unità base della collocazione: se la parola è da intendersi in senso stretto come lemma e di questo considerare separatamente ciascuna forma e per ogni forma analizzare la variazione dei collocati al fine render conto delle differenze tra i collocati stessi (ad es. i collocati di *frutto*, *frutti* e *frutta*), oppure in senso ampio per cui, secondo Halliday, *strength*, *strong* e *strengten* (trad. *forza*, *forte* e *rafforzare*) sono varianti di uno stesso item lessicale, cioè un gruppo di parole unite dalla relazione di

derivazione morfologica; un'altra questione dibattuta è il numero dei lessemi in collocazione e se debbano essere adiacenti o meno; infine, altri elementi di variazione sono il grado di *rigidità* (*figement, frozeness*) sintattica e lessicale, la *non-composizionalità* semantica (2.1) e se la collocazione costituisca o meno un'unità di senso.

Halliday in *Cohesion in English* (Halliday-Hassan, 1976) per primo utilizza il termine (*lexical*) *co-occurrence* ricondotto alla coesione testuale e alle nozioni di *prossimità* e *frequenza*; nello specifico, i fattori che influiscono sulla forza coesiva di due costituenti lessicali si rintracciano nelle relazioni interne al sistema linguistico da un lato e in quelle del testo dall'altro:

“[...] There are degrees of proximity in the lexical system, a function of the relative probability with which one word tends to co-occur with another. Secondly, in the text there is relatedness of another kind, relative proximity, in the simple sense of the distance separating one item from another, the number of words or clauses or sentences between. [...] There is a third factor influencing the cohesive force between a pair of lexical items in a text, and that is their overall frequency in the system of the language.” (Halliday-Hassan, 1976: 290)

Le collocazioni generate dall'associazione di elementi lessicali che co-occorrono regolarmente sono un elemento chiave della *lexical cohesion* e l'enfasi affidata al  *collocational environment* ancora l'accezione di un termine al contesto testuale, situazionale e al senso distribuzionale; Halliday in *Lexis as a linguistic level* del 1966, superando il classico dualismo fra semantica e sintassi, constata l'esistenza di una *lexicalness* complementare alla *grammaticalness* ed evidenzia la difficoltà per i due ambiti di descrivere le collocazioni in termini di linguistica formale, poiché le relazioni instaurate fra i costituenti coinvolti non sempre attengono a quelle tradizionali della grammatica (Squillante, 2016: 22); la definizione puramente distribuzionale (probabilità di co-occorrenza) e statistica di collocazione da lui formulata alcuni anni prima si rivela valida:

“Collocation is the syntagmatic association of lexical items, quantifiable, textually, as the probability that there will occur at n removes (a distance of n lexical items) from an item x, the items a,b,c...Any given item thus enters into a range of collocation, the items with it is collocated being ranged from more to less probable.” (Halliday, 1961:276, cit. in Squillante, 2016: 22-3)

Grazie alla prospettiva elaborata da Halliday, gli spogli condotti sui corpora di Sinclair e Hoey mettono in luce la pervasività, la natura dinamica e la proprietà delle collocazioni di creare significato. Dal progetto COBUILD (1.1.3) promosso da Sinclair,

si sviluppa l'area di ricerca di Hoey, che in prima istanza recepisce le collocazioni in termini puramente statistici, ma successivamente evidenzia che questa definizione ha il limite di confondere il metodo con lo scopo e non fornire alcuna informazione riguardo ai motivi della loro esistenza e al loro funzionamento (Hoey, 2005: 3-4).

Hoey elabora una teoria del lessico che lui stesso definisce provocatoria e avvincente, nella quale rovescia i ruoli tradizionali di lessico e grammatica, sostenendo che il lessico è strutturato in modo complesso e sistematico e che la grammatica rappresenta il risultato di tale struttura; la teoria poggia sui risultati emersi dallo spoglio di un corpus generale composto da 95 milioni di parole provenienti da articoli d'attualità, 3 milioni da testi scritti e 230.000 da testi orali contenuti nel *British National Corpus*; all'interno di questa cornice teorica, le collocazioni rappresentano un fenomeno psicolinguistico che offre indizi sul modo in cui è organizzata una lingua:

“a psychological association between words (rather than lemmas) up to four words apart and is evidenced by their occurrence together in corpora more often than is explicable in terms of random distribution.” (Hoey, 2005: 5)

Le collocazioni, per Hoey, partecipano tanto alla significazione quanto alla naturalezza di un enunciato e sono strettamente connesse al processo psicologico di *lexical priming*, un meccanismo che trae origine dai ripetuti incontri fra parole all'interno dello stesso contesto e proprio tali incontri sono funzionali alla costruzione di concordanze mentali, dove determinate parole co-occorrono innescandosi a vicenda (Hoey, 2005: Cap. 1); il *lexical priming* in parte deriva dalla nozione psicolinguistica di *semantic priming*, secondo cui la conoscenza di una parola si sviluppa in virtù dell'effetto cumulativo dei ripetuti incontri con determinate parole; le premesse importanti alla teoria sono che:

“Words are individually primed and it would seem that they remain individually primed. [...] Priming is always matter of probability rather than requirement. [...] Priming is genre and domain specific in the first instance, though there are many primings that apply across generic and domain boundaries. Predictably, the specificity of the priming is at its greatest when the priming relates to discourse properties, since these have been shown to vary greatly according to genre and text purpose.” (Hoey, 2005: 73, 140, 115)

Hoey sottolinea che il *priming* si attiva in fase di acquisizione linguistica e colpisce la parola, non il suo significato, ed è in grado di modellare il lessico mentale del parlante, il quale immagazzina l'unità lessicale non in isolamento, bensì entro *chunks*, in una costruzione che giace dormiente nella memoria del parlante, ma che suscita aspettative di



co-occorrenza o *primes* ogni volta che un termine viene prodotto o interpretato (López Ferrero-Battaner, 2017: 268).

### 1.1.3 Sinclair e la Corpus Linguistics

Sinclair è il fondatore della Linguistica dei Corpus (*Corpus Linguistics*) in seno alla Birmingham School e nel 1974 la coppia Jones-Sinclair per prima procede all'analisi computazionale delle collocazioni su un corpus di testi orali (135.000 parole) e scritti (12.000 parole) (Jones-Sinclair, 1974: 17).

Sinclair è anche il promotore del progetto COBUILD (Collins Birmingham University International Language Database), che a partire dagli anni '60 impiega gli sviluppi della tecnologia computazionale per progredire nel campo della lessicografia computazionale e per mezzo di parametri statistici egli valuta il grado di significazione di una co-occorrenza al fine di fornire buoni candidati e migliorare i metodi di estrazione delle collocazioni (Orlandi 2016: 20); il progetto COBUILD iniziato nel 1980, nel 1987 culmina nella prima edizione del *Collins COBUILD English Dictionary*, per il quale viene sviluppato un nuovo stile di presentazione dell'informazione lessicografica, improntato sull'impiego della lingua d'uso sia per la citazione di esempi che per la descrizione del significato del lemma. Il progetto COBUILD è fortemente legato allo sviluppo dell'*idiom principle*, dove la nozione di collocazione gioca un ruolo centrale.

La *Corpus Linguistics* applica i postulati teorici di Firth all'utilizzo di strumenti computazionali e tecniche di processamento statistico, che permettono di reperire automaticamente il contesto di occorrenza di una parola in corpora di lingua d'uso equilibrati e di taglia sufficiente; l'importanza imprescindibile dell'uso e del contesto linguistici, trova giustificazione nei principi metodologici *accept the evidence* e *reflect the evidence*; il primo, *accept the evidence* sottolinea il contrasto, specialmente in lessicografia, tra i dati catturati in modo oggettivo dal supporto elettronico e quelli raccolti nei dizionari, sui quali grava il giudizio soggettivo dei compilatori che decidono ciò che dovrebbe esservi incluso o meno; il secondo, *reflect the evidence* afferma che un esempio inventato, per quanto possa essere plausibile, non potrà mai essere proposto come esempio autentico della lingua d'uso e denuncia l'irragionevolezza della pratica di preparare appositamente esempi, dal momento che le istanze reali sono ampiamente disponibili; il prestigio acquisito dagli esempi autentici promuove l'idea che ogni

campione della lingua dipenda dal contesto circostante, riassumibile nella nozione di genuinità (*naturalness*).

Il metodo *bottom-up corpus driven* nel restituire le concordanze, e quindi esempi genuini, offre i vantaggi di informare riguardo all'interdipendenza fra i lessemi che condividono un contesto, di far emergere le co-occorrenze lessicali che non incontrano categorie linguistiche predefinite e di mostrarne la pervasività, tanto da promuoverle da fenomeno periferico a fenomeno centrale della lingua.

Le prime ricerche di Sinclair ruotano attorno al problema di descrivere le collocazioni in modo obiettivo e individuare il tipo di relazione che esiste fra la loro evidenza fisica e la percezione psicologica del significato; a tal fine, rifacendosi alla definizione statistica di Halliday (1.1.2) e occupandosi di  *collocational behaviour*, introduce le terminologie di *node*, *collocate*, *span* e  *collocational pattern*:

“A ‘node’ is an item whose total pattern of co-occurrence with other words is under examination; a ‘collocate’ is any item which appears with the node within a specified environment. [...] A ‘span’ is the amount of text within which collocation between items is said to occur. [...] Conventionally, span positions of collocates are numbered according to their distance from the node and whether they precede or follow it. [...] it was ascertained that for any node, every high proportion of relevant information could be obtained by examining collocates at position node  $\pm 4$ . [...] The ‘ *collocational pattern*’ of an item consists of all words  *appearing significantly often* in its environment with information about them.” (Jones – Sinclair, 1974: 16, 21, 34; il corsivo è mio)

L'espressione  *appearing significantly often* chiama in causa la frequenza e stabilisce che quanto maggiore è la frequenza di co-occorrenza rispetto a quella dei lessemi presi singolarmente, tanto più si può trattare di collocazioni:

“Collocation is the co-occurrence of two items in a text within a special environment. ‘Significant’ collocation is regular collocation between items, such that they occur more often than their respective frequencies and the length of text in which they appear would predict.” (Jones – Sinclair, 1974:19)

Ad esempio, in  *la fuga rocambolesca*,  *la* non è considerato un collocato significativo di  *fuga*, perché la sua frequenza generale è estremamente alta, infatti può accompagnarsi a tutti i sostantivi femminili singolari, mentre  *rocambolesca* è un collocato significativo, perché è una parola poco frequente e tendenzialmente si trova in contesti in cui co-occorre la parola  *fuga*.

Il calcolo delle frequenze induce Sinclair a operare un'importante distinzione fra  *upward* e  *downward collocations*; nel primo gruppo rientrano le parole che abitualmente si combinano con altre che risultano molto più frequenti, viceversa nel secondo ci sono

le parole che abitualmente si combinano con altre meno frequenti; Sinclair cita l'esempio della parola inglese *back*:

- se *back* co-occorre con le Preposizioni *down, at, from, on* tutte più frequenti dell'Avverbio, allora in riferimento a *back* le collocazioni sono *upward*;

- se *back* co-occorre con i Verbi *bring* o *arrive* che hanno frequenza minore rispetto all'Avverbio, allora in riferimento a *back* le collocazioni sono *downward*.

Sinclair definisce anche la diversa natura dei due tipi di collocazione: le *upward collocations* sono pattern statisticamente deboli e generalmente costituiti da associazioni sintattiche (Preposizioni, Congiunzioni, Avverbi), mentre le *downward collocations* costituiscono pattern marcati da un tratto semantico, i cui collocati sono Sostantivi o Verbi (Sinclair, 1991: 115-21).

Sul come le parole si combinano fra loro, Sinclair lo spiega attraverso due modelli di produzione linguistica e interpretazione del significato contrapposti ma cooperanti: *open-choice principle* e *idiom principle*. L'*open-choice principle* stabilisce che il parlante nel formulare un'unità linguistica dispone di un'ampia gamma di scelte lessicali, vincolata solamente dalle restrizioni grammaticali previste dal contesto (per es. *Maria parte per le vacanze* vs *\*Maria parte nelle vacanze*); l'*idiom principle*, il cui funzionamento è ben visibile nelle collocazioni, prevede che il parlante utilizzi costrutti già pronti non scomponibili o analizzabili in segmenti (*semi-constructed phrases*), che rappresentano una scelta reiterata e riutilizzabile (per es. *Gianni è un fumatore incallito* vs *\*Gianni è un fumatore impenitente*) (Casadei-Basile, 2019: 160); la motivazione dell'*idiom principle* in parte risiede nel *context of situation* (1.1.1): "To some extent, this may reflect the recurrence of similar situations in human affairs; it may illustrate a natural tendency to economy effort; or it may be motivated in part by the exigencies of real-time conversation." (Sinclair, 1991: 110).

Il contributo fornito dall'*idiom principle* non si esaurisce nel lavoro di Sinclair e la sua efficacia chiarisce in parte l'esistenza nel lessico mentale di *pre-fabricated items* (Pawly-Syder, 1983) e di *symbolic units* (Langacker, 1987, 1991) elaborati dalla Linguistica Cognitiva (1.3). Sinclair al termine della sua carriera, sottolinea la centralità dei *phrases* di cui le collocazioni fanno parte, perché idonei a esprimere "the place where structures are engineered to allow meanings to take shape." (Sinclair, 2008: 408).

## 1.2 Il formalismo continentale e le interpretazioni di *collocazione* più influenti in ambito fraseologico

L'approccio formalista alle collocazioni sostiene che la scelta di una determinata parola implichi delle restrizioni su quella delle parole attigue; questa tesi, seppur da una prospettiva epistemologica e metodologica divergente, in parte anticipa l'*idiom principle* di Sinclair, reso evidente dall'approccio empirico; l'importanza acquisita dai dati statistici non preclude quindi, un approccio focalizzato sulla struttura interna delle collocazioni, orientato verso uno scopo descrittivo e classificatorio, teso a indagare il funzionamento dell'*idiom principle* al fine di razionalizzare un'enorme massa di dati.

Il formalismo continentale (o europeo) origina da quello sovietico ed est-europeo, che annovera fra i suoi aderenti Vinogradov e Amasova, i quali occupandosi di uno specifico sotto-insieme di MWE nell'ambito della lingua letteraria russa, riprendono la distinzione introdotta da Bally tra *unités phraséologiques* e *séries phraséologiques* (Cap.1), per disporla lungo un continuum che va dalle combinazioni di parola più trasparenti e variabili a quelle più opache e fisse (Granger–Paquot, 2008: 28). Il formalismo europeo recepisce l'idea di continuum e promuove la ricerca di criteri linguistici (semantici, sintattici, pragmatici) validi per descrivere e classificare i diversi tipi di *phrases*; l'approccio formalista (detto anche fraseologico, semantico-sintattico, *significance-oriented, semantically-based approach*) intende descrivere il funzionamento del lessico attraverso le strutture sintattiche nelle quali esso si realizza ed elegge le collocazioni a oggetto d'indagine, talvolta facendone il principio stesso dell'ordinamento sintattico, come previsto dai modelli cognitivi *Pattern Grammar* e *Construction Grammar*.

L'approccio fraseologico mira alla valutazione qualitativa, di conseguenza le caratteristiche formali e il tipo di restrizioni lessicali, semantiche e sintattiche applicate alle collocazioni sono essenziali per definirle, classificarle e distinguerle dagli altri tipi di MWE: l'eterogeneità delle MWE si riflette nelle molte analisi, tassonomie e sotto-categorizzazioni prodotte e nella difficoltà di formalizzare la nozione di collocazione; inoltre, la centralità delle collocazioni nell'insegnamento delle lingue e nella traduzione dei testi è testimoniata dalla crescente attenzione che i lessicologi pongono nel formalizzare e presentare questo fenomeno nei dizionari.

Nel *BBJ Dictionary of English Word Combinations* (Benson-Benson-Ilson, 1986), Benson affidandosi alle proprie intuizioni osserva che le collocazioni sono caratterizzate da ripetitività, parziale predicibilità del significato e grammaticalità; tali qualità permettono di distinguerle dalle *free combinations* (massima possibilità di variazione paradigmatica, trasparenza semantica e modificabilità sintattica) e dagli *idioms* (espressioni cristallizzate, semanticamente opache e talvolta sintatticamente anomale). La tassonomia di Benson, suddivide le collocazioni in *lexical collocations*, formate da due costituenti di pari statuto lessicale, di cui il tipo più frequente è Verbo+Articolo+Nome (ad es. *coltivare una passione, lanciare un missile, versare una somma di denaro*), e *grammatical collocations*, composte da un termine dominante (Verbo, Nome, Aggettivo) seguito da una parola grammaticale, tipicamente una Preposizione (per es. *una discussione su, dipendere da, pronto per*) o una struttura linguistica (per es. *evitare di + Verbo infinito*); per inciso, il modello contestualista per riferirsi al contesto grammaticale utilizza il termine *colligation* (pattern grammaticale ricorrente e lessicalmente vincolato), coniato da Firth e successivamente approfondito nel  *collocational framework* da Renouf e Sinclair o nella *pattern grammar* da Hunston e Francis (1.3).

Il florilegio di esposizioni originato entro l'orientamento formale si spiega sia con la mancanza di accordo riguardo al tipo di restrizioni applicate, sia con l'intento prescrittivo funzionale alla lessicografia e all'organizzazione dell'entrata lessicale nei repertori; astraendosi dalla terminologia impiegata per nominare le collocazioni, la discrepanza fra le esposizioni può essere sanata dal loro confronto. In questa sezione sono presentate le interpretazioni di collocazione più influenti sulla lessicologia e lessicografia proposte dagli autori considerati classici e pertanto non può considerarsi esaustiva.

### **1.2.1 Cowie e il continuum dell'idiomaticità**

Cowie distingue le MWE in *composites* e *formulae*; queste ultime sono espressioni autonome con funzione pragmatica (ad es. *ci vediamo presto* o *ma certo!*), mentre i *composites* hanno funzione sintattica e sono ulteriormente suddivisi in sotto-categorie disposte lungo un continuum, che va dalle combinazioni di parola più trasparenti e variabili a quelle più opache e fisse: il continuum dell'*idiomaticità* si configura come un insieme di espressioni non divise da confini netti, ma con punti di contatto e sovrapposizioni che segnano lo sfumare dell'una nell'altra; per Cowie le MWE, sono *unità fraseologiche*, cioè unità lessicalizzate costituite da gruppi di parole bi- o pluri-

lessemici di uso comune, riproducibili e relativamente stabili dal punto di vista sintattico-semantic; Cowie esclude le *free combination* dal continuum, perché su di esse non si applica alcuna restrizione (Giacomini, 2012; 309):

Free combination		Restricted collocation	»	Figurative idiom	»	Pure idiom
<i>saltare l'ostacolo</i>		<i>saltare la fila</i>	»	<i>saltare sulla sedia</i>	»	<i>saltare di palo in frasca</i>

In questo frame le *restricted collocations* o semplicemente *collocations* sono:

“A composite unit which permits the substitutability of items for at least one of its constituents elements (the sense of the other element, or elements, remaining constant)” (Cowie, 1981, cit. in Giacomini 2012: 309)

I criteri adottati per distinguere i diversi tipi di *composites* sono: a) la presenza di una restrizione sulla combinazione, b) la variabilità lessicale e sintattica, c) la predicibilità o meno del significato, ossia l'interpretazione letterale o figurata degli elementi della MWE; in Cowie si hanno quindi:

*Free/Open Collocation (saltare l'ostacolo)*: totale apertura di collocabilità di ogni elemento in relazione all'altro; la sostituzione è limitata solo a livello semantico e tutti gli elementi denotano significato letterale;

*Restricted Collocations (saltare la fila)*: forte restrizione sulla collocabilità; la sostituzione è possibile, ma arbitrariamente limitata (*scavalcare la fila* vs *?sorpassare la fila*); almeno un elemento esprime senso figurato (*saltare*), mentre l'altro mantiene significato letterale (*fila*) e la combinazione è semanticamente trasparente;

*Figurative Idioms (saltare sulla sedia, nel senso di 'meravigliarsi, stupirsi')*: consistente restrizione sulla collocabilità; la sostituzione è raramente possibile (*sobbalzare sulla sedia*); l'espressione è parzialmente motivata e mostra senso figurato, ma conserva anche l'interpretazione letterale;

*Pure Idioms (saltare di palo in frasca)*: totale restrizione sulla collocabilità; la sostituzione è impossibile e quindi l'espressione è invariabile (*\*saltare di ramo in frasca* o *\*balzare di palo in frasca*); la combinazione esprime senso figurato ed è totalmente opaca.

Ježek (2016), pur condividendo la necessità di differenziare i tipi di combinazioni e l'assunto che sui *composites* si applica una regola di restrizione responsabile dell'innesco di un processo di lessicalizzazione, mette in discussione l'idea di *continuum*,

inteso come la semplice correlazione tra presenza di una restrizione e perdita di composizionalità del significato; innanzitutto, la presenza di una restrizione semantica non necessariamente correla con una minore predicibilità di significato, ad es. *pastorizzare il latte*, è una *Restricted Collocation*, dato che solo poche cose possono essere *pastorizzate*, solo certe azioni si applicano al *latte*, ma il significato è composizionale; inoltre, a fronte della rigidità sintattica sono numerose le sequenze i cui elementi pur non essendo sintatticamente autonomi, mantengono la trasparenza del significato, per es. *fuggire via*, *fare tardi*, *perdere peso*. Ježek stima che i tre criteri utilizzati da Cowie per distinguere i diversi tipi di MWE, correlano tra loro ma non in modo unidimensionale e giunge alla conclusione che:

“Being idiomatic is not the result of an increase in the restriction operating on the combination but of mechanism of a different nature. Despite the fact that idiomaticity manifests itself by blocking our ability to substitute the members of the sequence [...], the shift from literal to metaphorical meaning, which characterizes idiomatic expressions does not concern, except in a peripheral way, the phenomenon of word combination.” (Ježek, 2016: 215)

### 1.2.2 Hausmann e lo status dei costituenti in collocazione

Un aspetto divisivo all'interno dell'approccio fraseologico è la percezione del legame fra gli elementi della collocazione; per Cowie la natura dei costituenti in collocazione è indifferenziata, l'unica condizione da soddisfare è la restrizione applicata su uno degli item lessicali (anche se non specifica quale); Hausmann e Mel'čuk invece, sottolineano la diversa natura semantica delle *lexical units* (unità lessicali) in collocazione.

Hausmann a proposito delle *Kollokationen* (collocazioni) osserva che i lessemi in combinazione sono affini, cioè esibiscono la tendenza di comparire assieme nello stesso contesto (Hausmann, 1984 cit. in Giacomini, 2012: 315) e li distingue in: “ ‘base’ de la collocation le paternaire caractérisé...et ‘collocatif’ paternaire caractérisant” (Hausmann 1989, cit. in Giacomini, 2012: 315).

In questa prospettiva le collocazioni costituiscono una struttura bipartita e vincolata dall'uso/convenzionalità, dove gli item lessicali in conseguenza di un diverso status semantico esibiscono un rapporto gerarchico, per cui una funge da *base* (*Autosemantikon*, *base*) che seleziona unilateralmente l'altra, il *collocator* (*Synsemantikon*, *collocato*): la *base* è cognitivamente sovraordinata, semanticamente autonoma e non necessita di un secondo costituente per funzionare nella lingua, mentre il *collocato* è la componente

subordinata, semanticamente dipendente, che realizza pienamente il suo significato solo assieme alla *base*; ad es. *intavolare un discorso*, *intavolare* è il *collocato* che acquisisce significato solo unito a *discorso*, il quale costituisce la *base* e conserva il suo significato indipendentemente dalla costruzione in cui è inserita; per Hausmann quindi, le collocazioni sono semanticamente orientate, perché la *base* seleziona il collocatore, e limitate/ristrette, poiché per specificare un determinato significato la *base* mostra preferenze di collocabilità limitate; l'intero peso semantico dell'espressione poggia dunque sulla *base* e questo è particolarmente evidente nelle costruzioni a Verbo Supporto, per es. *prendere sonno* (3.3). Hausmann nota che in una determinata combinazione sintattica, la relazione fra *base* e *collocato* è costante: tendenzialmente i Sostantivi sono le *basi* dei Verbi e Aggettivi, mentre gli Aggettivi sono le *basi* per gli Avverbi; questo criterio rende conto delle collocazioni binarie, ma incontra dei limiti nel caso delle collocazioni non binarie, per le quali il rintracciamento di *base* e *collocato* richiederebbe la scomposizione della MWE in unità linguistico-concettuali più piccole (Giacomini, 2012: 164).

La classificazione di Hausmann, pur riprendendo i criteri di variabilità lessico-sintattica e trasparenza semantica introdotti da Cowie (1.2.1), li declina secondo il grado di autonomia dei lessemi, per distinguere *Ko-Kreation* e *Kontra-Kreation* dai fraseologismi; i fraseologismi a loro volta sono ripartiti in *Kollokationen*, *Teil-Idiome* e *Idiome* (Giacomini, 2012: 314-5) e precisa che:

“la collocation se distingue de la combinaison libre [...] par la combinabilité restreinte (ou affinité) des mots combinés [...]; se distingue d'autre part des locutions [...] par son non-figement et par sa transparence” (Hausmann, 2003: 311, cit. in Giacomini, 2012: 310)

Hausmann stabilendo la diversa natura semantica dei costituenti in collocazione si approssima al concetto di *meaning by collocation* di Firth (1.1.1) e procede in modo simile a Sinclair quando afferma che le combinazioni libere fanno parte del livello del sistema, della *langue*, cioè del livello dell'*open choice principle*, mentre le collocazioni attengono alla norma, all'uso, alla *parole*, ossia dell'*idiom principle* (1.1.3) (Casadei, 2019: 161).

### 1.2.3 Mel'čuk e le *Lexical Function*

Per Mel'čuk, le collocazioni costituiscono la maggioranza dei fraseologismi e al fine di descriverle nel modo più rigoroso, sistematico ed esaustivo possibile, a partire da metà



anni '60 e con progressivi sviluppi, elabora la *Meaning-Text Theory* (MTT), che modella il lessico e le unità fraseologiche avvalendosi dell'apparato delle *Lexical Functions*.

Il modello teorico di Mel'čuk poggia sull'analisi delle collocazioni dal punto di vista della produzione e costruzione del messaggio da parte del parlante (vs interpretazione del ricevente) e sull'allargamento della nozione saussuriana di doppia articolazione del segno linguistico (*significante* e *significato*) alla *proprietà combinatoria*, intesa come la capacità di un segno di combinarsi con gli altri presenti nello stesso sistema linguistico e rappresentata dall'insieme delle restrizioni applicate sulle co-occorrenze del segno stesso (Mel'čuk, 1998: 25).

Le lexical function consentono la rappresentazione astratta delle relazioni semantiche e sintattiche fra item lessicali e sono espresse attraverso la formula matematica  $f(x)=y$ , dove  $x$  è la lexical unit argomento della funzione  $f$  e  $y$  è il set di lexical unit o pattern di relazioni. Le lexical function sono classificate secondo criteri di natura diversa: una prima distinzione si compie tra *paradigmatic lexical functions*, che si riferiscono alla selezione sull'asse verticale della lingua, e *syntagmatic lexical functions*, che rinviano alla combinazione dei lessemi sull'asse orizzontale, perciò solo queste ultime consentono la modellizzazione delle collocazioni; una seconda ripartizione introduce le *standard lexical functions*, entità informazionali corrispondenti a un legame lessicale molto generale e universale, e *non standard lexical functions*, responsabili di una relazione lessicale estremamente specifica, limitata a un ristretto numero di lessemi, se non a uno solo (per es. *anno bisestile*, *burro rancido*); infine una terza divide le *simple lexical functions*, costituite da un'unica lexical function, dalle *complex lexical functions*, una serie di *lexical functions* combinate fra loro secondo regole di composizione. (Mel'čuk, 2003:28-30).

La classificazione di Mel'čuk analogamente a Cowie si basa sui criteri di commutabilità dei costituenti e trasparenza di significazione, esclude i *free phrase* dai fraseologismi e separa i *semantic phrasemes* (*composites* di Cowie), dai *pragmatic phrasemes* (*formulae* di Cowie); le collocazioni rientrano nei *semantic phrasemes* e sono chiamate *semi-phrasemes*, ossia espressioni parzialmente fisse, che lungo il continuum dell'idiomaticità (espresso in termini di *figement*, *frozeness*, rigidità) occupano una posizione intermedia fra *free phrases* e *full phrasemes*; secondo Mel'čuk:

“la combinazione di A e B assume lo status di collocazione se e solo se soddisfa le seguenti condizioni:

1. il significato di AB include il significato di uno dei suoi componenti, A ad esempio, come sua dominante semantica (*keyword*);
2. A è scelto dal parlante per il suo significato indipendentemente da B e senza restrizioni;
3. B (*value*) è selezionato in modo restrittivo non per il suo significato ma come funzione di A per esprimere uno specifico significato.” (Mel’čuk, 2007: 120, cit. in Casadei, 2019: 163)

In riferimento alla terza condizione, Mel’čuk interpreta le collocazioni come *cooccurrences lexicales restreintes* (Mel’čuk, 2003: 27), cioè co-occorrenze di lessemi sulle quali si applicano una o più restrizioni: i termini che si trovano combinati nello stesso contesto più frequentemente rispetto a quanto ci si aspetterebbe per caso, lo fanno in virtù del fatto che sulla loro combinazione si applica una limitazione, che agevola la loro co-occorrenza e contemporaneamente vincola la loro compatibilità; nonostante questa definizione sia più raffinata di quelle formulate entro l’approccio statistico, non è completamente soddisfacente, perché non coglie la peculiarità delle restrizioni in atto nelle collocazioni rispetto a quelle che gravano su altri tipi di MWE, ad es. le *solidarietà lessicali* (2.1). Sempre nella terza condizione, la nozione di *value* evoca quella del *collocator* di Hausmann, ma fra i due termini vi è una differenza sostanziale: *value* indica il set completo delle lexical unit che co-occorrono con la *keyword* per esprimere un determinato senso, invece *collocator* si riferisce a un solo elemento (Nesselhauf, 2005: 18); ad es. nelle collocazioni *avanzare/condurre/eseguire un’analisi*, i Verbi assieme rappresentano il *value* della lexical function ‘do/perform’ della *keyword analisi*, ma singolarmente esprimono i collocati della base *analisi*.

Le *lexical functions standard* postulate dalla MTT sono essenzialmente degli Universali Linguistici, vale a dire l’insieme delle unità o proprietà linguistiche che risultano comuni a tutte le lingue del mondo, o perlomeno a un alto numero di esse (ad es. Verbi e Nomi); questa proprietà attribuisce loro validità interlinguistica recepita a livello internazionale e promuove anche la loro applicazione in diversi ambiti: nella Linguistica Computazionale per la traduzione e generazione automatica di testi, nell’apprendimento delle lingue (lingua materna e lingue straniere) e nella creazione di opere di riferimento, per es. grammatiche di consultazione e dizionari di collocazioni; Mel’čuk in collaborazione con altri autori ha compilato il *Dictionnaire Explicatif et Combinatoire* (DEC o ECD), dizionario specifico di *phrases*, le cui entrate intendono conferire tutte le informazioni lessicali sulla base delle lexical function. Infine, la validità

delle lexical function è confermata dagli studi psicolinguistici intorno alla memorizzazione della forma di un phrase, che avviene contestualmente alla memorizzazione della funzione o funzioni ad esso associate.

### **1.3 Apporti della Linguistica Cognitiva e della Psicolinguistica: le collocazioni come *pre-fabricated chunks***

Le collocazioni sono un fenomeno testuale ma anche psicologico e gli approcci cognitivi integrano la dimensione cognitiva a quella della lingua d'uso: la consapevolezza maturata da un parlante madrelingua circa le collocazioni usuali del suo idioma è una capacità che fa parte della sua competenza comunicativa.

La Linguistica Cognitiva sostiene che il linguaggio rappresenta un sistema simbolico in grado di riflettere l'organizzazione concettuale dell'esperienza (linguistica e non), all'interno del quale la semantica gioca il ruolo di principio generatore della grammatica, e quindi morfologia e sintassi sono semanticamente motivate. L'impostazione cognitiva si impernia su un nucleo di assunti inerenti al linguaggio umano alla luce dei quali il significato non costituisce un puro fenomeno linguistico, ma è il risultato di un processo cognitivo dove i concetti/significati possiedono un radicamento esperienziale e corporeo (*embodiement*): la cognizione di una lingua si basa su singole istanze memorizzate, si sviluppa a partire dall'esperienza che il parlante fa della lingua e si manifesta attraverso l'uso; in sintesi, solo l'impiego di co-occorrenze specifiche in occasioni d'uso tipiche crea categorie e strutture a tutti i livelli linguistici e la non (totale) composizionalità del significato o le idiosincrasie sintattiche vengono compensate dalle regolarità pragmatiche (Arduini-Fabbri, 2008: 52-6).

Sulla base di questi argomenti la Linguistica Cognitiva sostiene l'idea dell'automatizzazione del linguaggio e del riutilizzo da parte del parlante di blocchi precostituiti (*chunks, pre-fabricated units, pre-fab*) ed elabora diversi modelli, fra cui la *Construction Grammar* di Fillmore e Kay (1988) e la *Pattern Grammar* di Hunston e Francis (2000). La *Pattern Grammar* riconoscendo l'associazione fra pattern e significato, si pone in un rapporto di continuità con gli studi di Sinclair e del contestualismo; il modello prodotto descrive gli schemi distribuzionali delle parole, ossia i contesti sintattici tipici delle singole unità lessicali ottenuti mediante lo studio delle concordanze nei corpora, esempi di *patterns* sono V+N+Inf ([*Gianni*] *ha visto Maria correre*) e V+come+N ([*Gianni*] *lavora come informatico*); ogni parola possiede un proprio set di

*patterns*, ulteriormente suddivisi secondo le diverse accezioni del lessema stesso (Legallois-Tutin, 2013: 14-5).

### **1.3.1 Langacker e le *symbolic units***

La Linguistica Cognitiva al suo interno prevede più modelli e in questa panoramica si fa riferimento alla versione di Langacker (1987, 1991) (Gries, 2008: 12-5).

La Linguistica Cognitiva elegge la *symbolic unit* a unico elemento contenuto nel sistema linguistico e la definisce:

“A unit is a structure that a speaker has mastered quite thoroughly, to the extent that he can employ it in largely automatic fashion, without having to focus his attention specifically on its individual parts for their arrangement [...]; he has no need to reflect on how to put it together.” (Langacker, 1987: 57, cit. in Gries 2008:13)

Questa descrizione, pur essendo generica, sottolinea che lo status di unità si associa positivamente con il parlante/ascoltatore, il quale sia nella fase della produzione che dell'interpretazione non analizza la struttura interna dell'espressione.

La Linguistica Cognitiva elegge a parametro fondamentale la frequenza d'uso; il modello Langacker è esplicitamente basato sull'uso: innanzitutto perché l'esposizione alle *symbolic units* e il loro uso (*Performance*) modellano il sistema linguistico sia del parlante che dell'ascoltatore, e poi perché la frequenza di co-occorrenza è il requisito necessario per il rafforzamento dello *status* dell'espressione linguistica a unità.

Il contributo di Langacker è significativo anche per aver favorito la transizione delle collocazioni, da una condizione di *periphery* (marginalità) al *core* (nucleo) del sistema linguistico.

### **1.3.2 Linguistica Cognitiva e la teoria dei prototipi**

La Linguistica Cognitiva, interessata a indagare come gli esseri umani categorizzano l'esperienza del mondo/realtà, ha introdotto il concetto di prototipo e la teoria a esso collegata, seppur con alcuni aggiustamenti, ha incontrato una crescente applicazione nell'ambito della Semantica Lessicale che se ne serve per chiarire la molteplicità dei referenti, la polisemia lessicale e l'uso letterale o figurato di una parola.

La Teoria dei prototipi avanzata da Rosch negli anni '70 e successivamente ampliata con i contributi di Lakoff, Geeraerts, Langacker e Fillmore, stabilisce che la categorizzazione delle conoscenze umane è fondata sulla *similarità* (o *familiarità*) rispetto un *prototipo*, definito come il miglior esemplare o *cluster* di migliori esemplari

della categoria selezionati in ragione dell'alta frequenza della loro enunciazione; sulla base del criterio della *similarità*, i membri di una categoria si configurano come centrali o periferici rispetto al prototipo e condividono tra loro una rete di somiglianze parziali; inoltre, ogni elemento può appartenere a categorie diverse, poiché queste, non essendo arbitrarie, possiedono confini labili (*fuzzy*). Il prototipo è un costrutto mentale (vs istanza reale) composto di proprietà prototipiche, che si presenta in forme diverse (*effetti prototipici*) in conformità alla categoria in cui compare ed è descrivibile in termini scalari (Giacomini, 2012: 39-42).

L'adozione della nozione di prototipo da parte della Semantica Lessicale determina passaggio dall'idea del "significato come concetto" a quella di "significato come concetto con struttura prototipica", ossia anche il significato, al pari dei concetti, consta di un nucleo prototipico e di un'area circostante o periferica che mostra effetti prototipici, i quali a un certo punto si esauriscono (Ježek, 2016: 71-2).

Nella Semantica Lessicale la nozione di prototipo è attrattiva, perché spiega l'attitudine delle parole di variare il significato a seconda del contesto in cui occorrono e quindi del loro uso.

### **1.3.3 Psicolinguistica e il *lessico mentale***

Le teorie psicolinguistiche suggeriscono che le MWE sono stoccate come unità nel *lessico mentale*, inteso come:

“il vocabolario interiore che immagazzina forme e significati nella memoria a lungo termine, per accedervi al momento opportuno e nel quale le parole sono connesse fra loro secondo un principio di abitudine linguistica.” (Giacomini, 2012: 114)

Ogni persona adulta, tenuto conto di alcune differenze dovute all'ambiente sociale in cui vive, alla scolarità e agli interessi coltivati, dispone di un patrimonio di vocaboli, una sorta di repertorio sistematicamente organizzato, atto a contenere tutte le informazioni (semantiche, sintattiche, morfologiche, fonologiche, ortografiche e di struttura argomentale) contenute nei vocaboli della lingua utilizzata; l'accesso a questo bagaglio è estremamente rapido ed efficiente, perché ogni parlante è capace di elaborare tre-quattro parole al secondo, indipendentemente dal grado di istruzione e dalla lingua parlata, e correla con la frequenza, cosicché quanto più spesso si incontra una combinazione di parole nel contesto comunicativo quotidiano, tanto più aumenta la sua attivazione in produzione e in comprensione nella memoria.

Il linguaggio consiste di un repertorio lessicale e di un insieme di regole grammaticali, le quali producono non solo sequenze grammaticalmente corrette, ma anche, e per la maggior parte, MWE linguisticamente anomale, frequenti nell'uso e sensibili a particolari contesti dai quali dipendono sia per l'apprendimento che per il riuso, similmente a quanto avviene per le forme lessicali (Denes, 2009: 84-7); solo di recente lo studio del lessico si è orientato sull'elaborazione di combinazioni di parole calate in contesto linguistico e situazionale.

Il modello neurocognitivo del linguaggio proposto da Ullman, derivato da quello di Schacter-Tulving, distingue la memoria in dichiarativa e in procedurale; la memoria dichiarativa è attiva nel processo di apprendimento, trattamento e richiamo delle conoscenze enciclopediche e include il lessico mentale; invece la grammatica è inglobata nella memoria procedurale, alla base dei processi di apprendimento implicito di sintassi, morfologia, alcuni compiti fonologici e di programmazione articolatoria; nel modello di Ullman quindi, il linguaggio nelle sue componenti dichiarativa e procedurale, oltre a essere modulare, partecipa a un sistema più ampio di memoria (Denes, 2009: 33-4).

In ambito psicolinguistico lo studio condotto da Wray sul linguaggio formulaico (Legallois-Tutin, 2013: 17-8) ha il pregio di rilevare l'eterogeneità del lessico mentale, articolato in cinque tipi: 1) grammaticale, 2) referenziale, 3) interazionale, 4) memorizzato e 5) riflessivo.

Il costo cognitivo della produzione di enunciati decresce secondo l'ordine 1-5, e se il lessico grammaticale (tipo 1) si mobilita con gli enunciati liberi, che richiedono una maggiore competenza analitica, al contrario, il lessico riflessivo (tipo 5) fornisce espressioni prodotte con minor carico cognitivo. Il modello proposto da Wray rende conto del fatto che le unità linguistiche sono immagazzinate in modo differente, per es. *buongiorno a voi* è un'unità appartenente al lessico interazionale (tipo 4), ma allo stesso tempo *buongiorno* e *voi* sono unità del lessico referenziale (tipo 2). Una costruzione può appartenere al lessico interazionale (tipo 3) e allo stesso tempo almeno una parte dei suoi costituenti al lessico referenziale (tipo 2), ad es. *Perché non andiamo al cinema?*, oppure *Perché non andiamo a mangiare?*: «perché non andiamo + Prep (+Art.) + N» è una costruzione semi-idiomatica aperta, dato che lo slot N va saturato con un Sostantivo o Verbo sostantivato, mentre *cinema* e *mangiare* appartengono al lessico referenziale.

Da questi esempi si evince che il lessico mentale suddiviso in cinque tipi è essenzialmente interattivo e quindi secondo Wray i contesti tipici in cui molti termini co-occorrono portano i parlanti ad acquisirli come *chunks* (*pre-fabricated, pre-fab units*) inanalizzati, che vengono scomposti in costituenti nelle fasi più tarde dello sviluppo della competenza linguistica e solo se necessario (Benigno-Kraif, 2016: 240-1); la disponibilità di un vasto numero di *pre-fabricated units* presenta i vantaggi di ridurre il carico di processamento, di facilitare la comprensione, di rendere scorrevole il linguaggio e la possibilità di riconoscere l'appartenenza di un parlante a specifiche comunità linguistiche. Questi vantaggi uniti alla frequenza d'uso e alla conoscenza del contesto conducono all'ipotesi che i vengano immagazzinate in modo olistico (Pawly-Syder, 1983: 218).

A supporto dell'idea che le collocazioni e i phrase sono memorizzati e recuperati nel lessico mentale in modo olistico, Sosa-MacFarlane (2002: 227-30, 233-35) mettono in luce che il fattore alla base di questi processi è la frequenza (*collocational* o *string frequency*) con cui gli items co-occorrono nelle abituali conversazioni; nella loro ricerca monitorano i tempi di risposta dell'elaborazione esplicita da parte dei parlanti dei diversi livelli di frequenza delle funzioni (semantiche, sintattiche pragmatiche) della parola target *of* in collocazione con altri termini e rilevano un marcato e significativo effetto della *collocational frequency*: le latenze di risposta alla parola target in collocazione risultano più lunghe/lente quando essa occorre in collocazioni molto frequenti (*kind of, lot of*), viceversa le latenze sono più brevi/veloci quando essa compare in collocazioni poco frequenti o nuove combinazioni (*piece of, example of*); lo scarto nei tempi di risposta è il segnale di una diversa modalità di accesso al lessico mentale della parola target, che nelle collocazioni poco frequenti avviene in modo diretto, mentre in quelle ad alta frequenza l'accesso alla parola target è consentito previa operazione di scomposizione del *chunk*. I maggiori tempi di risposta sono quindi, il segnale che, all'interno degli enunciati, le co-occorrenze più frequenti costituiscono delle unità e queste unità, impedendo l'accesso diretto alla parola target come unità indipendente e obbligandola a essere processata come elemento fuso entro il blocco stesso, da un lato sono responsabili del rallentamento nell'elaborazione e dall'altro confermano la memorizzazione e il recupero olistici dei lessemi nel lessico mentale. L'aumento massiccio della frequenza d'uso caratterizza queste associazioni lessicali come *routines* linguistiche, fino a renderle grammaticalmente anomale e semanticamente opache e in tal caso l'accesso alla parola

target della collocazione avviene esclusivamente attraverso la conoscenza linguistica esplicita (vs. implicita). L'effetto della  *collocational frequency*  è compatibile con un modello di memorizzazione e processamento lessicali basati sull'uso, nei quali la rappresentazione lessicale e la struttura interna del lessico sono determinate dalla lingua d'uso; inoltre, conferma l'ipotesi che le unità più grandi della parola vengono immagazzinate nel lessico mentale in modo olistico e come item indipendenti.

Gli studi psicolinguistici e quelli sull'Intelligenza Artificiale decretano che l'esperienza e l'uso sono superiori alle regole linguistiche: in Psicolinguistica emerge chiara l'idea che il cervello umano è meglio equipaggiato per la memorizzazione che per il processamento (Nesselhauf, 2005: 2), mentre nell'ambito dell'Intelligenza Artificiale è stato dimostrato che i parlanti elaborano le conoscenze sulla base di esperienze ripetute.

#### **1.4 Approccio adottato nel presente lavoro**

Le collocazioni sono un fenomeno linguistico pervasivo nella lingua, tanto da risultare sfuggenti e non sempre riconoscibili nell'immediato sia dai parlanti madrelingua che dai lessicografi, perché all'apparenza sembrano una normale combinazione di parole dettata dalla sintassi e il loro notevole valore nella lingua scritta e parlata si riscontra nel confronto interlinguistico o negli spogli dei corpora linguistici elettronici.

Il presente lavoro muove dall'ipotesi che la descrizione dell'italiano offerta nei dizionari generali e specialistici accusi un certo grado di soggettività e sia fondata più sull'intuizione e sull'introspezione del compilatore, piuttosto che sulla rigorosa osservazione dell'uso reale che i parlanti fanno della lingua; la parola scelta a paradigma delle collocazioni è  *notizia*  e la ricerca instaura un raffronto tra l'uso effettivo della lingua attestato in un corpus di riferimento e ciò che è recensito nei repertori, allo scopo di verificare ed eventualmente ampliare i  *pattern collocazionali lessicali*  citati nei glossari stessi; un  *pattern*  (modello, schema ricorrente, struttura ripetitiva) è la configurazione regolare assunta da particolari sequenze di parole all'interno di un insieme di risultati sperimentali e in ambito linguistico è la struttura sintagmatica minima nella quale tutte le parole sono stabili nel senso, o meglio non ambigue (Ježek, 2016: 77).

L'approccio adottato in questo lavoro è statistico-quantitativo in parte mitigato da quello fraseologico: il trattamento quantitativo delle collocazioni della Corpus Linguistics



(1.1.3) si intreccia con alcuni criteri (morfo)sintattici (ad es. la flessione della parola, l'ordine dei lessemi) e semantici idonei a distinguere i vari tipi di *pattern collocazionali*.

Il primo modello è calibrato sulla frequenza (*frequency-oriented approach*) e permette di estrarre automaticamente le co-occorrenze di parole dai corpora digitali di grandi dimensioni, utilizzarne statisticamente i risultati e condurre osservazioni empiriche in ottica esplorativa (*corpus driven*) e/o confermativa (*corpus based*); questo lavoro opta per l'approccio *corpus-driven*, che procede all'analisi di un corpus secondo un metodo induttivo, ossia partendo dai dati per sviluppare teorie e ipotesi linguistiche. Indubbiamente questo approccio restituisce un materiale grezzo (*collocazioni empiriche* di Firth), deprivato dei tratti delle circostanze in cui avviene l'evento comunicativo (*context of situation*) e nulla o poco rivela circa gli autori e i destinatari del messaggio; inoltre, contiene tutti i tipi di MWE, oltrech  le combinazioni libere molto usate, ma ha il pregio di fornire dati puntuali, perch  ottenuti dalla lingua realmente in uso presso una comunit  linguistica; il *caveat* da tener presente   che un corpus, indipendentemente dalla taglia e variet , alle basse frequenze di co-occorrenza   sempre rappresentativo solo di s  stesso, quindi   auspicabile non sovrastimarne le possibilit : il fatto che un item non sia presente in un corpus non prova che esso non esista, semplicemente non   cos  frequente nel tipo di testi collezionati in quel corpus (Partington, 1998: 146); uno dei maggiori pregi dell'utilizzo dei corpora   di aver agevolato il passaggio delle collocazioni da una zona di anomalia e periferia della lingua alla centralit  del lessico.

L'analisi fraseologica, non rende conto di tutti i possibili casi di collocazioni, le quali sono un fenomeno non prevedibile facendo ricorso solamente al significato (cangiante in base al contesto e determinato dalla specificit  dei partners collocazionali), al lessico (ambiguo e polisemico quando usa le parole in isolamento) e alla grammatica (spesso anomala riguardo la selezione lessicale imposta dal predicato sugli argomenti); la descrizione qualitativa, pur dimostrandosi frammentaria e insufficiente, resta comunque uno strumento imprescindibile.

L'integrazione tra l'approccio quantitativo e quello fraseologico funge da avallo teorico adeguato a formulare una definizione funzionale di collocazione e procura un apparato descrittivo adatto a individuare e cogliere le relazioni e il funzionamento dei costituenti in collocazione.



## Capitolo 2

### Definizione funzionale, strumenti e metodologia di ricerca

Dopo aver introdotto nel Cap. 1 gli approcci teorici allo studio delle collocazioni e aver individuato il modello statistico-quantitativo come il più adatto allo scopo della ricerca, nel Cap. 2 viene formulata una definizione funzionale di collocazione, si illustrano gli strumenti utilizzati (repertori e corpus) e il metodo applicato per compiere il raffronto tra i pattern collocazionali emersi dallo spoglio del corpus e quelli citati nei glossari.

La scelta del lessema su cui condurre la ricerca è ricaduta su *notizia* ed è motivata da vari fattori: innanzitutto sulla base delle stratificazioni statistiche e di uso rientra nel “vocabolario fondamentale, formato da circa 2.000 vocaboli di massima frequenza, che coprono da soli in media il 90% di qualunque testo scritto o parlato [...] e sono noti a qualunque italiano in possesso di almeno un’istruzione elementare, il 90% della popolazione” (Casadei, 2003: 132); in secondo luogo, appartiene alla classe grammaticale dei sostantivi, i quali “si riferiscono a gruppi di proprietà correlate che rispecchiano una porzione dell’esperienza percettiva di ogni individuo. I Verbi, invece, rappresentano relazioni tra le entità circoscritte dai nomi riassumendo tutte le informazioni che tali relazioni implicano [...], inoltre per il nome si può parlare in generale di una più frequente semplicità concettuale in termini di autonomia semantica.” (Giacomini, 2012: 51); è un lemma che nel Dizionario delle Collocazioni (Tiberii, 2018) riceve una voce ricca di lessemi in combinazione; infine, nell’era della comunicazione e dell’informazione, le collocazioni di *notizia* sono molto diffuse e possono andare incontro a ulteriori sviluppi.

#### 2.1 Definizione funzionale di collocazione lessicale e proprietà delle collocazioni

L’interesse multidisciplinare suscitato dalle *collocazioni*, se da un lato testimonia la vitalità dell’argomento, dall’altro ha prodotto una proliferazione di definizioni del termine, fino a farne un “umbrella/cover term” che indica le combinazioni di parole tipiche di una lingua e vantaggiose per la produzione e l’interpretazione del messaggio; il fenomeno sembra non essere descrivibile in modo universalmente valido, eppure la debolezza, la vaghezza e l’instabilità inerenti al concetto offrono il vantaggio di adattare

la definizione allo scopo del lavoro e colmare in questo modo il divario tra teoria linguistica e prassi lessicografica.

Sulla base di queste premesse e dell'approccio adottato, la definizione di *collocazione lessicale* che meglio si adatta a questo studio riprende quella di Jones – Sinclair (1.1.3), che la descrivono come co-occorrenza di due unità lessicali in un testo all'interno di un contesto specifico e i cui elementi costitutivi sono: node, span e collocational pattern; inoltre, la collocazione è: “una combinazione limitatamente compositiva, limitatamente sostituibile e limitatamente modificabile” (Giacomini, 2012: 112-3). Si tratta di una definizione statistica, fondata sul parametro della frequenza (Jones-Sinclair), integrata da delle regole di restrizione lessicale basate su una solidarietà consolidata dall'uso.

Tutte le MWE sono presenti nel lessico mentale come unità coese, caratterizzate da familiarità e culturalmente tipizzate, ma dalla definizione appena fornita è possibile dedurre le specificità linguistiche ed extra-linguistiche che contraddistinguono le collocazioni nel panorama delle combinazioni di parole.

Le principali qualità extralinguistiche delle collocazioni sono:

- frequenza (valore statistico): la co-occorrenza dei costituenti in collocazione mostra una maggior frequenza rispetto alla frequenza con cui questi si presentano nel contesto di altri lessemi (*delitto efferato, erogare un mutuo*);
- convenzionalità (*salience*) (valore distribuzionale): la forza di coesione esistente tra determinate parole. Ad es. *pioggia forte* vs *pioggia torrenziale*: *forte* è l'Aggettivo più usato per esprimere l'intensità dell'evento atmosferico, ma *pioggia torrenziale* è più convenzionale e consolidata rispetto la prima, perché nonostante *forte* compaia con *pioggia* più spesso rispetto a *torrenziale*, si combina anche con un'ampia gamma di parole, mentre *torrenziale* occorre con un numero ristretto di vocaboli, uno dei quali è *pioggia*; la spiccata tendenza di *torrenziale* a combinarsi con *pioggia* piuttosto che con altre parole è più convenzionale o istituzionalizzata rispetto a quella di *forte*. La forza di coesione tra lessemi viene stabilita attraverso le misure di associazione lessicale (2.4).

Le caratteristiche linguistiche sono:

- relazione *base-collocato*: i due elementi stabiliscono un rapporto semantico gerarchico tale per cui il primo è dominante e funge da innesco per il secondo; in

ambito distribuzionale costituiscono rispettivamente *node* e *collocate* e si trovano in uno *span* di  $\pm 4$ ; lo *span* è la finestra predefinita di contesto fraseologico, cioè la distanza massima in termini di stringhe di caratteri (in questo caso 4) che separa il *node* dai *collocati*, i quali possono precederlo o seguirlo;

- configurazione in *pattern collocazionali*: lo studio si concentra sulle *collocazioni lessicali*, combinazioni di lessemi che si profilano in una serie di pattern; un pattern (modello, schema ricorrente, struttura ripetitiva) è la configurazione regolare assunta da particolari sequenze di parole all'interno di un insieme di risultati sperimentali e in ambito linguistico è la struttura sintagmatica minima nella quale tutte le parole sono stabili nel senso, o meglio non ambigue (Ježek, 2016: 77). I tipi di strutture (morfo)sintattiche ricorrenti sono: N+Agg, Agg+N, V+Compl, Sogg+V, Avv+V, Avv+Agg. Un tipo particolare di Pattern V+Compl sono le costruzioni a Verbo Supporto (*Light Verb Construction*) (3.3);
- limitata composizionalità: il significato della collocazione è generalmente semi-composizionale, caratterizzato da una relativa trasparenza e desumibile (componibile) dai significati dei membri che la costituiscono e dalle regole sintattiche con cui essi si combinano. La *base* mantiene la trasparenza semantica, mentre il *collocato* esprime un uso figurato del significato (ad es. *sorpresa amara*), per cui si verifica solo la parziale violazione del *principio di composizionalità*, riconducibile a Frege;
- limitata sostituibilità: la sequenza dei lessemi è stabile, non casuale, e permette di valutare la variabilità distribuzionale; le collocazioni si distinguono per la *fissità* lessicale, per cui la sostituzione lungo l'asse paradigmatico con un elemento della stessa categoria genera un'espressione inappropriata ma sanabile con la sostituzione della parola che viola la restrizione lessicale con quella consolidata dall'uso, ad es. *?scolare le patate vs scolare la pasta*, dato che solo la *pasta* può essere sottoposta all'azione espressa dal Verbo;
- limitata modificabilità (morfo)sintattica: questa caratteristica introduce svariati comportamenti: la possibilità o meno di modificare l'ordine dei costituenti e/o inserire materiale linguistico tra essi e/o bloccare la flessione e di conseguenza i membri, essendo parti variabili del discorso, si cristallizzano e perdono la propria flessione;

- autonomia sintattica: gli item della collocazione sono sintatticamente autonomi e ammettono le trasformazioni tipiche degli elementi delle combinazioni libere, ad es.

la collocazione *bandire un concorso*, può essere sottoposta a:

passivizzazione *il concorso è bandito dal Comune;*

ripresa anaforica *il concorso, il Comune lo ha bandito;*

topicalizzazione *il concorso ha bandito il Comune;*

relativizzazione *il concorso che il Comune ha bandito,*

*il concorso che è stato bandito.*

Le collocazioni sono abbinamenti preferenziali di parole che i parlanti a qualsiasi livello di competenza utilizzano per veicolare determinati concetti, nonostante il sistema linguistico offra loro altre combinazioni possibili, ad. es. in *bandire un concorso*, per esprimere l'atto di annunciare pubblicamente, *concorso* (la base) si abbina di preferenza con il Verbo specifico *bandire* (il collocato), anziché con altri plausibili, come *pubblicare* e il significato acquisito nella combinazione specifica non è altro che *il meaning by collocation* di Firth (1.1.1); sulle collocazioni si applica una regola di restrizione lessicale consolidata dall'uso, la cui violazione pur producendo un effetto di inappropriatezza, è sanabile con la sostituzione della parola che viola la restrizione lessicale con quella consolidata dall'uso (Ježek, 2016: 193), come nell'esempio riportato sopra *?scolare le patate vs scolare la pasta*.

Il tipo di restrizione lessicale a cui sono sottoposte le *collocazioni lessicali* è particolarmente significativo per distinguerle dalle *solidarietà semantiche lessicali*, teorizzate da Coseriu nel 1971; entrambi i fenomeni esibiscono una solidarietà fra lessemi, ma nel primo caso è istituzionalizzata dall'uso, mentre nel secondo origina da restrizione ontologica e un'implicazione sintagmatica di contenuto, ad es.:

*Collocazioni lessicali:*

da solo *bandire* non implica *concorso*

*sporgere* non implica *denuncia*

*florida* non implica *azienda*

*amara* non implica *sorpresa*

*Solidarietà semantiche lessicali:*

anche da solo *indossare* implica *vestito*

*miagolare* implica *gatto*

*aquilino* implica *naso*  
*riccio* implica *capello*

Nelle *solidarietà semantiche* la connessione fra i due termini è salvaguardata anche quando il collocato, monosemico, compare in isolamento (da solo *miagolare* implica sempre *gatto*), diversamente da quanto avviene nelle collocazioni (da solo *sporgere* non implica *denuncia*), nelle quali il collocato è polisemico; inoltre, la restrizione lessicale opera in direzioni diverse: nelle solidarietà semantiche è imposta dal Verbo o Aggettivo al Nome, viceversa nelle collocazioni è imposta dal Nome al Verbo o Aggettivo (Ježek, 2016: 201-2).

## **2.2 Repertori e trattamento lessicografico delle collocazioni**

La letteratura sul trattamento lessicografico delle collocazioni sottolinea la necessità e l'utilità della compilazione di dizionari speciali per le combinazioni di parole e della loro inclusione nei dizionari generali (Orlandi 2016, Torner-Bernal 2017), ma allo stesso tempo ne evidenzia il trattamento non sistematico e la selezione affidata più all'introspezione del lessicografo anziché allo spoglio di un corpus.

La ricognizione dei pattern collocazionali è stata condotta, in primo luogo, su una serie di repertori, generali o specifici; si tratta di dizionari con ordinamento semasiologico e in formato cartaceo monolingui generali, combinatori, analogici e un dizionario elettronico bilingue italiano-tedesco.

La fonte primaria per il reperimento dei pattern è stato il Dizionario delle Collocazioni (DdC, Tiberii, 2018), al quale sono stati affiancati altri glossari al fine di allargare la base dei dati; DdC è anche il testo di riferimento per operare il confronto tra uso e descrizione della lingua, perché essendo l'unico a occuparsi specificatamente dell'argomento è il più selettivo all'interno del panorama lessicografico italiano, inoltre è il più recente tra i dizionari speciali di MWE. La tabella inserita a fine paragrafo riporta le definizioni di *notizia* recensite nei vocabolari e i pattern collocazionali rintracciati nei repertori nella tabella riportata in Appendice.

### **Dizionari di combinazioni lessicali**

I quattro dizionari di collocazioni disponibili per l'italiano si rivolgono innanzitutto a utenti madrelingua che intendono raggiungere una maggiore efficacia comunicativa

attraverso l'arricchimento e l'appropriatezza del linguaggio, e tre di questi (DdC, Lo Cascio e Urzi) fanno specifico riferimento ai non madrelingua, per i quali la produzione delle collocazioni nella lingua di arrivo rappresenta una difficoltà inevitabile.

Nei quattro dizionari, le *teste* appartengono alle categorie grammaticali dei Nomi, Aggettivi e Verbi, conformemente alle classificazioni comunemente accettate (Benson-Benson-Illson, 1986), le quali prevedono che solo queste tre parti del discorso possano fungere da *base* della collocazione; solo DdC cataloga le collocazioni in riferimento alla *base*, mentre gli altri tre le registrano sia sotto la *base* che sotto il *collocato*; il criterio di Tiberii si dimostra più rigoroso ed evita le ridondanze, ma non agevola la consultazione ai non esperti della lingua.

Dato che in italiano esistono molte eccezioni alle regole generali, i quattro dizionari non solo segnalano i contesti dove l'uso del sostantivo esige il singolare o il plurale, ma forniscono anche le indicazioni sintattiche che governano le sequenze lessicali. Questa informazione diventa essenziale nel caso del pattern Verbo+Complemento, dove il Nome tipicamente ricopre la funzione di Complemento Oggetto o Complemento Indiretto; nel primo caso i repertori avvertono se il sostantivo prende l'articolo determinativo, indeterminativo o è usato senza articolo, mentre nel secondo indicano le *reggenze* del Verbo per mezzo delle Preposizioni.

Nell'introduzione che segue per ciascun glossario si mettono in rilievo i criteri generali che hanno guidato la selezione e il trattamento delle collocazioni.

*Dizionario delle collocazioni. Le combinazioni delle parole in italiano* (DdC, Tiberii, 2018): esclude le collocazioni rare o estremamente specifiche e quelle troppo comuni e generali; la seconda edizione del 2018 (che fa seguito alla prima, del 2012) chiarisce il metodo della selezione:

“e revisione [delle collocazioni], oltre alla consultazione di banche dati e lessici specialistici, ci si è serviti di un sistema di validazione ed estrazione automatica delle collocazioni basato su corpora [PAISÀ], con successiva valutazione e selezione manuale dei risultati ottenuti attraverso metodi statistici.” (DdC: 4)

All'interno di ogni voce i pattern sono raggruppati in base all'uso e alla categoria grammaticale di appartenenza; ciascuna categoria segue l'ordine alfabetico e segnala l'uso esclusivo del sostantivo al singolare o plurale; per i nomi vengono indicate le collocazioni con Aggettivi e Verbi, questi ultimi distinti in base alla funzione sintattica svolta (Verbo+Complemento e Soggetto+Verbo) e le Costruzioni (locuzioni e usi



fraseologici). Il maggior pregio di DdC è la distinzione degli Aggettivi a seconda della loro posizione, libera o fissa, mentre il suo limite è il non fornire esempi, fondamentali per i non madrelingua.

*Dizionario Combinatorio Italiano* (DCI, a cura di Lo Cascio, 2013): per rendere completo il quadro delle preferenze combinatorie di una parola, DCI registra sia le collocazioni ristrette che le forme idiomatiche; le combinazioni sono sotto-classificate in base alla categoria grammaticale a cui appartiene la parola che si combina con l'entrata lessicale e sono elencate in ordine alfabetico; riguardo la posizione degli Aggettivi, qualora questa sia libera, il DCI presenta quella che ricorre più frequentemente in un motore di ricerca Internet (Google). Il metodo di selezione delle MWE si è avvalso di "diverse fonti: dalle conoscenze e intuizioni dei parlanti nativi ai dizionari italiani monolingui, al corpus Sketchengine, ma soprattutto Internet [...]. Si è fatto inoltre uso di una banca dati realizzata in più di 30 anni per la redazione di un dizionario elettronico bilingue" (DCI: xxiii). DCI correda l'articolo con una breve definizione e ove necessario si avvale di marche diafasiche, per segnalare registri e sottocodici, etichette diatopiche, brevi spiegazioni se sussistono difficoltà di comprensione ed esempi del contesto d'uso.

*Dizionario delle combinazioni Lessicali* (DCL, Urzi, 2009): "prende in esame tutte le Combinazioni Lessicali per le quali, data per nota la base (Nome, Verbo o Aggettivo), la modalità di ricerca orizzontale (sintagmatica) della parola da abbinare risulti *più rapida ed efficiente di qualunque altra modalità di ricerca* (ricerca per sinonimi, mediante i dizionari tradizionali o altro)" (DCL: 3); DCL opera una distinzione semantica delle collocazioni e tramite appositi discriminatori di significato, raggruppa i collocati che hanno tra loro rapporti paradigmatici di sinonimia e antonimia. La ricchezza di esempi "opportunamente adattati, che sono stati tratti da fonti giornalistiche e da Internet" (ibidem) è indubbiamente il pregio di questo glossario, ma l'affermazione "si può pertanto considerare in larga misura *corpus-based*" (ibidem) risulta inopportuna, perché se da un lato Internet non è una collezione ragionata di testi, dall'altro il DCL non ricorre all'uso sistematico dei corpora, comprovato anche dall'omissione del criterio di selezione delle collocazioni, che pertanto potrebbe essere fondato in larga parte sull'introspezione del lessicografo.

*Modi di Dire, Lessico Italiano delle Collocazioni* (MdD, Russo D., 2010): non riporta alcuna indicazione sul metodo di selezione delle combinazioni. La tipologia delle

stesse, diversamente dagli altri dizionari combinatori, non è condotta sulle categorie grammaticali, ma sui criteri di frequenza d'uso e coesione interna, tuttavia non specificando se per coesione intende modificabilità sintattica o sostituibilità paradigmatica; questo sistema produce quattro tipi di collocazioni (A,B,C,D) determinati da un rapporto di inversa proporzionalità tra i due parametri: collocazioni di alta frequenza sarebbero caratterizzate da bassa coesione e viceversa, ma si tratta di una correlazione arbitraria o non rigidamente definibile, dato che sono due variabili indipendenti; un ulteriore punto critico di MdD è l'assenza di esempi.

### **Dizionari generali dell'uso**

I dizionari generali registrano la lingua nella dimensione sincronica, cioè nel funzionamento e nei caratteri attuali, ma solo recentemente hanno rivolto l'attenzione agli aspetti fraseologici; anche se la tendenza attuale privilegia gli esempi fondati sull'uso più che le citazioni letterarie, sono ancora pochi quelli che isolano in modo sistematico le collocazioni distinguendole chiaramente dalle combinazioni libere, mentre la maggior parte di essi offre informazioni generiche e poco gestibili da parte del fruitore, dato che cela il fenomeno tra gli esempi. Sul versante metodologico tutti i repertori generali dichiarano di essere orientati dal criterio dell'uso, ma nessuno di loro palesa le fonti da cui ricava la fraseologia esemplificativa.

*Nuovo Devoto-Oli: il vocabolario dell'italiano contemporaneo* (DO, 2018): l'indicazione delle reggenze sintattiche ha comportato un arricchimento della fraseologia e la versione on-line mette in risalto le espressioni idiomatiche, le polirematiche e la terminologia, ma negli esempi le collocazioni si confondono con le combinazioni libere;

*Grande Dizionario Italiano dell'Uso* (GraDIt, 2000): è l'unico dizionario generale che riserva uno spazio dedicato alle MWE: le collocazioni assieme alle espressioni polirematiche, a quelle idiomatiche e ai termini tecnici, vengono elencate alla fine della voce nella sezione "Polirematiche"; in questa sezione, oltre che alcuni tecnicismi, è citato solo *fare notizia*, mentre tra gli esempi delle accezioni compaiono combinazioni molto frequenti.

*Palazzi-Folena* (PF, 1992): nella Presentazione si qualifica come "un vocabolario dell'uso vivo" e ricorda che uno dei caratteri genetici dell'opera è "la ricchezza colloquiale e locuzionale, per cui la parola veniva presentata nei suoi plurimi valori d'uso

dentro un contesto [...], si è valorizzata al massimo quella che era una delle caratteristiche più positive del dizionario: l'esemplificazione viva di locuzioni, modi di dire, costrutti attestati attraverso l'uso odierno" (PF, 1992:4-5). Il PF dichiara di abbondare nella fraseologia (locuzioni particolari e proverbi) e di procurare gli esempi sia con l'intento di mostrare la parola usata nei contesti che meglio ne risaltano il senso, sia per illustrare i comportamenti sintattici possibili, ma, nonostante s'intuisca un criterio semantico nel loro ordinamento, appaiono caotici nella presentazione, per cui alla prima accezione di *notizia* si trova "dare, diffondere una notizia, buone, cattive notizie, leggere, ascoltare le ultime notizie; fare notizia, essere tale da suscitare l'interesse generale: un matrimonio che farà notizia"; un punto di forza delle voci in PF è la ricca sezione "Nomenclatura", nella quale sono delimitati i vocaboli che possono figurare assieme alla parola-lemma, tipicamente Aggettivi e Verbi nel caso dei Sostantivi.

*Sabatini-Coletti* (SC, 2013): fin dalla sua prima edizione del 1997, si prefigge di aiutare l'utente a padroneggiare l'uso della lingua, ma da questo punto di vista non adotta un criterio coerente dato che mette in evidenza le espressioni idiomatiche, ma nasconde le collocazioni, parte relegandone tra gli esempi e parte elencandole a fine articolo in "Combinazione con i verbi", dove sebbene sia praticata la distinzione del ruolo sintattico del Nome (Soggetto o Complemento), i pattern collocazionali non sono sottoposti a un sistematico isolamento.

*Il nuovo Treccani* (NTr, 2018): l'ultima edizione è del 2023, però è stata pubblicata a spoglio già effettuato. Uno dei motivi per cui si dichiara *nuovo* è per l'accoglienza riservata di esempi tratti dall'uso reale della lingua; il formato cartaceo segnala con un pallino colorato le espressioni e frasi fatte (solo *fare notizia*) delle quali viene sciolto il significato e riportato un esempio ben contestualizzato, ma allo stesso tempo *per tua, per vostra notizia*, non sono segnalate come fraseologismi; la versione on-line del vocabolario Treccani abbonda di esempi e citazioni d'autore, che illustrano le sfumature di significato e specifica se l'uso in combinazione richiede la forma singolare o plurale (ad es. *correva la notizia che...vs leggere le notizie del mattino*).

*Zingarelli* (Zin, 2022): l'attenzione per il contesto e l'uso più appropriato delle parole si rintraccia nelle schede (circa mille) dedicate alle "sfumature di significato" dove vengono trattate le famiglie di parole analoghe; l'esemplificazione delle accezioni del vocabolo è carente e disomogenea, poiché nella prima accezione si trova solo il pattern

Verbo+Complemento (Oggetto e Indiretto) e nella terza, tranne un caso, solo quello Nome+Aggettivo; entro le diverse accezioni isola e rende graficamente evidenti gli usi idiomatici, ma cela le collocazioni.

*Grande dizionario Hoepli italiano* (GDIH, 2015): afferma di esporre e spiegare, non semplicemente definire, le voci anche “attraverso una ricca rassegna di esempi tratti dall’uso e alla contestualizzazione del vocabolo in esame [...], tenendo in considerazione gli aspetti dinamici della parola, cioè le variazioni flessive che questa subisce concretamente nel momento dell’uso” (GDHI: ix-x) utili a chiarire i rapporti che essa intrattiene con le altre parole; in GDIH l’esemplificazione delle diverse accezioni è molto essenziale, limitata a un esempio.

### **Dizionari analogici**

Le parole di una lingua formano una rete e i dizionari analogici si propongono di elencare una serie di voci-guida attorno alle quali vengono raggruppate le parole o espressioni a essa collegate e per questo motivo si sono rivelati uno strumento utile per individuare gli Aggettivi e i Verbi connessi con il lessema in esame; anche se generalmente nella microstruttura non è contemplata la definizione del lemma, questa è presente nel *Thesaurus Treccani* e nel dizionario di Simone.

*Dizionario dei sinonimi e dei contrari analogico e nomenclatore* (SCAN,1993): in uno specifico raggruppamento, chiamato genericamente “nomenclatura”, raccoglie un’ampia serie di Aggettivi e Verbi che hanno con l’esponente un riferimento nel momento dell’uso e risultano preziosi per arricchire un’espressione e necessari per costruire correttamente un enunciato; fra questi raggruppamenti, il limite maggiore si riscontra nella categoria dei Verbi in combinazione con la voce-guida, perché l’elencazione strettamente alfabetica e priva delle reggenze, non separando i predicati in cui il Nome funge da Soggetto da quelli in cui è Complemento, in qualche caso genera confusione e dubbi circa il pattern che si viene a creare (ad es. il Verbo “confermare”: *confermare la notizia vs la notizia conferma*).

*Thesaurus Treccani, Dizionario analogico della lingua italiana* (ThTr, 2018): l’entrata lessicale si compone di una mappa, che illustra e rende visualmente disponibili le relazioni di significato (in questo lavoro si considerano solo le categorie Azioni-Verbi e Qualità o Relazioni), completata poi da tre elementi: “Presentazione” della parola e dei

suoi significati che ricorrono nella comunicazione corrente corredati da esempi, un riquadro dedicato alle “Parole espressioni, modi di dire” in cui è presente la parola-cardine e una “Citazione d’autore”. L’inclusione di ThTr a fianco di NTr nella ricognizione dei pattern collocazionali è motivata dal fatto che la definizione del significato e le relative esemplificazioni ricevono un’elaborazione originale, nonostante entrambi siano editi dall’Istituto dell’Enciclopedia Italiana.

*Grande Dizionario Analogico della Lingua Italiana* (DAU, 2010): le sezioni Verbi e Aggettivi sono articolate in “campi”, che recano nel titolo l’indicazione della natura della relazione che lo collega al lemma (ad es. “Aggettivi che si combinano con –”); i “campi” a loro volta sono distinti in “segmenti” che riuniscono le parole tra loro semanticamente affini. Entrambe le sezioni offrono un’ampia gamma di termini, che si combinano con l’esponente e questo sistema si rivela particolarmente efficace nella presentazione dei Verbi, segmentata in: “Azioni compiute con –”, “Azioni di -”, “Azioni in cui – è strumento”. L’inconveniente di DAU è la citazione degli Aggettivi con la forma del lemma, per cui nel caso di *notizia*, viene a mancare la coordinazione della flessione; viceversa, un punto a suo favore è di dedicare l’ultimo “campo” ai fraseologismi distinti in due “segmenti”, dove nel primo trovano posto locuzioni e proverbi e nel secondo le combinazioni tipiche di parole con *notizia*.

### **Dizionario elettronico ELDIT**

*Elektronisches Lernerwörterbuch Deutsch-Italienisch* (ELDIT, 2002): il programma rappresenta un prototipo di risorsa lessicale e lessicografica on-line creata per apprendenti delle lingue tedesca e italiana; consente di reperire informazioni di varia natura per il vocabolo ricercato, tra cui le combinazioni di tipo collocativo e idiomatico; purtroppo, sotto questo aspetto la risorsa si è dimostrata inservibile, perché ad oggi questa sezione non è ancora stata compilata. Nella schermata principale accanto alla definizione seguita da esempi contestualizzati compaiono una serie di altri esempi del pattern Verbo+Nome tratti dallo spoglio del corpus di ELDIT seguiti da una lista di Aggettivi che frequentemente co-corrono con il Nome; il vantaggio della risorsa è la disponibilità di un corpus composto da 400 testi, che è possibile interrogare attraverso la ricerca avanzata e che nel presente

lavoro ha consentito l'ulteriore ampliamento della ricognizione dei pattern collocazionali.

<b>Dizionario</b>	<b>Definizioni</b>
DCI	Informazione, comunicazione, delucidazione relativa a fatti, avvenimenti, ecc..
MdD	Informazione su qcn/qcs, comunicazione, nozione, conoscenza
DO	1 Informazione su persone o eventi appresa in prima persona, comunicata da qualcuno o diffusa tramite mezzi di comunicazione 2 Informazione, documento, testimonianza su persone, popoli o eventi del <i>presente</i> 3 ; • Nozione, spec. di carattere scientifico; • Breve comunicazione, bibliografica o di cronaca, in pubblicazioni periodiche a carattere scientifico 3 LETT. Conoscenza specifica, riconducibile sia al concetto di 'informazione' che a quello di 'nozione' 4 N. di reato, la conoscenza di un fatto da parte dell'autorità giudiziaria che dà l'avvio all'accertamento di eventuali reati
GraDIT	1a informazione su qcs. o qnc. 1b comunicazione di un fatto di pubblico interesse, spec. recente, diffusa attraverso i mezzi di informazione 1c informazione, ragguglio pubblicato su una rivista specialistica 1d nota apposta dall'autore di un testo per fornire indicazioni o chiarimenti su un'opera 2 nozione, conoscenza, dato specifico 3 CO nozione relativa a una disciplina, a un'arte o a un'attività pratica   OB solo sing., competenza in una data materia 4 LE familiarità, dimestichezza
PF	1. annuncio di un avvenimento, un fatto, gen. attuale e di qualche interesse   ragguglio   informazione, delucidazione 2. lett. conoscenza
SC	1 informazione relativa a persona o un avvenimento perlopiù recente, appresa direttamente o tramite altri; in partic. (e spec. al pl.) informazione di interesse generale trasmessa dai mezzi di massa 2 informazione su fatti del passato tramite tracce tangibili o memoria; nozione, perlopiù di carattere scientifico 3 lett. conoscenza, cognizione di qlco.
NTr	1.a. Informazione su qualcosa o qualcuno b. Annuncio, informazione data al pubblico attraverso i mezzi di comunicazione, specialmente di fatti accaduti di recente 2. N. di reato, nel linguaggio giuridico, relativa a fatti che costituiscono reato 3.a. Conoscenza come sapere acquisito, in quanto se ne conserva traccia, o ne è trasmessa o recepita la memoria; informazione b. Nel linguaggio letterario, conoscenza
Zin	1 lett. cognizione, conoscenza 2 nozione 3 informazione relativa a un fatto spec. recente 4 comunicazione, ragguglio, spec. su riviste tecniche o specialistiche 5 dir. n. di reato, informazione che il pubblico ministero o l'autorità giudiziaria acquisisce in merito a fatti penalmente rilevanti 6 nell'alto Medioevo, documento probatorio di un'azione giuridica perfetta in sé stessa
GDIH	1 Cognizione, conoscenza 2 Oggetto di conoscenza   Nozione 3 Informazione, annuncio, spec. di avvenimenti recenti, per mezzo di giornali, televisione e sim.   Comunicazione che porta a conoscenza di qualcosa 4 Breve comunicazione in pubblicazioni a carattere scientifico   Cenno, menzione 5 ST In epoca medievale, dichiarazione che attestava un'azione giuridica perfetta in sé stessa
ThTr	1. Una NOTIZIA è un'informazione su di sé o su altri, data da chi è direttamente coinvolto oppure riferita o ricevuta tramite altre persone 2. Generalmente una notizia si riferisce a fatti di interesse pubblico, per lo più recenti, ed è comunicata dalla stampa o da altri mezzi di comunicazione di massa (radio, televisione, Internet) o comunque resa nota con qualche altro sistema 3. Una notizia può anche essere un'informazione relativa a fatti di interesse più limitato e specialistico, pubblicata specialmente in periodici, riviste scientifiche o bollettini; spesso il termine viene utilizzato all'interno di titoli 4 La parola notizia, infine, può indicare una conoscenza relativa a fatti vicini o lontani nel tempo, di cui si conserva traccia o si trasmette la memoria
DAU	Informazione per lo più recente, relativa a qno e, in partic., resa pubblica attraverso una fonte di informazione (spec. un giornale e sim.)
ELDIT	1 Il racconto o la comunicazione di un fatto o un avvenimento, specialmente se è avvenuto da poco tempo (trad. Nachricht) 2 Ciò che si conosce riguardo determinati argomenti (trad. Angaben)

### 2.3 Corpus utilizzato per il censimento dei pattern collocazionali

Un *corpus* linguistico, al plurale *corpora*, è una collezione di dati testuali autentici in formato elettronico, provenienti da fonti scritte, orali o multimediali, rappresentativi di una lingua o una varietà di essa, al fine di condurre indagini linguistiche empiriche ripetibili, sia quantitative che qualitative.

Per questa ricerca è stato utilizzato utilizza il corpus CORIS (CORpus dell'Italiano Scritto Contemporaneo), sviluppato presso il Centro Interdipartimentale di Linguistica Teorica e Applicata dell'Università di Bologna da F.Tamburini sul modello dei grandi corpora elettronici delle lingue europee (ad es. Brown University Standard Corpus of Present-Day American English, o BROWN, e il British National Corpus, o BNC) con la finalità di costituire un insieme di testi informatici rappresentativi dell'italiano attuale.

Il CORIS è un corpus generale di riferimento sincronico di lingua scritta, i cui testi costitutivi sono stati raccolti tra il 1980 e il 2000 per una base originaria di 150 milioni di *tokens* (per *token* si intende la singola unità linguistica o l'elemento linguistico minimo separati da uno spazio, detto *blank* e va specificato che sono calcolati come token anche i segni diacritici) e sottoposto a costanti aggiornamenti con cadenza triennale tramite un corpus di monitoraggio, il cui ultimo aggiornamento risale all'estate 2021 per un totale di 165 milioni di token.

La scelta di utilizzare CORIS è motivata dalla bontà delle caratteristiche costitutive che sono il formato elettronico, l'autenticità dei dati, le dimensioni, la finitezza, la rappresentatività e il bilanciamento (Lenci-Montemagni-Pirrelli, 2005); CORIS è rappresentativo, perché rispecchia qualitativamente e quantitativamente la lingua e contiene al suo interno una distribuzione dei fenomeni linguistici analoga a quella della lingua italiana ed è idoneo quindi alla possibilità di formulare delle generalizzazioni riguardanti l'idioma oggetto d'analisi; è bilanciato, dato che include al suo interno una vasta gamma di tipologie testuali differenti, rappresentative della lingua, provenienti dalla stampa (quotidiani, periodici, supplementi sia nazionali che locali), dalla narrativa (romanzi e racconti italiani e stranieri, sia per adulti che per ragazzi, appartenenti a generi quali poliziesco, di avventure, di fantascienza, delle donne), dalla prosa accademica e giuridico-amministrativa, da vari tipi di volumi e di saggi identificabili nella loro varietà come "miscellanea" e da vari testi a mano, a stampa e in formato elettronico, caratterizzati

dalla loro breve permanenza confluiti nella sezione "ephemera"; i testi sono stati inseriti nella loro totalità (Rossini Favretti -Tamburini -De Santis, 2002).

Il linguaggio di interrogazione (*query*) usato è il Corpus Query Processor (CQP) e l'istruzione per il reperimento delle *concordanze* è esplicitata nella formula generale: "word1" [[]{0,AFT} "word2" (Appendice, 1); i risultati vengono restituiti sotto forma di *concordanze* Key-Word In Context (KWIC) (Appendice, 2) fino a un massimo di 1000 (o, in alternativa, un massimo di 30, default, o 100, o 300), ma qualora le concordanze nel corpus siano più del massimo impostato, se ne vede solo un campione. Le *concordanze* sono le liste di *contesti linguistici* (l'insieme dei *collocates* che co-occorrono con il *node*) in cui una determinata parola (*node*) appare, esse forniscono informazioni sull'uso della parola-perno e agevolano l'individuazione dei casi di uso più frequente. Nel CORIS l'ampiezza del contesto consta di 30 caratteri precedenti e 30 caratteri seguenti; per ogni risultato viene fornita l'indicazione del sottocorpus di provenienza, ma non del testo di provenienza; inoltre, il menu a tendina "Unsorted" consente di ordinare le concordanze in base alle occorrenze a destra e a sinistra della forma cercata (fino a un massimo di 4 parole), in questo caso i segni di punteggiatura vengono trattati come tokens a sé stanti.

Il programma prevede anche la possibilità di visualizzare le classifiche dei collocati, calcolati sfruttando le misure di associazione lessicale, intese come degli indici statistici inferenziali impiegati per la selezione dei dati di un corpus, che originandosi entro la teoria della probabilità formulano previsioni sul comportamento delle parole analizzate e consentono di quantificare l'attrazione reciproca delle occorrenze sulla base del confronto tra *frequenza aspettata* e *frequenza osservata* della co-occorrenza; le misure di associazione inserite in CORIS sono T-score, Log-Likelihood Ratio (LL) e Mutual Information (MI); il ricorso alle misure di associazione intende sopperire al problema dell'insufficienza della *frequenza assoluta* (Raw frequency) di mettere in risalto la capacità di coesione dei costituenti.

## **2.4 Metodologia di ricerca**

Il presente lavoro adotta l'approccio *corpus-driven* (1.4), che procede all'analisi di un corpus secondo un metodo induttivo, ossia partendo dai dati per sviluppare teorie e ipotesi linguistiche, di conseguenza si è fatto ricorso alla versione del corpus non annotato



(*raw corpus*), il quale non contenendo alcun tipo di annotazione non inficia la purezza e l'integrità dei risultati.

Le concordanze consentono di analizzare l'uso del node nei contesti in cui ricorre e sono ottenibili attraverso il comando "word1"[]{0,AFT}"word2", ad es.: "bella"[]{0,4}"notizia","arrivano"[]{1,1}"notizie","pubblicare"[]{2,2}"notizia|notizie", dove:

- le co-occorrenze restituite sono solo quelle che compaiono nella sequenza indicata, quindi "bella" "notizia" restituisce la concordanza <bella notizia> ma non <notizia bella>;
- la ricerca è *case sensitive*, ossia lo stesso termine è calcolato come differenti token solo per il fatto di essere scritto con lettere maiuscole o minuscole (*notizia* vs *Notizia*, *arrivano* vs *Arrivano*);
- si possono indicare più alternative usando la barra verticale |, ad es. "notizia|notizie";
- si può selezionare lo *span* con {0,AFT}, quindi in {0,4} la distanza massima che intercorre tra node e collocate è di 4 token (<notizia preoccupante>, <notizia molto preoccupante>, <notizia ancora più preoccupante>, <notizia comincia a diventare preoccupante>); invece {1,1}, fornisce solo i risultati del tipo <notizia molto preoccupante>, ma non <notizia preoccupante>. Lo span massimo impostabile è  $\pm 20$ , ma qui si utilizza l'intervallo massimo  $\pm 4$ , come previsto nella definizione funzionale di collocazione (2.1).

Riguardo al censimento del pattern Verbo+Complemento e Soggetto+Verbo sono necessarie alcune precisazioni, dato che il comando per la ricerca attraverso la forma del lemma verbale, [lemma="..."], non sempre prevede la possibilità di aggiungere il nome in funzione di Soggetto o Complemento; in questo caso si è deciso di procedere alla ricerca delle coniugazioni dei predicati limitatamente all'Indicativo (Presente, Imperfetto, Passato Remoto e Futuro Semplice), al Congiuntivo Presente, al Gerundio Presente e al Participio Passato; inoltre, per il pattern Verbo+Complemento, lo spoglio è stato effettuato solo per le terze persone, singolare e plurale.

Nella definizione di collocazione, sono contemplati il criterio descrittivo della frequenza e il criterio inferenziale della forza di associazione, oggettivamente calcolabili per mezzo delle misure di associazione; questo lavoro sfrutta principalmente la MI, un indice adatto a stimare la coesione tra due lessemi, intesa come la forza con cui due

vocaboli sono attratti l'uno dall'altra, però non è in grado di informare circa la direzionalità dell'attrazione, che non è mai simmetrica. Gli ideatori della MI Church e Hanks, nell'articolo *Word Association Norms, mutual information and lexicography* (1990), osservano che solo le coppie di parole con span massimo di  $\pm 5$  e un valore soglia  $> 3$  tendono a essere interessanti. I valori di MI sono inseriti in fasce di grandezza che accolgono i *bi-grammi* (sequenze di due parole) a seconda del loro punteggio:

MI	Valori soglia
$>9 - \leq 11$	AA = molto alta
$>7 - \leq 9$	A = alta
$>5 - \leq 7$	M = media
$>3 - \leq 5$	B = bassa
$\leq 3$	NC = non collocazione

In alcuni pattern si fa riferimento anche al LL, il quale in modo molto generale può essere inteso come la misura in grado di quantificare le evidenze di un'associazione positiva tra i termini rispetto alla probabilità che tale associazione sia del tutto casuale.

Dato che la presente ricerca intende operare il raffronto tra uso e descrizione lessicografica della lingua, la disanima dei pattern collocazionali di *notizia* si articola in due fasi: il punto di partenza sono le combinazioni statisticamente significative censite dallo spoglio del corpus dal punto di vista delle qualità extralinguistiche (2.1) della frequenza assoluta e della MI; sulla base di questi risultati sperimentali viene stabilito il confronto con i pattern collocazionali ottenuti dalla ricognizione condotta nei repertori.

## Capitolo 3

### Descrizione dei pattern collocazionali censiti nel corpus e raffronto con la ricognizione condotta nei repertori

Dopo aver introdotto nel Cap. 2 la definizione operativa di collocazione, gli strumenti utilizzati e il metodo applicato per l'analisi, il Cap. 3 si concentra sulla descrizione dei pattern collocazionali del lessema *notizia*, node della collocazione, in grado di creare relazioni forti e stabili con Aggettivi (3.1 e 3.2) e Verbi (3.3 e 3.4), i collocati. Le occorrenze del lemma *notizia* sono 20.890, su un tot di 164.599.694 di token del corpus: *notizia* 11.058 occorrenze, *notizie* 9.356, *Notizia* 135 e *Notizie* 341.

L'illustrazione dei pattern è incentrata sul confronto tra le concordanze ottenute dallo spoglio del corpus e il trattamento lessicografico dei costituenti in combinazione; un'attenzione speciale è stata dedicata ai collocati individuati per mezzo delle misure di associazione, principalmente Mutual Information (MI), i cui valori sono consultabili nelle tabelle in Appendice. Ciascun tipo di pattern collocazionale è descritto attraverso alcune considerazioni generali sui termini in collocazione, la discussione di alcuni casi interessanti e l'ampliamento dei possibili collocati non citati nei repertori.

La tipologia dei pattern è mutuata dal Dizionario delle Collocazioni (DdC, Tiberii, 2018) che li suddivide in: Nome+Aggettivo (in posizione libera o obbligata), Verbo+Complemento e Soggetto+Verbo; in questo lavoro si trascura la categoria Costruzioni che include Locuzioni (o Polirematiche), ad es. *fuga di notizie*, *notizia di attualità/cronaca/ecc*, le quali pur essendo delle combinazioni di parole, contengono un sovrappiù di compattezza semantica e sintattica tali per cui sono assimilabili alle monorematiche. Il DdC è anche il repertorio di riferimento per la descrizione e la discussione dei termini in combinazione, per i motivi esposti in 2.2.

Le tabelle in Appendice contengono: i pattern lessicali collocazionali ottenuti dalla ricognizione nei repertori e distinti tipologia, riportati prima in ordine alfabetico e poi in ordine di frequenza assoluta di co-occorrenza; i dati ricavati dallo spoglio del corpus a scopo di verifica e ampliamento; le classifiche dei valori nelle misure di associazione utilizzate nel lavoro; in riferimento a queste ultime, la graduatoria dei valori di MI è stata "ripulita" dalle interferenze (rumore o *noise*) degli elementi linguisticamente non

rilevanti, ad es. nomi propri (ANSA, Times e Striscia con riferimento al programma televisivo) e quella di LL dai segni di punteggiatura, apostrofo, congiunzioni, preposizioni, articoli, pronomi.

Per agevolare la lettura, qualora un Verbo sia citato in forma lemmatizzata si è optato per la grafia in carattere tondo maiuscolo (ad es. ESSERE, FARE).

### **3.1 Pattern Aggettivo in posizione libera**

Il pattern collocazionale del lemma *notizia* (*n.*) combinato con Aggettivo (Agg), si configura con il Nome (N) in qualità di testa sintattica del sintagma nominale e l'Agg che indistintamente può seguire o precedere il Nome.

L'Agg qualificativo è quella parte variabile del discorso che serve a modificare il significato del N a cui si riferisce per esprimerne una particolare qualità (Agg) o una determinazione per individuarlo o specificarlo meglio (Agg determinativo) e con il quale si accorda nel genere e nel numero, svolgendo una funzione attributiva (*la notizia attendibile*) o predicativa (*la notizia è attendibile*); l'Agg qualificativo costituisce una classe aperta di elementi con la possibilità di ampliamento per mezzo di nuove coniazioni. Generalmente, quando più Agg sono posposti al N il loro ordine di successione deriva dal grado di dipendenza logica degli Agg rispetto al N, ad es. *n. inattesa e preoccupante, n. false esagerate e tendenziose*. Un tipo particolare di Agg è l'Agg. di relazione, il quale se usato in combinazione con altri Agg esibisce delle preferenze riguardo la sequenza degli elementi (3.2).

#### **3.1.1 Descrizione del pattern Aggettivo in posizione libera**

Gli Aggettivi in posizione libera inventariati nei repertori in totale sono 118 e la ricerca condotta sul corpus ne restituisce 7 con un'unica occorrenza (hapax) e 17 con nessuna; i più citati sono *recente* (in 8 dizionari, diz) e *infondata* (in 9 diz), mentre sotto le 4 occorrenze si nota una parcellizzazione delle citazioni fra i repertori.

Riguardo la tipologia, si hanno: Agg qualificativi (ad es. *falsa, ghiotta, gradita*), di relazione (ad es. *clamorosa, catastrofica, generica*), Participi Presenti (*interessante, dirompente, scottante*) e Passati che ricevono trattamento aggettivale (*dettagliata, fondata, gradita*); tra questi ultimi nei repertori si trovano sia *attesa* (Agg in posizione

libera) e *inventata* (pattern Sequenza Obbligata, 3.2), che ATTENDERE e INVENTARE, attribuiti al pattern Verbo + Complemento (V+Compl) (3.3).

Il trattamento riservato ai Participi Passati nei dizionari non è sempre uniforme, ad es. DdC attribuisce al pattern V+Compl sia *attesa*, Agg derivato dal Participio Passato del V transitivo, che ATTENDERE, mentre negli altri dizionari lo si trova o con funzione aggettivale (ThT), oppure verbale e incluso fra le forme coniugate del Verbo (PF, SCAN); anche per *documentata*, menzionato solo in DdC, nelle concordanze del corpus emerge l'uso esclusivo in diatesi passiva e solo alla flessione plurale.

DdC: fornisce 73 Agg su 118 totali e include tutti gli Agg che mostrano valori significativi di MI (*brutta, tristi, sensazionale, allarmante, attendibile, inattesa, scarse, splendida, contraddittorie*) tranne *contrastanti* (MI 6,892 media coesione), citato unicamente in DCL e *sicure* (MI 6,690 media coesione) citato da DCL, SCAN e DAU; viceversa, DdC è l'unico a introdurre *curiose* (MI 8,122 alta coesione). DdC e pochi altri si sono dimostrati perspicaci per aver presentato tra gli Agg i collocati di *notizia/e*:

- *splendida*, in DdC e ThTr, a media coesione MI con “notizia” 6,455;
- *contraddittorie* (MI 7,640 alta coesione) DdC, SCAN e ThTr; *inattesa* (MI 7,556 alta coesione) e *tristi* (MI “notizie” 7,281 alta coesione) DdC, DCL e ThTr.

DdC, inoltre, inserisce *brutta/e* e *scarse* tra gli Agg in posizione libera, mentre l'uso più frequente osservato è costituito dalle sequenze *brutta n.* e *scarse n.*, Agg+N, ordine suggerito anche nei sei repertori che presentano la collocazione fra gli esempi.

Sulla base della frequenza assoluta, gli Agg forniti da Ddc sono dispiegati lungo tutta la classifica, dai ranghi più alti fino a quelli che contano 4 occorrenze ed è il solo dizionario a riferire Agg molto frequenti (*positiva, dettagliata, drammatica*), *avvilente* in hapax e *cruda, dirompente, rincuorante, stringata* con zero occorrenze:.

Il censimento del corpus ha consentito di mettere in luce il comportamento di alcuni Agg in riferimento a parametri importanti, quali:

- la forma flessiva: Agg solo al singolare *splendida, feroce*; Agg solo al plurale *frammentarie, tendenziose, contrastanti, scarse, contraddittorie documentate*;
- l'ordine dei costituenti: Agg+N *brutta, lieta, feroce notizia* e N+Agg *notizia/e allarmistiche, contraddittorie, contrastanti, deprimenti, documentate, dubbie, eclatanti, fuorvianti, inaspettata, incoraggianti, inedita, negativa, positiva, precisa, rilevante*;

- l'inserzione sintagmatica: *falsa, inattesa, infondata, interessante, rassicurante* sono Agg in posizione libera che esibiscono uno span fino a  $\pm 4$ , indicativo della possibilità di inserimento di elementi linguistici tra il node e il collocato, ad es. span +: *n. approssimative e spesso false, n. a dir poco inattese, n. quanto mai utili e interessanti, n. e informazioni utili o confortanti* span -: *infondata e oltremodo calunniosa la n.*;
- il preponderante uso in combinazione con l'Avverbio (Avv) *più: notizie più dettagliate, n. più importante, n. più precisa.*

### 3.1.2 Descrizione di alcuni casi del pattern Aggettivo in posizione libera

#### *brutta/e*

Aggettivo	Occorrenze CORIS		Span							
			-4	-3	-2	-1	+1	+2	+3	+4
brutta	min. s.	110; 4				110		1	2	1
	pl.	86; 4				86		2	1	1
	Ma. s.	-								
	pl	31; 1				31			1	
	TOT	<b>236</b>								

*brutta/e* è l'Agg più frequente tra quelli inclusi nella categoria Agg in posizione libera, presenta 236 occorrenze su 4.034 totali, inoltre *Brutte* raggiunge un valore di MI molto alta, pari a 9,317, si attesta come terzo Agg qualificativo che ha un legame forte e coeso con *notizie* e anche il valore ottenuto di Log Likelihood (LL) avvera l'ipotesi che esista una dipendenza tra le occorrenze del bi-gramma; la collocazione ammette la variabilità flessiva e con buon equilibrio tra forme singolari e plurali. Il significato della collocazione è semi-composizionale, perché l'Agg esprime un uso figurato di tipo metaforico.

L'aspetto interessante riguarda la modificabilità sintagmatica: lo span -1, *brutta|e notizia/e*, coprendo quasi tutte le concordanze, rivela un pattern Sequenza Obbligata (3.2), ulteriormente supportata dal discreto numero di occorrenze di *Brutte* con l'iniziale maiuscola, segnale che l'Agg in combinazione compare a inizio della frase. Tra le uniche 9 occorrenze con span positivo, nessuna è a +1 (*n. brutta*) e le rimanenti inseriscono: a span +2 il Verbo ESSERE, che determina il passaggio dell'Agg dalla funzione attributiva a quella predicativa, *la brutta n. vs la n. è brutta*, oppure un Avv di intensificazione, *la n. molto/davvero brutta*; a span +3: in coppia con altri Agg, *la n. bella e brutta, la n. buona*

o brutta; a span +5 una proposizione relativa restrittiva, *le notizie che arrivano sono solo brutte*.

La collocazione è citata in sei repertori: DdC la inserisce nella categoria N+Agg in posizione libera, ma dalle evidenze statistiche si evince che pertiene alle Sequenze Obbligate; in MdD la troviamo nelle collocazioni di tipo B, ad alta frequenza e media coesione; tuttavia, se per coesione si intende solo la modificabilità sintattica e non la sostituibilità paradigmatica (ma su questo aspetto l'autore non è chiaro) la collocazione è ad altissima coesione; nel SC è tra gli esempi in forma di enunciato, *essere rattristato da una brutta notizia*, con il Verbo (V) presentato nella forma del lemma e non isolata nella sezione fraseologica; N.Tr., *comunicare, ricevere una bella, una brutta notizia* è inserito fra gli esempi del significato 1.a.; il Thesaurus Treccani (ThTr) nella mappa linguistica cita l'Agg nel riquadro riservato alle qualità del lessema, non specificandone la posizione rispetto al Nome; DAU segnala *brutto*, nel campo "parole che si combinano con -"; infine la collocazione emerge come concordanza nel corpus di ELDIT, ma non compare fra gli Agg del lemma.

***triste/i***

Aggettivo	Occorrenze CORIS		Span							
			- 4	-3	-2	-1	+1	+2	+3	+4
triste	min. s.	14; 5				14	1	3		1
	pl.	3; 6				3	3	2	1	
	TOT	28								

*triste/i* ricorre in 28 occorrenze su 5.789 occor. e con MI "notizie" 7,281 raggiunge un livello medio di coesione, anche se nel corpus 2 occorrenze su 3 mostrano l'uso al singolare. Dal punto di vista sintagmatico presenta forte coesione, soprattutto a span -1, *triste notizia*, e pochi casi a span +1, *notizia triste*; le occorrenze con span maggiore di +1, sono costituite da Avverbi utili a intensificarne il significato: span +2: *n. ben, molto, così triste*, span +3: *n. le più tristi*, ed eventualmente in combinazione con altri Agg, span +4: *una n. così angosciata e triste*.

Il significato è semi-composizionale e il dato interessante è che nel contesto linguistico, la collocazione in 25 occorrenze su 28 assume il significato espresso dall'iperonimo "morte"; questo dato consente di considerare la variabilità paradigmatica dell'Agg in collocazione, che ammette una limitata sostituibilità con un cluster di Agg riuniti in una solidarietà semantica, i quali se combinati con *notizia* prendono lo stesso

senso, anche se nessuno fra questi è un collocato vero e proprio, dato che nessuno di essi compare nelle graduatorie di MI. Gli Agg in questione sono tutti caratterizzati da un'intensa coesione sintagmatica: tranne pochi casi a span +2, le co-occorrenze si trovano a span  $\pm 1$ ; inoltre, il materiale linguistico inserito è conforme a quello *triste/i*; gli Agg che appartengono al cluster e vengono elencati di seguito in ordine di frequenza assoluta delle occorrenze:

- *terribile* al decimo posto nella frequenza assoluta degli Agg, span  $\pm 1$  bilanciato (cioè il numero delle occorrenze a span +1 e il numero di quelle a span -1 all'incirca si equivalgono) e uso prevalentemente singolare, significa “morte” in 28 occorrenze su 34 totali;
- *tragica* predilige la sequenza *tragica n.* e l'uso singolare, significa “morte” in 21 occorrenze su 27 totali;
- *drammatica* span  $\pm 1$  bilanciato, alterna in egual misura la forma singolare e plurale, significa “morte” in 10 occorrenze su 14 totali;
- *dolorosa* privilegia la *dolorosa n.* (span -1), alterna in egual misura la forma singolare e plurale, significa “morte” in 5 occorrenze su 7 totali;
- *tremenda* con span  $\pm 1$  bilanciato, significa “morte” solo nella forma singolare, 5 occorrenze al singolare, mentre l'unica plurale assume un significato diverso;
- *funesta* significa “morte” in 3 occorrenze su 4 totali in cui prevale la forma plurale e l'ordine dei costituenti cambia in base alla forma flessa: al singolare *funesta n.* (span -1), mentre al plurale *n. funeste* (span +1).

L'Agg *triste* si trova in tre repertori: ThTr nella mappa al riquadro dedicato alle qualità del lemma e DdC tra gli Agg, anche se tendenzialmente la collocazione si presenta con sequenza *triste n.* (span -1), come in DCL *una triste notizia*; DCL quindi, è l'unico a riportare la collocazione autentica, in apertura di una stringa di Aggettivi semanticamente affini: “una triste, dolorosa, infausta, funesta, nefasta tragica, luttuosa, ferale n.”.

***curiosa/e***

Aggettivo	Occorrenze CORIS		Span							
			- 4	-3	-2	-1	+1	+2	+3	+4
curiosa	min. s.	3; 12		1		2	8	1	3	
	pl.	2; 8				2	4	2	2	
	Ma. s.	-								
	pl	1					1			
	TOT	26								



*curiosa/e* ricorre in 26 occorrenze su un totale di 1.720 e si posiziona al quattordicesimo posto per la frequenza assoluta fra gli Agg; inoltre, *curiose* avendo un alto punteggio di MI (8,122), si classifica come quinto Agg tipicamente in collocazione con *notizie*, nonostante che nelle occorrenze emerse la flessione singolare superi di poco quella plurale (15 sing. vs 11 plur.).

L'Agg presenta un uso figurato, per cui il significato della collocazione è limitatamente compositivo.

La coesione fra i costituenti è molto forte (span  $\pm 1$ ) e in metà delle occorrenze mostra l'ordine N+Agg (span +1); la collocazione ammette l'inserzione di Avv modificatori dell'Agg, *n. più/piuttosto curiosa* (span +2), *n. ancora più curiosa* (span +3), ma anche di altri Agg *n. fresca e curiosa*, o ESSERE, *la n. è già curiosa di per sé*, in cui l'Agg da qualificativo diventa predicativo; solo 1 occorrenza è a span -3, *curiosa è la n.*, dove l'ordine sintatticamente marcato dei costituenti nella costruzione predicativa (Agg+ ESSERE+N) a livello pragmatico ha lo scopo di porre il focus sulla qualità del sostantivo.

L'unico dizionario in cui compare la collocazione *notizia curiosa* è DdC, che in questo caso si rivela più intuitivo rispetto agli altri repertori.

### **scarse**

Aggettivo	Occorrenze CORIS		Span							
			-4	-3	-2	-1	+1	+2	+3	+4
n. scarse	min. pl.	9; 3	2			7	1		1	1
	Ma. pl.	1				1				
	TOT	<b>13</b>								

L'Agg *scarse* conta 13 occorrenze su 783 totali; *scarse* raggiunge un valore medio di MI (6,989) e la collocazione ricorre esclusivamente alla forma plurale. Prevalde decisamente lo span -1, cioè l'ordine marcato Agg+N, che può espandersi fino a span -4, ad es. *scarse e poco sicure n.*; le occorrenze con span positivo maggiore a 1 sono: *le n. documentarie sono s. e le n. da scarse diventano commodity*.

Il trattamento lessicografico della collocazione registra l'ordine marcato in ben sei repertori; DdC in controtendenza è l'unico a indicare la posizione libera dell'Agg, ma le evidenze del corpus supportano l'idea che *scarse n.*, analogamente a *brutta/e n.*, rientri nel pattern delle Sequenze Obbligate, consigliate o più frequenti nell'uso.

### 3.1.3 Ampliamento del pattern Aggettivo in posizione libera

#### *priva/e di*

Aggettivo	Occorrenze CORIS		Span							
			- 4	-3	-2	-1	+1	+2	+3	+4
<i>prive</i>	min. s.	2; 9	2				2	3	1	3
<i>di fondamento/interesse/importanza</i>	pl.	3; 11	3(1:+5)				5	1	2	3(1:+5)
	TOT	<b>25</b>								

L'impiego delle misure statistiche supporta l'ampliamento degli Agg in collocazione con il termine in esame e non citati nei repertori. *Priva/e* compare in 25 occorrenze su un totale di 4.561 e *prive* raggiungendo un valore di MI 7,368 con "notizie" si inserisce tra i collocati ad alta coesione, ma dall'osservazione della frequenza delle co-occorrenze si constata che è ben presente anche la flessione al singolare. Il significato della collocazione è semi-composizionale, dato che l'Agg viene utilizzato in senso figurato.

L'ordine dei costituenti è tipicamente N+Agg (span+) e i contesti in cui compare nell'ordine marcato Agg+N, risponde all'esigenza pragmatica di mettere in rilievo la qualità della *notizia*, operazione possibile per mezzo del focus sintattico, ad es. *è priva di fondamento la n.* (span -4), *definisce prive di ogni fondamento le n.* (span -5). Nell'ordine non marcato le occorrenze con span maggiore o pari +2 consentono l'aggiunta di Agg, Avv e proposizioni, oppure l'espressione della funzione predicativa dell'Agg. Ad es.:

- span +2: *n. così priva di..., n. strillate prive di...;*
- span +3: *la n. è assolutamente priva di..., le n. pubblicate sono prive di...;*
- span +4: *n. discreta e apparentemente priva di..., n. sensazionali ma spesso prive di..., n. che, considerate generalmente prive di..., n. del genere sebbene prive di...;*
- span +5: *n. di genere diverso comunque prive di....*

A livello sintattico l'Agg è sempre accompagnato da un Complemento di Privazione che ammette la variabilità paradigmatica dei sostantivi *interesse, importanza, fondatezza e fondamento*; *fondamento* co-occorre con *notizia* 16 volte su 25 occorrenze e considerato che ottiene un valore di MI 6,121 in combinazione con *notizia*, si può concludere che *notizia/e priva/e di fondamento* costituiscono una collocazione a medio-alta coesione.

### 3.2 Pattern Sequenze Obbligate

Il pattern delle Sequenze Obbligate coinvolge la categoria lessicale dell'Agg, il quale, diversamente dalla posizione libera, si trova in una sequenza obbligata consigliata o più frequente nell'uso e che tipicamente o segue o precede il N.

In italiano l'Agg qualificativo può sia seguire che precedere il sostantivo, ma la permuta di posizione produce una serie di cambiamenti che interessano diversi livelli della lingua, ad es.:

- 1) Gianni ha divulgato la *n. telegrafica*, più dettagliata di quella telefonica
- 2) La *telegrafica n.* ha sconvolto tutti

A livello semantico si verifica un radicale mutamento di senso, nell'es. 1) *telegrafico* denota più oggettivamente la caratteristica del N, e l'Agg ha un significato letterale, mentre in 2) evidenzia una maggiore soggettività di giudizio e riceve interpretazione metaforica. A livello sintattico l'Agg dell'es. 1) svolge funzione restrittiva e designa il referente del Nome; invece, in 2) ha funzione descrittiva e riferisce a una qualità accessoria del sostantivo. L'ordine Agg+N, in pragmatica, generalmente indica un maggior coinvolgimento emotivo ed infine nella dimensione diafasica in riferimento ai registri, la sequenza Agg+N quando non obbligatoria, esprime l'elaborazione stilistica del messaggio (Dardano-Trifone:1997, 451).

Un tipo specifico di Agg qualificativi sono gli Agg di relazione, i quali sono denominali e mettono in relazione il Nome a cui si riferiscono con quello da cui derivano; il legame tra Agg relazionali e N può essere semplice da dedurre (*n. economica* → *n. di economia*), ma in altri casi necessita di una parafrasi (*n. angosciosa* → *n. piena di angoscia*); inoltre, molti di essi non possiedono il comparativo e il superlativo (*\*n. più economica*, *\*n. economicissima*), non possono né essere anteposti al N (*\*economica n.*), né essere utilizzati in funzione predicativa (*\*la n. è economica*), altrimenti acquistano lo status di Agg qualificativi, che esprimendo una qualità del N mutano radicalmente il significato del sintagma (Dardano-Trifone: 1997, 197-8).

La compresenza di un Agg qualificativo e un Agg di relazione prevede le sequenze N+Agg relazione+Agg qualificativo, *n. giornalistiche riservate*, e Agg qualificativo+N+Agg relazione, *importanti notizie economiche*.

### 3.2.1 Descrizione del pattern Sequenze Obbligate

Le Sequenze Obbligate (Seq Obbl) recensite nei repertori in totale sono 40 e l'esplorazione compiuta sul corpus ne restituisce 5 in hapax e 5 con nessuna occorrenza. Tra le Seq Obbl più frequenti le più citate sono *ultima n.* (13 diz), *buona n.* (9 diz), *bella n.* (8 diz), *n. certa* (8 diz), di cui solo *certa* non ottiene un valore rilevante nelle misure di associazione; contemporaneamente sono molti gli Agg in Seq Obbl menzionati in un solo dizionario.

Le concordanze mostrano una correlazione tra il tipo di Agg e l'ordine costituenti in combinazione: qualificativi+N e viceversa N+di relazione.

Sempre sulla base delle co-occorrenze del corpus, è possibile osservare il comportamento dei Participi Passati che nel trattamento lessicografico ricevono un'interpretazione aggettivale: *anticipata* in DCL e *accreditata* menzionato unicamente in MdD, ma le rispettive concordanze restituiscono solo forme nel contesto di diatesi passiva; *esagerate*, citato in NTr, è effettivamente utilizzato in funzione aggettivale e con cospicua presenza nella sequenza *n. false esagerate e tendenziose*, originaria del linguaggio giuridico poi estesa a quello della stampa; *riservata/e* mostra un uso aggettivale e i quattro dizionari che lo presentano lo includono tutti tra gli Agg; *inventata*, è utilizzato sia come Agg sia come V, ma se DAU ne illustra entrambe le funzioni, in DdC e ThTr si trova solo tra gli Agg.

DdC riporta 17 Sequenze Obbligate e cita tutti gli Agg che mostrano un valore significativo nelle misure di associazione (*cattive* MI 11,562, *riservate* “notizie” MI 6,901, *buona* LL e *bella* LL), però non registra *prime*; tra le collocazioni di “notizia” vi è *notizia choc*, presente solo in DAU al campo dedicato agli Aggettivi con la variante grafica *shock*.

Tra le prime quattordici Seq Obbl più frequenti in assoluto, in DdC sono assenti solamente *prime notizie* e *ottima notizia* e al contrario recensisce *notizia esclusiva* e *notizia scandalistica*, che però non trovano alcuna concordanza.

In DdC è specificato che *notizia bomba* è un'espressione invariabile e sebbene le evidenze del corpus mostrino un uso prevalente alla forma singolare (29 occorrenze), non si può escludere totalmente quella plurale, rinvenuta in 3 occorrenze.

Il censimento del corpus ha permesso di illustrare l'uso di alcune Seq Obbl, in riferimento ai seguenti parametri:

- flessione: *prime n.* e *notizie sportive* compaiono solo al plurale, e vengono citate in questa forma nei dizionari di lingua d'uso ed ELDIT, e al singolare in DdC e ThTr, assieme a DAU che però impiega la forma lemmatizzata dell'Agg (*sportivo*).
- ordine dei costituenti in combinazione: tutti gli Agg si trovano o con sequenza Agg+N o N+Agg, tranne *buona* che oltre al vasto uso davanti al N, presenta un modesto numero di concordanze N+Agg;
- inserzione: la maggioranza degli Agg in Seq Obbl non ammettono l'aggiunta di ulteriori elementi linguistici tra node e collocato a esclusione di *buona*, *cattiva*, *riservata*, ad es. *la n. del ritrovamento doppiamente buona*, *n. comparabilmente cattiva*, *n. molto importante e riservata*. È interessante il caso di *esagerata/e* (ThTr) che su 8 occorrenze solo 1 ha span +1, mentre tutte le altre presentano uno span  $\geq 2$ : *la n. della mia morte è esagerata*, *la n. è grossolanamente esagerata*.

### 3.2.2 Descrizione di alcuni casi del pattern Sequenze Obbligate

#### *ultima/e notizia/e*

Sequenza Obbligata	Occorrenze CORIS		Span							
			- 4	-3	-2	-1	+1	+2	+3	+4
ultima n.	min. s.	50	1		2	47				
	pl.	188			9	179				
	Ma. s.	1				1				
	pl.	15				15				
	TOT	254								

*ultima/e* conta 254 occorrenze su 28.050 totali ed è l'Agg più frequente in combinazione con *notizia/e* dopo *buona/e*; *ultime*, raggiungendo un valore significativo di LL, misura che esprime la non casuale co-occorrenza degli elementi considerati, è un collocato di “notizie” e sebbene consenta la variabilità flessiva, le occorrenze al plurale sono in rapporto 3 a 1 con quelle al singolare. Il significato è compositivo e le evidenze del corpus mostrano lo span -1, che da un lato conferma la posizione prenominali dell'Agg e dall'altro prova la forte attrazione reciproca dei costituenti, soprattutto al plurale, e la limitata possibilità di inserzione, generalmente rappresentata da altri Agg che a loro volta adottano l'ordine marcato: *ultima buona n.*, *ultime drammatiche n.*, *l'ultima è una brutta n.*

La Seq Obbl è tra le più menzionate nei dizionari (tutti tranne PF e SCAN), con l'ordine sintagmatico Agg+N e la forma plurale *ultime notizie*, solo DdC ed ELDIT in

controtendenza riportano la versione singolare. Fra i dizionari combinatori, in DCI è l'unico Agg citato in questa categoria e viene riproposto all'esempio *leggere le ultime n.*, alla categoria Verbi; in MdD *sentire le ultime n.* è inserito fra le collocazioni ad altissima frequenza e minima coesione, impiegata nella lingua standard (collocazioni di tipo A); DAU, dizionario analogico, *ultime notizie* è riportato nel segmento "Combinazioni tipiche di parole con - (notizia/e)".

Analogamente a *ultime notizie*, anche *prime notizie* prevede esclusivamente la posizione prenominali dell'Agg e le 92 occorrenze (solo flessione plurale) possiedono quasi tutte span -1, indice di forte coesione tra i costituenti; diversamente da *ultime n.*, la combinazione compare unicamente in Zin nell'esempio al senso di "nozione-notizia": *le prime n. di una scienza.*

***notizia/e riservata/e***

Sequenza Obbligatoria	Occorrenze CORIS		Span							
			- 4	-3	-2	-1	+1	+2	+3	+4
n. riservata	min. s.	8					4	1	2	1
	pl.	54					47	5	1	1
	Ma . s.	-								
	pl.	1					1			
	TOT	63								

*riservata/e* si trova in 63 occorrenze su 4.195 occor e il significato è semi-composizionale.

L'ordine sintagmatico è rigido: in tutti i contesti l'Agg segue il Nome; la maggioranza dei casi mostra span +1, per cui essendo la coesione fra i costituenti molto forte, la Seq Obbl solo raramente compare con l'aggiunta di ulteriori elementi, i quali possono essere altri Agg, *n. giudiziarie riservate*, *n. molto importanti e riservate*, V che esprimono l'aspetto durativo, *la n. rimane riservata*, e Avv *n. per ora riservata*.

L'Agg *riservate* raggiungendo un valore di MI "notizie" (6,901) rappresenta un collocato di media coesione e questo dato coadiuvato dalla frequenza d'uso delle occorrenze plurali, porta alla conclusione che nei repertori la miglior forma di citazione sarebbe notizie *riservate*.

I repertori che recensiscono la collocazione, sono solo tre: Tiberii fra le Seq Obbl e con uso indifferenziato della forma flessiva, Th.T al singolare e senza indicazioni

riguardo la posizione dell'Agg e DAU che cita l'Agg in forma lemmatica al campo "Aggettivi che si combinano con –"

### 3.2.3 Ampliamento del pattern Sequenze Obbligate

#### *notizia/e scomoda/e*

Sequenza Obbligata	Occorrenze CORIS		Span							
			-4	-3	-2	-1	+1	+2	+3	+4
scomode	min. s.	2					1	1		
	pl.	10					9			1
	Ma. s.	-								
	pl	1					1			
	TOT	13								

*scomoda/e* presenta 13 occorrenze su un totale di 753 e pur essendo un collocato a alta coesione con valore di MI "notizia|notizie" di 8,441, non è censito nei repertori.

L'Agg è usato in senso figurato e quasi esclusivamente alla forma plurale; l'ordine dei costituenti è unicamente N+Agg con possibilità di ampliamento sintagmatico, ad es. V con aspetto resultativo, *la n. risulta scomoda*, o Avv dal valore concessivo, *la n. utile anche se scomoda*.

### 3.3 Pattern Verbo + Complemento

Nel pattern Verbo+Complemento (V+Compl), *notizia/e* ricopre il ruolo di Complemento, inteso come "qualcosa che si aggiunge a un Nome per determinarne il significato, in qualunque modo ciò possa avvenire" secondo la definizione data da Beauzée nel XVIII secolo (Dardano-Trifone: 1997, p. 122).

*Notizia/e* nelle concordanze del corpus e negli esempi citati nei repertori, costituisce un elemento nucleare del sintagma verbale e la sua presenza è necessaria per dare senso compiuto all'espressione; il criterio per capire se un costituente funge da Complemento del Verbo è la valenza del Verbo stesso e il contesto linguistico in cui quest'ultimo è inserito.

In questo studio *notizia/e*, a seconda del Verbo considerato, svolge il ruolo di Oggetto Diretto (*pubblicare la/le n.*) oppure di Argomento Strumentale (*avvisare con e informare con*).

Nel pattern V+Compl sono comprese anche le co-occorrenze in cui *notizia/e* a livello superficiale sintattico è il Soggetto grammaticale della diatesi passiva del V, ma a

livello di struttura semantica profonda esercita il ruolo semantico di Tema, ossia subisce l'azione e può essere da essa modificato.

All'interno del pattern V+Compl si distinguono le costruzioni a Verbo Supporto (*Light Verb Constructions*, LVC), in cui il Verbo andando incontro a uno svuotamento semantico assume sempre un significato generico (ad es. DARE, FARE, PRENDERE), mentre il significato lessicale poggia quasi completamente sul Nome; diversamente dalle altre collocazioni, dal punto di vista sintattico i componenti di queste costruzioni non sono totalmente autonomi e quindi le trasformazioni sintattiche risultano inibite.

### 3.3.1 Descrizione del pattern Verbo + Complemento

I repertori recensiscono 94 Verbi e l'interrogazione del corpus ne riporta 6 in hapax e 17 con nessuna occorrenza; in 92 V *notizia/e* è l'Oggetto Diretto e in 2 argomento del Verbo con funzione strumentale, *avvertire con* e *avvisare con*.

I Verbi più frequenti mostrano anche valori significativi nelle misure di associazione (MI e/o LL) e per la maggior parte sono ben presenti nei dizionari: (*non*) *avere, dare, diffondere, chiedere, ricevere, fare notizia, riportare, apprendere, fornire, raccogliere, commentare, comunicare, confermare, smentire, accogliere, trasmettere*; i V con minor frequenza, ma che costituiscono collocazioni con *notizia/e* sono: *annunciare, attingere* e *spargere*, i quali assieme a *chiedere* e *riportare*, ricevono scarso trattamento lessicografico. *Prendere notizia*, con 32 conc., non è compreso in nessun repertorio.

*Fa* (voce del V FARE) raggiunge un valore significativo di LL "notizia", ossia la misura che calcola la probabilità che due costituenti non compaiano casualmente nello stesso contesto; FARE è quindi un collocato di *notizia*, ciononostante non è possibile attribuirlo al pattern descritto in questo paragrafo o a quello Soggetto+Verbo (*la n. fa clamore/scalpore*, 3.4.1) sulla base delle sole misure statistiche. A dirimere la questione e assegnarlo al pattern V+Compl in qualità di LVC è utile il ricorso all'analisi qualitativa, che evidenzia l'indeclinabilità del node (*fare notizia* vs *\*fare notizie*), l'inseparabilità dei due costituenti (*\*fare la n.*, *\*fare molta n.*, *\*fare fresca n.*), l'inibizione delle trasformazioni sintattiche (*\*la n. è fatta da Gianni*, *\*la n. che ha fatto Gianni*) e la semantica non compositiva della collocazione *FARE notizia*, il cui significato è parafrasabile con "essere tale da suscitare l'interesse generale" (PF). *fare notizia* in



ambito lessicografico è una delle combinazioni più menzionate ed è recensita in dodici repertori:

- DdC la include tra i V del pattern senza fornire ulteriori informazioni;
- MdD la attribuisce alle collocazioni di tipo C, collocazioni di media o bassa frequenza d'uso generale e di alta coesione, assolutamente necessarie nell'uso scritto e parlato;
- DCL la propone nell'esempio "la politica ambientale non fa più n.";
- GraDIt la colloca tra le poliremariche, ThTr nel riquadro "parole, espressioni e modi di dire" e DAU nel campo "Locuzione e proverbi";
- invece DCI, DO, PF, SC, NTr, appurandone il carattere idiomatico, avvertono la necessità di segnalarla in quanto locuzione e istruire l'utente attraverso spiegazioni, commenti e/o esempi.

DdC elenca 50 V e più precisamente cita tutti i V in collocazione, quelli cioè che raggiungono valori significativi in MI e/o LL, tranne *sapere* (MI *saputa* 10,022) che compare in ThT e DAU, e *attingere* (MI "notizie" *attinte* 12,424) presente in DCL, PF, GDIH e SCAN; invece, tra i V più frequenti DdC trascura *portare* e *sentire*. DdC e pochi altri citano le combinazioni in collocazione:

DdC+2 diz: *fornire* in MdD e DAU, *annunciare* in MdD e ThT;

DdC+1 diz: *commentare* DAU, *accogliere* in MdD.

DdC è l'unico a inserire *tenere nascosta* e in questo si rivela intuitivo, perché *nascosta* è un collocato di *notizia* e la forma participiale conserva il valore verbale: infatti, nella maggioranza delle occorrenze si trova la combinazione *tenere nascosta la n.* o *la n. è tenuta nascosta*, dove la costruzione con il V *tenere* esprime l'aspetto durativo dell'azione.

Del trattamento lessicografico di *attesa* e *inventata*, Participi Passati con funzione aggettivale dei V *attendere*, si è discusso rispettivamente in 3.1.1 e 3.2.1.

Nelle concordanze del pattern V+Compl si osservano i fatti seguenti:

- flessione del node: *apprendere*, *confermare la n.* al singolare e *ascoltare aspettare*, *fornire*, *leggere le n.* e *chiedere*, *mandare n.* al plurale;
- variabilità paradigmatica: *ascoltare* (DCI, DO, PF, ThTr), *sentire* (DCI, MdD-collocazioni tipo A, GraDIt, SC, DAU) e *udire* (DAU); *avvisare con* e *avvertire con* entrambi recensiti solo in DAU.

### 3.3.2 Descrizione di alcuni casi di pattern Verbo + Complemento

#### *Verbi ad alternanza forma transitiva ~ pronominale*

Verbo + Complemento	Occorrenze CORIS	T. semplici	T. composti	N + Particípio Pass. concordato	Part. Passato concordato + N
diffondere	280	180	40	58	2
spargere	21	12	1	6	2

I due Verbi si rivelano interessanti, perché entrambi presentano l'alternanza della forma transitiva (*n.* Oggetto Diretto), con quella pronominale (*n.* Sogg).

Soggetto + Verbo	Occorrenze CORIS	T. semplici	T. composti
si diffonde	48	4	44
si sparge	59	27	32

Grazie ai dati forniti dalle misure di associazione si può osservare che le collocazioni con *notizia/e* prediligono l'uso di una forma rispetto l'altra.

*Diffondere* ha 280 occorrenze vs 48 di *si diffonde* e il V in collocazione è DIFFONDERE, dato che alle forme *diffondendo la/le*, *diffonde le n.*, *diffondere la n.* raggiunge valori di MI medio-alti, come pure *le n. diffuse*; il Particípio Passato in *n. diffuse* è ambiguo (potrebbe trattarsi sia della forma transitiva in diatesi passiva che del tempo composto del V pronominale), ma dallo spoglio delle concordanze nel contesto l'uso che emerge è il primo; anche la parola *diffondersi* (MI 6,865) rappresenta una collocazione di media coesione con *notizia/e*, ma nella graduatoria occupa un rango più basso rispetto ai valori di MI delle forme coniugate del V transitivo.

DIFFONDERE in collocazione con *notizia/e* consente l'aggiunta di altri elementi quali N, Agg, Avv o altri V: *diffondere dati e n.*, *diffondere le allarmanti n.*, *diffondere sempre più spesso n.*, *diffonde e commenta le n.*, *diffonde con aria allusiva n.*, *le n. errate e gravemente fuorvianti diffuse da ....* DIFFONDERE è ben rappresentato e si trova in tutti i dizionari combinatori, in PF, SC, SCAN, DAU e ELDIT invece *si diffonde* solo in DAU e ELDIT.

Riguardo alla seconda alternanza *spargere ~ si sparge*, il collocato significativo *sparsa* (MI 7,100) è riferibile a entrambe le forme, poiché nel corpus si trova sia *sparsa la n.* (2 concordanze) nella diatesi passiva del V transitivo, sia *si ESSERE sparsa la n.*,

(23 concordanze) e *la n. si ESSERE sparsa* (7 concordanze), Partecipio Passato dei tempi composti alla forma pronominale del V.

Il V alla forma transitiva compare in DCL, PF, SCAN e DAU, mentre in quella pronominale in DAU ed ELDIT, invece DdC non cita né *spargere* né *spargersi*.

Il corpus restituisce altre occorrenze che coinvolgono Verbi ad alternanza transitiva ~ pronominale, nonostante: *sgonfiare* ~ *si sgonfia* e *propagare* ~ *si propaga* (3.3.3); benché le due alternanze non raggiungano il valore minimo restituito dalle misure di associazione, l'aspetto interessante è che i repertori recensiscono solo la forma transitiva, ma l'uso emerso dalle concordanze è quello pronominale.

### 3.3.3 Ampliamento del pattern Verbo + Complemento

#### *scambiarsi e reperire*

Verbo + Complemento	Occorrenze CORIS	T. semplici	T. composti	N + Partecipio Pass. concordato	Part. Passato concordato + N
scambiarsi	19	18		1	
reperire	12	12			

*scambiarsi* e *reperire* sono le forme dei rispettivi lemmi verbali che presentano un valore MI: ad alta forza coesiva il primo (MI 7,330) e a media il secondo (MI 6,732).

Se SCAMBIARSI predilige la coniugazione al modo Indicativo e ai tempi semplici, REPERIRE nel corpus si trova solo all'Infinito Presente in combinazione con i lemmi POTERE (*si possono reperire recensioni e notizie*), CONSENTIRE (*consente di reperire facilmente notizie*), oppure come Infinito Sostantivato (*per/di reperire notizie*).

Nelle concordanze entrambi i V selezionano solo la flessione plurale (*notizie*) usata con ellissi dell'articolo determinativo.

Ulteriori ampliamenti e particolarmente degni di nota dal punto di vista lessicografico, riguardano *propagare* e *sgonfiare* ad alternanza transitiva e pronominale; *propagare* ~ *si propaga* non sono collocati veri e propri di *notizia*, in quanto non raggiungono il valore minimo restituito dalle misure di associazione; il V alla forma transitiva è recensito in DdC e DCL, ma il corpus restituisce solo concordanze alla forma pronominale, assente in tutti i dizionari; lo stesso ragionamento vale per *sgonfiare* ~ *si sgonfia*, recensito alla forma transitiva da DCL e DAU, ma le 2 occorrenze rinvenute sono alla forma pronominale e non presenti nelle combinazioni dei repertori.

### *coperte*

	Occorrenze CORIS		Span							
			-4	-3	-2	-1	+1	+2	+3	+4
n. coperte	min. s.	2								2 (+5)
<i>da segreto professionale/istruttorio</i>	pl.	13					12			1
	TOT	15								

*coperte* possiede una forza di coesione media con “notizie” (MI 6,602), e dalla ricerca condotta sul corpus, il dato interessante osservato è l’uso preponderante del V *coprire* alla diatesi passiva, mentre la coniugazione attiva appare solo in 2 concordanze: *le emittenti coprono le n. e nessun segreto possa coprire fatti, notizie o documenti concernenti reati.*

La collocazione ammette l’inserzione di Avv (*la n. era stata a lungo coperta dalle autorità sanitarie*), anche in funzione di modificatori dell’Agg (*la n. politicamente rilevante, non coperta da segreto di Stato*).

Su 15 concordanze alla costruzione passiva, in 13 il Compl. d’Agente è espresso dalla parola *segreto*, sempre seguito da un Agg qualificativo (*professionale, istruttorio, investigativo, d’indagine*) o da un Complemento di Specificazione (*di Stato*) e da queste evidenze si può affermare quindi, che l’uso più probabile della collocazione *coprire la/le n.* in entrambe le diatesi richieda il N *segreto* in qualità di Sogg logico.

### **3.4 Pattern Soggetto + Verbo**

Il pattern Soggetto+Verbo (Sogg+V) include solo i casi in cui il Soggetto grammaticale coincide con il Soggetto logico dell’enunciato e quindi *notizia/e* ricopre il ruolo semantico di Agente reale dell’azione, diversamente da quanto avviene nella costruzione passiva dove i due tipi di Soggetto non si corrispondono (3.3).

Il Soggetto è un costituente del sintagma verbale e completa il significato del Verbo; il Sogg non occupa sempre il primo posto nella frase (*fa rumore la n. di qsa*) e generalmente precede il V (*la n. sconvolge qno*), ma può anche essere postposto al fine di metterlo in particolare rilievo (*si sparge la n. di qsa*) o in contrasto con un altro Sogg (*pervengono indiscrezioni, non notizie*); in alcuni V intransitivi il Sogg tendenzialmente segue il Verbo (*trapela la n, giungono le n.*) (Dardano-Trifone:1997, 113).

### 3.4.1 Descrizione del pattern Soggetto + Verbo

I repertori recensiscono 41 Verbi totali, di cui 19 transitivi, 16 intransitivi, 5 con *si* pronominale e 1 (*interessare*) transitivo ~ intransitivo; inoltre, la forma participiale di alcuni di questi V riceve un'interpretazione aggettivale (ad es. ALLARMARE → *allarmante/i* Participio Presente in 3.1.1).

I V più frequenti giungono a valori significativi nelle misure di associazione (MI e/o LL): *fa*, *arriva*, *giunge*, *trapela*, *circola*, *si sparge*, *rimbalza*, *è apparsa su*, *filtra*, *si diffonde*, *interessa*, *sconvolge*; nell'uso INTERESSARE e SCONVOLGERE prevalgono sugli Agg da essi derivati (*interessante/i* e *sconvolgente/i*), i quali pur avendo maggior frequenza assoluta di occorrenze, non risultano significativamente coesi con *notizia/e*; fra i Verbi meno frequenti del pattern Sogg+V, ve ne sono 5 in hapax e 6 con nessuna occorrenza.

FARE è il Verbo più frequente in assoluto del pattern ma, nonostante l'altissima frequenza e il possesso di un valore significativo in LL "notizia", difficilmente rientra tra le collocazioni di questa struttura, a causa della semantica molto generica del V rintracciabile nelle combinazioni *la n. fa clamore/scalpore/etc...* (DCI, GDIH) e *la n. ha fatto il giro del mondo* (DCI). Per la descrizione di FARE in collocazione con *notizia*, si veda in 3.3.1.

In ambito lessicografico il modello Sogg+V è il più negletto fra i pattern e i repertori appaiono discordanti e disomogenei, valgano a titolo di esempio *arriva*, che pur essendo il più frequente in assoluto dopo *fa* (3.3.1) e possedendo valori significativi di MI e LL, è registrato solo da DdC; viceversa *corre*, nonostante la frequenza sporadica e i valori non rilevanti nelle misure di associazione, è il V più citato nei repertori: DdC e altri sette.

DdC fornisce il maggior apporto di V con 15 voci, mentre i più scarsi sono DCI, MdD e tutti i dizionari generali (tranne SCAN), che citano un unico lemma o anche nessuno (DO e Zin); DdC è anche il repertorio che nomina il maggior numero di V in collocazione con *notizia/e* in funzione di Soggetto, ma non segnala: *è apparsa su* (DCI), *si diffonde* (DAU e ELDIT) e *interessa* (ThT); al contrario, DdC è il solo a inserire *arriva* e *sconvolge* (vs sinonimo *ha scioccato*, in DCI) e assieme a pochi altri cita *filtra* (DCL) e *ribalza* (DCL, DAU).

Per alcuni V è possibile una certa variabilità paradigmatica, riscontrata sia nel corpus che nei repertori:

- il V copulativo *crea* (DdC) accompagnato da *scalpore, sconcerto, sgomento, allarme, clamore, panico*, è possibile sostituirlo con i V *fa* (DCI, GDIH), *desta* (corpus di ELDIT) e *suscita* (DAU);
- *arriva* (DdC) e *giunge* (DdC+6), entrambi raggiungono valori significativi nelle misure statistiche: *Arrivano* in MI e LL assieme ad *arriva* e *arrivata*, mentre *giungono* in MI e LL “notizie”.
- DCI tra gli esempi del pattern cita *la n. è apparsa/uscita sul giornale*: le occorrenze confermano l’uso più frequente di entrambi i Verbi ai tempi composti, però APPARIRE a differenza di USCIRE, conta più co-occorrenze e rappresenta un collocato ad alta coesione (*apparse* MI 7,827).

### 3.4.2 Descrizione di alcuni casi di pattern Soggetto + Verbo

#### *trapela e filtra*

Soggetto + Verbo	Occorrenze CORIS	T. semplici	T. composti
trapela	73	37	36
filtra	49	39	10

*trapelare* è il Verbo e anche il lemma che esprime un valore di massima coesione con *notizia/e*, in particolare alle forme *trapelò* (MI 11,714), *trapela* (MI 9,044) e *trapelata* (MI 8,834) e in quanto V intransitivo, il Sogg tendenzialmente è posposto; inoltre, nei contesti in cui il V appare coniugato ai tempi composti, vi è sempre specificazione di tempo: *la n. è trapelata da qualche minuto, la n. è trapelata soltanto ieri*.

TRAPELARE è recensito da DdC e altri otto dizionari, ma alcuni di essi lo propongono in costruzione causativa: *far trapelare una n.* (MdD) e *lasciar trapelare una n. segreta* (NTr e ThTr).

Su 73 occorrenze, 21 mostrano la costruzione con V causativo analitico (FARE/LASCIARE + Infinito), di cui 13 con FARE e 8 con LASCIARE, ad es.: *qualche talpa ha fatto trapelare la n., ho lasciato trapelare la n. in redazione*(io, Sogg sottinteso); questo tipo di costruzione si compone di due frasi e quindi di due Agenti, che stabiliscono fra loro un rapporto gerarchico, dove il primo (Iniziatore, *la talpa* e *io*) è istigatore dell’agire del secondo (Esecutore, *la n.*), e a livello pragmatico tale costruzione serve per mettere in rilievo che la responsabilità dell’evento è attribuibile all’Iniziatore, mentre quella dell’Esecutore risulta minima; a livello sintattico l’Esecutore assume sia la forma

che il comportamento dell'Oggetto Diretto e per questo alla forma passiva diventa il Sogg grammaticale della frase (*la n. era stata fatta trapelare da fonte ignota*). FARE e LASCIARE manifestano un diverso grado di forza causativa, forte il primo, perché impone o comanda, e debole il secondo, perché autorizza o permette (Simone, 2010).

Anche *filtrare* (MI 7,488) è un collocato ad alta coesione con *notizia/e* e anch'esso presenta occorrenze alla forma causativa: 5 con FARE e 1 con LASCIARE; i dizionari che contengono *filtrare* sono solo DdC e DCL, che in questo caso si rivelano particolarmente perspicaci.

### 3.4.3 Ampliamento del pattern Soggetto + Verbo

*compare, proviene, raggiunge e risale*

Soggetto + Verbo	Occorrenze CORIS	T. semplici	T. composti
compare	20	7	13
proviene	21	21	
raggiunge	23	16	7
risale	19	19	

I quattro Verbi mostrano un valore significativo di MI rispettivamente alle forme: *risalgono* (MI 6,825 e *risale* MI “notizia”), *provengono* (MI “notizie” 7,377), *raggiunse* (MI “notizia” 6,367) e *comparsa* (MI “notizia” 6,130).

Questi dati trovano corrispondenza nelle concordanze del corpus, dalle quali emerge che la flessione più tipica del Sogg rispetto a COMPARIRE e RAGGIUNGERE è *notizia* singolare e al contrario PROVENIRE e RISALIRE selezionano *notizie* al plurale nei collocati con MI citati sopra.

L'osservazione del contesto sintagmatico rivela che COMPARIRE e RISALIRE non ammettono l'aggiunta di altri elementi tra il Sogg e il V, fatta eccezione per le concordanze di *risalgono* (ad es.: *n. certe risalgono, n. del pittore risalgono n., che la riguardano risalgono*); viceversa, PROVENIRE e RAGGIUNGERE consentono l'inserzione di altri elementi, ad es.: *la n. eclatante proviene, le n. come sempre provengono, n. che a questo (rif. a persona) provenivano, la n. ci raggiunge, la n. della disfatta raggiunse, le n. che potevano interessargli lo raggiungevano*.

*riguardante e concernenti*

Soggetto + Verbo	Occorrenze CORIS		Span							
			- 4	-3	-2	-1	+1	+2	+3	+4
n. riguardante	min. s.	12					9	2*		1 (+5)
*n. diffusa/magari r.	pl.	61					51	1	6	3
	TOT	73								
n. concernenti	min. s.	2					2			
	pl.	11					7		2	2
	TOT	13								

*riguardante* (MI 6,711 “notizia”) e *concernenti* (MI 6,576 “notizie”) mantengono la funzione verbale, diversamente da altri Participi Presenti (ad es. *confortante*, *circolante*), che rientrano nel pattern con Aggettivo in posizione libera (3.1).

RIGUARDARE conta 183 occorrenze, tutte coniugate ai tempi semplici di cui 73 al Partecipio Presente e fra queste la flessione più frequente è quella al plurale (61 plurale vs 12 singolare); l’ordine dei costituenti N+V è fisso, altrimenti nell’ordine inverso V+N, il N da Sogg passa a Oggetto Diretto. La collocazione ammette elementi di diverse categorie al suo interno: *n. essenziali riguardanti*, *n. sulle vicende riguardanti*, *n. e documenti segreti riguardanti*, *n. apparse sulla stampa riguardanti*; riguardo la variabilità paradigmatica *riguardante* con il significato di “avere una relazione” trova proprio in *concernente* un suo sinonimo.

*Concernente/i* conta solo 13 occorrenze e la ricerca nel corpus di *notizia/e* in combinazione con CONCERNERE restituisce solo ed esclusivamente la forma al Partecipio Presente, con la flessione plurale in netta maggioranza rispetto quella singolare; le concordanze di questa collocazione mostrano l’inserimento esclusivo di altri sostantivi, ad es. *notizie e informazioni concernenti*.

In conclusione, sulla base delle misure di associazione si può affermare che le collocazioni più appropriate sono *notizia riguardante*, per la forma al singolare, e *notizie concernenti*, per quella plurale.



## Conclusioni

Il lessico di una lingua non è costituito soltanto dall'insieme delle parole consolidate, da quelle potenziali che in ogni momento possono essere regolarmente formate, né dalle novità assolute che progressivamente vengono introdotte e neppure solo dai vocaboli presi singolarmente: rientrano a tutti gli effetti nel lessico anche le collocazioni o le *callidae iuncturae* come scriveva Orazio, cioè accoppiamenti giudiziari o giunture più stabilizzate nell'uso (De Mauro, 2000: viii), come ad es. *ricco sfondato, cogliere l'occasione, regna la quiete, diametralmente opposto*.

La Linguistica teorica, successivamente all'attenzione iniziale di Firth, si è occupata in modo sistematico dei pattern collocazionali dalla seconda metà del '900 e nel corso degli anni '90 ha incrementato l'interesse per le collocazioni lessicali, fino a portarli al nucleo del dibattito scientifico, perché si sono rivelati un fenomeno tanto pervasivo nella lingua quanto sfuggente e caleidoscopico; proprio in ragione della loro complessità sono stati affrontati da orientamenti di diversa natura, innanzitutto quelli quantitativo, formalista e cognitivo, che però non sono ancora stati in grado di formulare una teoria unitaria e di pervenire a un'impostazione comune.

Questi eventi, impattando sui diversi ambiti della Linguistica applicata, hanno sollevato questioni complesse e generato esigenze inderogabili; nello specifico, nella lessicografia a più riprese sono giunte sollecitazioni circa il reperimento delle collocazioni nei corpora elettronici e il loro inserimento e trattamento adeguato nelle voci dei repertori linguistici, siano essi generali della lingua d'uso, bi-plurilingue, terminologici o combinatori.

Sinclair, fondatore della Corpus Linguistics e del progetto COBUILD, osserva che i pattern collocazionali sono un'area del linguaggio trascurata dai dizionari e servendosi dei principi metodologici *accept the evidence* e *reflect the evidence*, decreta la forza granitica degli esempi genuini della lingua d'uso in contrasto con gli esempi fondati sull'introspezione del lessicografo; l'autorevolezza degli esempi autentici ottenibili dalla crescente mole di dati ricavabili dai corpora si fonda sui criteri oggettivi della significatività statistica: frequenza d'uso e misure di associazione.

Le prove evidenti prodotte dallo spoglio dei corpora e dagli studi cognitivi e psicolinguistici circa il largo impiego di *pre-fabricated chunks* da parte dei parlanti,

impongono l'urgenza ai lessicografi di offrire esempi d'uso provenienti da testi prodotti e riproducibili nella lingua, incrementando così lo spazio dedicato alle collocazioni di un lessema; tuttavia, l'esemplificazione e la rappresentazione del patrimonio collocazionale di una lingua, essendo quest'ultimo un sistema complesso, sfuggente e in continua evoluzione, rischiano di essere inficiate dall'intuizione e dal giudizio umano.

A partire da tali considerazioni, la domanda di ricerca su cui poggia il presente elaborato muove dall'ipotesi che la descrizione dell'italiano offerta nei glossari accusi un certo grado di soggettività e si basi più sull'intuizione e sull'introspezione del compilatore, che sulla rigorosa osservazione dell'uso reale che i parlanti fanno della lingua.

L'approccio quantitativo in parte integrato con quello fraseologico, funzionale alla distinzione delle diverse categorie di pattern collocazionali, e il metodo *corpus-driven* hanno consentito di svolgere il confronto puntuale tra le collocazioni lessicali di una parola-perno ottenute dall'esplorazione di un corpus e quelle censite nei dizionari; si è optato per il *frequency-oriented approach*, perché fissa l'uso e la frequenza come criteri rilevanti: infatti, il largo uso e ri-uso di una collocazione è uno dei fattori che ne decreta il passaggio dall'uso alla norma, "si volet usus / quem penes arbitrium est et ius et norma loquendi" (trad. "se lo vorrà l'uso, nel cui potere stanno le decisioni, il diritto e la norma del parlare", Orazio, *Ars Poetica*, cit. in De Mauro, 2000: viii).

In conseguenza dell'adozione di un approccio ibrido si è proceduto a formulare una definizione funzionale di collocazione. Gli strumenti utilizzati sono stati il corpus di riferimento CORIS e una selezione di dizionari monolingui generali, combinatori e analogici, in formato cartaceo, con l'aggiunta di un dizionario elettronico bilingue italiano-tedesco.

Il limite oggettivo dell'elaborato consiste nella natura stessa dello strumento corpus, che non può essere considerato un campione della lingua in grado di riprodurre la realtà dell'universo dei testi scritti e parlati di un idioma, dal momento che la lingua è inconoscibile nella sua totalità e costantemente in evoluzione; un limite soggettivo è dato dal grado di successo ottenuto dalla capacità di imbrigliare e governare la mole dei dati.

La disanima dei pattern collocazionali è stata articolata in due fasi: il punto di partenza sono le combinazioni statisticamente significative censite dallo spoglio del corpus dal punto di vista delle qualità extralinguistiche (frequenza assoluta e della Mutual

Information); sulla base di questi risultati sperimentali è stato compiuto il confronto con i pattern collocazionali ottenuti dalla ricerca condotta nelle opere lessicografiche.

Il raffronto tra la ricognizione nei repertori e il censimento di un corpus di riferimento dei pattern collocazionali lessicali della parola-perno *notizia* ha consentito innanzitutto di verificare e ampliare la copertura delle collocazioni presenti nei glossari e contestualmente di sviluppare alcune considerazioni riguardo all'utilità e all'opportunità di sfruttare le potenzialità delle risorse elettroniche per ricavare dati autentici ai fini descrittivi del comportamento di un lessema.

Riguardo la verifica e l'ampliamento dei pattern collocazionali, il lemma *notizia* occorre 20.890 volte su 164.599.694 token totali del corpus. La ricognizione effettuata sui repertori ha identificato 293 combinazioni totali, di cui 155 provenienti da Tiberii (2018, DdC), e il censimento condotto sul corpus ha fornito 53 collocati in possesso di valori significativi di Mutual Information (MI, e in alcuni casi anche di un valore significativo in Log-Likelihood, LL), di cui 11 non recensiti nei dizionari e che rappresentano l'ampliamento dei pattern collocazionali. Di seguito vengono riportati, in forma sintetica, i dati e i risultati più significativi distinti in base al tipo di schema collocazionale (n.b. occ = occorrenze):

- **Aggettivo in posizione libera:** 118 combinazioni rintracciate nei repertori di cui 73 provenienti da Tiberii (2018, DdC); gli Aggettivi in posizione libera più citati nei repertori sono *infondata* (9 diz., 24 occ) e *recente* (8 diz, 53 occ), mentre le co-occ più frequenti nel corpus sono *brutta* (236 occ, valori di MI e LL, 6 diz), *falsa* (221 occ, LL, 6 diz) e *utile* (128 occ, LL, 2 diz). Riguardo la tipologia, si hanno: Agg qualificativi, di relazione, Participi Presenti e Passati che ricevono trattamento aggettivale.

Il censimento del corpus ha consentito di mettere in luce il comportamento di alcuni Agg in riferimento a parametri importanti, quali: la forma flessiva (Agg solo al singolare o solo al plurale), l'ordine prevalente dei costituenti (Agg+N oppure N+Agg), l'ampiezza della possibilità dell'inserzione sintagmatica, il preponderante uso in combinazione con specifiche categorie di parti invariabili del discorso. Gli Agg in posizione libera, che rappresentano dei collocati interessanti sono *brutta* (le concordanze mostrano che l'Agg appartiene al pattern Sequenze Obbligate), *triste*

(per la possibile seppur limitata variabilità paradigmatica), *curiosa* (collocato citato solo in Tiberii, 2018) e *scarse* (prevalenza dell'ordine marcato Agg+N).

Lo spoglio del corpus ha fornito 14 collocati per il pattern Aggettivo in posizione libera, di cui 11 compresi in Tiberii (2018, DdC) e 1 in funzione dell'ampliamento della copertura dei dizionari: *priva/e*. *Priva/e* è un collocato ad alta coesione con "notizie", ma dall'osservazione della frequenza delle occorrenze è ben presente anche la flessione al singolare, *priva*. Il significato della collocazione è semi-composizionale e l'ordine dei costituenti è tipicamente N+Agg (span+), viceversa i contesti in cui compare nell'ordine marcato Agg+N, rispondono all'esigenza pragmatica di mettere in rilievo la qualità della *notizia*, operazione possibile per mezzo del focus sintattico, ad es. *definisce prive di ogni fondamento le n.* (span -5). Nell'ordine non marcato le occorrenze con span maggiore o pari +2, consentono l'aggiunta di Agg (*n. discreta e apparentemente priva di...*), Avv (*n. così priva di...*) e proposizioni (*n. che, considerate generalmente prive di...*), oppure l'espressione della funzione predicativa dell'Agg (*la n. è assolutamente priva di...*). A livello sintattico l'Agg è sempre accompagnato da un Complemento di Privazione che ammette la variabilità paradigmatica dei Sostantivi: *interesse, importanza, fondatezza e fondamento; fondamento* co-occorre con *notizia* 16 volte su 25 occorrenze e considerato che possiede un valore significativo di MI in combinazione con *notizia*, si può concludere che *notizia/e priva/e di fondamento* costituisce una collocazione a medio-alta coesione.

- **Sequenze Obbligate:** 40 combinazioni rintracciate nei repertori di cui 17 provenienti da Tiberii (2018, DdC); gli Aggettivi in Sequenza Obbligata più citati nei repertori sono *ultima n.* (13 diz, 254 occ, LL), *buona n.* (9 diz, 1.067 occ, LL) e *n. certa* (8 diz, 45 occ), mentre le co-occ più frequenti nel corpus sono *buona n.* (1.067 occ), *ultima n.* (254 occ) e *cattiva n.* (249 occ); al tempo stesso sono molti gli Agg in Sequenza Obbligata menzionati in un solo dizionario. Le concordanze mostrano una correlazione tra il tipo di Agg e l'ordine dei costituenti in combinazione: Agg qualificativi+N e viceversa N+Agg di relazione. Anche per questo pattern, il censimento del corpus ha permesso di illustrare l'uso di alcune Sequenze Obbligate, in riferimento ai parametri di flessione morfologica e la possibilità di inserzione tra i costituenti. I collocati inclusi fra gli Agg delle Sequenze Obbligate considerati

interessanti sono *ultima/e n.* (per la forte tendenza alla flessione plurale) e *n. riservata/e* (non ammette l'inserzione di altro materiale linguistico fra i costituenti). Lo spoglio del corpus ha fornito 5 collocati per questo pattern, di cui 4 compresi in Tiberii (2018, DdC) e 1 in funzione dell'ampliamento della copertura dei dizionari: *notizia/e scomoda/e*. *Notizia/e scomoda/e* è un collocato ad alta coesione con "notizia|notizie". L'Agg è usato in senso figurato e quasi esclusivamente alla forma plurale; l'ordine dei costituenti è unicamente N+Agg con possibilità di ampliamento sintagmatico, ad es. Verbi con aspetto resultativo, *la n. risulta scomoda*, o Avverbi dal valore concessivo, *la n. utile anche se scomoda*.

- **Verbo + Complemento:** 94 combinazioni rintracciate nei repertori di cui 50 provenienti da Tiberii (2018, DdC); i Verbi più citati nei repertori che presentano *notizia* in funzione di complemento sono *fare n.* (12 diz., 191 occ, valore rilevante in LL), *(non) avere n.* (11 diz, 686 occ, MI) e *dare n.* (10 diz, 586 occ, MI), mentre le co-occ più frequenti nel corpus sono *(non) avere n.* (686 occ), *dare n.* (586 occ) e *diffondere n.* (280 occ, MI, presente in 9 diz); il censimento del corpus ha messo in luce il caso problematico di *fa* (voce del V FARE), l'uso della parola-perno al singolare o al plurale con le relative coniugazioni e qualche caso di variabilità paradigmatica. I pattern collocazionali V+Complemento interessanti hanno riguardato i Verbi ad alternanza transitiva e pronominale (*diffondere ~ si diffonde*, *spargere ~ si sparge*). Lo spoglio del corpus ha fornito 23 collocati, di cui 17 compresi in Tiberii (2018, DdC) e 3 in funzione dell'ampliamento della copertura dei dizionari: *scambiarsi*, *reperire*, e la forma participiale *coperte*. *Scambiarsi* e *reperire* sono le forme dei rispettivi lemmi verbali che presentano un valore rilevante di MI, ad alta forza coesiva il primo e a media il secondo. Se SCAMBIARSI predilige la coniugazione al modo Indicativo e ai tempi semplici, REPERIRE nel corpus si trova solo all'Infinito Presente in combinazione con i lemmi POTERE (*si possono reperire recensioni e notizie*), CONSENTIRE (*consente di reperire facilmente notizie*), oppure come Infinito Sostantivato (*per/di reperire notizie*). Nelle concordanze entrambi i V selezionano solo la flessione plurale (*notizie*) usata con ellissi dell'articolo determinativo. Ulteriori ampliamenti del pattern V+Complemento e particolarmente degni di nota dal punto di vista lessicografico, riguardano *propagare* e *sgonfiare* ad alternanza transitiva e pronominale; *propagare ~ si propaga* non sono collocati veri

e propri di *notizia*, perché non raggiungono il valore minimo restituito dalle misure di associazione, tuttavia, il V alla forma transitiva è recensito in Tiberii (2018, DdC) e Urzi (2009, DCL), anche se il corpus restituisce solo concordanze alla forma pronominale, che invece è assente in tutti i dizionari; lo stesso ragionamento vale per *sgonfiare* ~ *si sgonfia*, recensito alla forma transitiva da DCL e Simone (2010, DAU), ma le 2 occorrenze rinvenute sono alla forma pronominale e non presenti nelle combinazioni dei repertori. Il Participio Passato *coperte* possiede una forza di coesione media con “notizie” e dallo spoglio condotto sul corpus, si osserva un dato interessante: l’uso preponderante del V *coprire* alla diatesi passiva, mentre la coniugazione attiva appare solo in 2 concordanze (*le emittenti coprono le n. e nessun segreto possa coprire fatti, notizie o documenti concernenti reati*). La collocazione ammette l’inserzione di Avv (*la n. era stata a lungo coperta dalle autorità sanitarie*), anche in funzione di modificatori dell’Agg (*la n. politicamente rilevante, non coperta da segreto di Stato*). Nelle concordanze della costruzione passiva, il Compl. d’Agente preponderante è espresso dalla parola *segreto*, sempre seguito da un Agg qualificativo (*professionale, istruttorio, investigativo, d’indagine*) o da un Complemento di Specificazione (*di Stato*): si può affermare quindi, l’uso più probabile della collocazione *coprire la/le n.* in entrambe le diatesi richieda il N *segreto* in qualità di Sogg logico.

- **Soggetto + Verbo:** in ambito lessicografico è il pattern più negletto fra tutti gli schemi e i repertori mostrano discordanze e disomogeneità rispetto alle combinazioni inserite negli articoli del lemma; sono 41 le combinazioni rintracciate nei repertori di cui 15 provenienti da Tiberii (2018, DdC); i Verbi più citati nei repertori che presentano *notizia* in funzione di soggetto sono *corre* (8 diz., 17 occ.) e *giunge* (7 diz, 329 occ, valori significativi in MI e LL), mentre le co-occ più frequenti nel corpus sono *fa (molto)scalpore/rumore/clamore* (764 occ, LL, 2 diz), *arriva* (624 occ, MI e LL, citato solo in Tiberii (2018, DdC) e *giunge* (329 occ, 7 diz); le evidenze del corpus hanno messo in luce la variabilità paradigmatica di alcune collocazioni riscontrata anche nei repertori. I Verbi collocati di *notizia*, che si sono rivelati interessanti sono, *trapela* e *filtra* (usati molto spesso in costruzione causativa).

Lo spoglio del corpus ha fornito 17 collocati, di cui 7 compresi in Tiberii (2018, DdC) e 6 in funzione dell’ampliamento della copertura dei dizionari: *compare, proviene,*

*raggiunge, risale, concernenti e riguardanti*. Per quanto concerne *compare, proviene, raggiunge, risale*, i quattro Verbi hanno un valore significativo di MI rispettivamente alle forme: *risalgono e risale, provengono, raggiunse, comparsa*. Questi dati trovano corrispondenza nelle concordanze del corpus, dalle quali emerge che la flessione più tipica del Sogg rispetto a COMPARIRE e RAGGIUNGERE è *notizia* singolare, mentre PROVENIRE e RISALIRE selezionano *notizie* al plurale. L’osservazione del contesto sintagmatico rivela che COMPARIRE e RISALIRE non ammettono l’aggiunta di altri elementi tra il Sogg e il V, fatta eccezione per le concordanze con *risalgono* (ad es.: *n. certe risalgono, n. del pittore risalgono n., che la riguardano risalgono*); viceversa, PROVENIRE e RAGGIUNGERE consentono l’inserzione di altri elementi, ad es.: *la n. eclatante proviene, le n. come sempre provengono, n. che a questo (rif. a persona) provenivano, la n. ci raggiunge, la n. della disfatta raggiunse, le n. che potevano interessargli lo raggiungevano*. Risultano poi rilevanti *riguardante e concernenti*. *Riguardante* (MI 6,711 “notizia”) e *concernenti* (MI 6,576 “notizie”) sono collocati di media coesione e, diversamente da altri Participi Presenti (ad es. *confortante, circolante*), mantengono la funzione verbale. Le occorrenze di RIGUARDARE sono tutte coniugate ai tempi semplici di cui la metà circa al Partecipio Presente e fra queste la flessione più frequente è quella al plurale; l’ordine dei costituenti N+V è fisso, altrimenti nell’ordine inverso V+N il N da Sogg passa a Oggetto Diretto; la collocazione ammette elementi di diverse categorie al suo interno: *n. essenziali riguardanti, n. sulle vicende riguardanti, n. e documenti segreti riguardanti, n. apparse sulla stampa riguardanti*; quanto alla variabilità paradigmatica *riguardante* con il significato di “avere una relazione” trova proprio in *concernente*, un suo sinonimo. Per quel che riguarda *concernente/i*: la ricerca nel corpus di *notizia/e* in combinazione con CONCERNERE restituisce solo ed esclusivamente la forma al Partecipio Presente, con la flessione plurale in netta maggioranza rispetto quella singolare; le concordanze di questa collocazione mostrano l’inserimento esclusivo di altri Sostantivi, ad es. *notizie e informazioni concernenti*. In conclusione, sulla base delle misure di associazione si può affermare che le collocazioni più frequenti e appropriate sono *notizia riguardante*, per la forma al singolare, e *notizie concernenti*, per quella plurale.

Il raffronto tra la ricognizione nei repertori e il censimento di un corpus di riferimento dei pattern collocazionali lessicali di *notizia* ha permesso inoltre, di sviluppare alcune considerazioni generali riguardo al grado del successo raggiunto dai dizionari nel presentare le collocazioni più frequenti e statisticamente significative, cioè quelle che mostrano la parola usata nei contesti che meglio ne risaltano il senso.

Complessivamente, i dizionari combinatori e quelli analogici presentano un alto livello di rappresentatività delle collocazioni, spiegabile con una duplice motivazione: i primi dichiarano esplicitamente di aver effettuato la selezione dei collocati su risultati ottenuti dallo spoglio dei corpora (tranne Russo, 2010, MdD) ed entrambi i tipi svolgono una funzione complementare rispetto ai dizionari generali, di cui integrano i dati; nello specifico, fra i dizionari di collocazioni, Tiberii (2018, DdC) oltre che procurare il più alto numero di combinazioni, è anche quello che recensisce il più alto numero di pattern statisticamente significativi, ma è penalizzato dalla mancanza di esempi in contesto; tra i dizionari analogici, il Simone (2010, DAU) si distingue per la menzione di buon numero di collocati, anche se citati in forma lemmatizzata, siano essi disseminati tra i diversi “segmenti” o inseriti nel “segmento” dedicato alle combinazioni tipiche di parole con *notizia*.

I dizionari generali, per loro natura non possono rendere conto di una gran mole di pattern o di indicazioni dettagliate su di essi, ma nella selezione di glossari generici considerati nel presente lavoro si può ravvisare un impiego limitato delle conoscenze raggiunte sull’argomento, riscontrabile in primo luogo nel non riconoscimento del fenomeno linguistico delle collocazioni nei metatesti di queste opere lessicografiche (presentazioni e guide per la consultazione) e in secondo luogo nell’inclusione di esempi di combinazioni infrequenti o cadute in disuso, indizio che il compilatore abbia fatto affidamento più sul proprio giudizio e introspezione, piuttosto che attingere da un bacino di dati oggettivi. A livello di micro-struttura, l’identificazione immediata del carattere collocazionale di un esempio d’uso risulta difficile se non impossibile, dato che la maggioranza dei repertori non isola o non segnala i dati più significativi; fra tutti, fa eccezione Palazzi Folena (1992, PF), che all’interno della voce include la sezione “Nomenclatura”, nella quale sono delimitati i vocaboli che possono figurare assieme alla parola-lemma.



Pur alla luce dei progressi fatti per arricchire qualitativamente e quantitativamente le entrate dei lemmi con pattern convenzionali, la situazione non è ancora soddisfacente e stride la non coincidenza tra intenti dichiarati e dati riportati. La fraseologia, di cui le collocazioni fanno parte, è un elemento indispensabile per integrare la definizione e per rendere evidenti i significati e gli usi della voce, però al fine di testimoniare l'uso corrente concretamente realizzabile nella lingua scritta e parlata, sarebbe auspicabile fondare la fraseografia su testi genuini e non su esempi preparati appositamente dalla redazione; nel processo semi-automatico dell'estrazione dei dati da un corpus, il ruolo dell'introspezione del lessicografo non viene messo al bando, semplicemente è orientato verso la decodifica dei dati estratti e il prendere decisioni circa il modo di inserirli nei repertori sottoforma di esempi o in apposite sezioni delle entrate lessicale. I dizionari poveri di collocazioni non sono cattivi dizionari, ma sicuramente uno dei pregi di un buon dizionario è la capacità di fornire anche informazioni riguardo alla frequenza d'uso e la forza di coesione tra lessemi, attraverso numerosi esempi che mettano in risalto le co-occorrenze istituzionalizzate dall'uso.

Con questa ricerca è stato svolto un approfondimento empirico e lessicografico che è riuscito a evidenziare come i principi metodologici *accept the evidence* e *reflect the evidence* della Corpus Linguistics siano imprescindibili nell'attività lessicografica.

In tal senso, le possibili linee future di indagine sono l'applicazione della metodologia ad altri lessemi, il confronto interlinguistico e la traduzione dei pattern collocazionali lessicali della parola-perno *notizia*, oppure, l'impiego nell'apprendimento delle lingue, considerato che i pattern collocazionali rientrano pienamente nel tema generale dell'insegnabilità del lessico.



## **Appendice**

- 1. Schermata iniziale Corpus CORIS e comando di ricerca**
- 2. Risultato delle concordanze restituite dal Corpus CORIS**
- 3. Ricognizione pattern collocazionali nei repertori in ordine alfabetico**
- 4. Ricognizione pattern collocazionali nei repertori in ordine di frequenza delle co-occorrenze restituite dal corpus**
- 5. Verifica e descrizione dei pattern collocazionali recensiti nei repertori sulla base dei risultati dello spoglio del corpus**
- 6. Ampliamento e descrizione dei pattern collocazionali non recensiti nei repertori e ricavati dallo spoglio del corpus**
- 7. Classifiche dei collocati di *notizia* ottenute dal calcolo delle misure di associazione**

## 1. Schermata iniziale Corpus CORIS e comando di ricerca

Corpus CORIS, annotated version (2021, 165Mw)  
- Corpus query form -

<b>User Authentication</b> CORIS access is free for research purposes (Please, read the footnote carefully). Now you can search CORIS specifying the Time Slice and/or the SubCorpus also for Monitor corpora.	<b>Query</b> (Query Language Help) "notizia" [{"0,4}] "curiosa" Time Slice: All Subcorpus: All
<b>Concordance Options</b> Show <input checked="" type="radio"/> 30 <input type="radio"/> 100 <input type="radio"/> 300 <input type="radio"/> 1000 lines.	Sort position: Unsorted
<b>Collocations</b> Get Collocates? <input checked="" type="radio"/> NO! <input type="radio"/> Yes.	Sort using <input type="radio"/> Log-Likelihood Ratio. <input checked="" type="radio"/> Mutual Information. <input type="radio"/> T-score. <input type="radio"/> Raw frequency.

Esegui Cancell

Interrogando il corpus CORIS/CODIS tramite questa procedura, l'utente dichiara e accetta che l'interrogazione è volta unicamente a scopi di ricerca scientifica e che non ne sarà tratto alcun beneficio economico. L'accesso al corpus è concesso esclusivamente per scopi di ricerca scientifica.  
By querying the CORIS/CODIS corpus through this procedure, the user declares and accepts that the query is for scientific research purposes only and that no economic benefit will be derived from it. Access to the corpus is granted exclusively for scientific research purposes.

[CORIS Credits](#)

## 2. Risultato delle concordanze restituite dal Corpus CORIS

### Corpus query result

Query: "notizia" [{"0,4}] "curiosa"

Sorted: 0

Number of concordances: 14 / 14

[CORIS1980\\_2000](#): ria dove è ancora peggio ( una <notizia curiosa> in pellicceria adesso hanno u  
[CORIS1980\\_2000](#): he , senza dimenticare qualche <notizia curiosa> come quelle sulla nascita del  
[MONITOR2005\\_07](#): uomo , beh questa sì che è una <notizia curiosa> . L ' idea di conoscersi lavo  
[MONITOR2005\\_07](#): mobilità cittadina 4 5 2005 La <notizia è già curiosa> di per sé , ma in tempi di cr  
[MONITOR2008\\_10](#): one o il divorzio . Infine una <notizia curiosa> che viene da Londra e che seg  
[MONITOR2008\\_10](#): .it Grazie per i dati e per la <notizia sulla curiosa> decisione del Garante della p  
[MONITOR2011\\_13](#): siderare l ' episodio solo una <notizia di cronaca curiosa> , altrettanto errato prenderl  
[CORIS1980\\_2000](#): dilungato nel commento di una <notizia curiosa> . L ' esperto era a conoscenz  
[CORIS1980\\_2000](#): ... ASTI La molto cantata Una <notizia curiosa> : " In Sicilia , i balli trad  
[CORIS1980\\_2000](#): , ora 70mila C ' è un ' altra <notizia curiosa> : il giorno 21 febbraio si è  
[CORIS1980\\_2000](#): do pochi anni di vita è ancora <notizia fresca e curiosa> per il pubblico non specialis  
[CORIS1980\\_2000](#): è un particolare che rende la <notizia ancora più curiosa> , trasformandola in una roman  
[CORIS1980\\_2000](#): proprio contro\_di noi " . Ma la <notizia più curiosa> arriva da Roma : la Caf ha ac  
[CORIS1980\\_2000](#): ncherebbe - o si ripiega sulla <notizia curiosa> . Del genere l ' ingegner Fra

### 3. Ricognizione pattern collocazionali nei repertori in ordine alfabetico

Aggettivi	Ddc DCI	DCL	MdD	DO	GraDIt	PF	SC	NTr	Zin	GDIH	SCAN	ThTr	DAU	ELDIT	CORIS
abbondanti								x							1
agghiacciante	x		x											c	4
allarmante	x	x pl									x	x			20
allarmistiche		x													9
attendibile	x	x				x			x		x	x			16
attesa	x										x				16
avvilente	x														-
brutta	x		x				x	x				x	x	c	236
calda													x		-
calunniosa													x		4
catastrofica	x	x pl											x		7
circolante													x		1
clamorosa	x	x	x												24
comunicabile													x		-
confortante	x	x									x		x		28
confusa		x									x				16
consolante	x	x									x				4
contraddittoria	x										x	x			11
contrastanti		x													14
credibile	x	x											x		5
cruda	x														-
curiosa	x														26
deprimente	x														3
dettagliata	x														29
diffamante													x		1
dirimpente	x														-
documentata	x														5
dolorosa	x	x						x						c	7
drammatica	x														14
dubbia		x pl									x				3
eclatante		x													3
elementare								x pl						x	-
esatta						x					x			e	5
esauriente									x pl						2
falsa	x	x pl		x pl				x pl			x		x	e	221
fausta											x				-
faziosa													x		-
felice	x														2
ferale		x		x											6
fondata	x	x					x				x		x		4
frammentaria								x						x	16
fresca	x	x pl				x			x		x		x	e	31
funesta	x	x						x pl				x pl			4
fuorviante	x														3
generica	x														4
ghiotta	x												x		11
gradita	x	x													4
grave	x					x					x				7
importante	x		x		x	x					x			e	114
imprecise		x	x sing												-
impressionante											x				2
inaspettata	x	x	x									x		x	8
inattendibile	x	x pl					x						x		4
inattesa	x	x										x			16
incerta	x	x pl	x			x					x				10
inconsistente													x		-
incontrollabili		x											x		2
incontrollate		x													1
incoraggiante	x	x pl													16
incredibile	x		x												10
inedita	x	x pl											x		9
inesatte		x													14
infamante													x		2
infausta		x									x				1
infondata	x	x pl	x		x	x	x		x			x	x		24
ingiuriosa													x		-

inquietante	x																9
interessante	x pl								x pl							x	65
inverosimili		x pl															3
lieta	x	x		x				x									19
luttuosa		x															-
magnifica										x							2
malsicura													x				-
meravigliosa											x						11
nefasta	x	x															1
negativa	x															e	19
parziale	x																7
piacevole	x	x				x					x						4
positiva	x																60
precisa						x					x						55
preoccupante	x	x pl											x				22
rasserente		x															-
rassicurante	x	x													x		22
recente	x	x pl	x			x		x pl			x	x pl			x		53
rilevante														x			20
rincorante	x																-
sbalorditiva												x					4
scama	x	x pl															14
scarse	x		x agg+n	x agg+n	x agg+n				x agg+n		x agg+n				xagg+n		13
scioccante		x															2
sconcertante	x													x			10
sconfortante	x																5
sconvolgente	x	x	x														22
scottante	x	x											x				1
sensazionale	x	x						x					x	x	x		27
sgradita		x															2
sicura		x										x		x			17
significativa	x																4
singolare	x																4
soprendente	x		x														4
specifica	x															x	6
spiacevole	x	x				x					x	x					4
splendida	x												x				13
stantia		x pl				x					x						-
strabilante		x															3
stringata	x																-
stupefacente		x									x						3
tendenziosa	x	x pl		x pl		x						x	x	x	x		16
terribile	x					x					x				x		34
tragica	x	x											x				27
tranquillizzanti		x												x			6
tremenda	x												x			e	6
triste	x	x											x			e	28
utile	x	x pl															128
vaga									x							x	10
vecchia	x	x pl							x				x				12
vera	x	x		x pl		x	x pl					x				e	118
veritiera	x																4
<b>Sequenze obbligate</b>																	
bella n.	x		x		x	x	x						x	x	x pl		149
buona n.	x		x	x		x	x	x pl		x		x			x		1067
cattiva n.	x					x	x					x		x	x		249
gioiosa n.		x															1
ottima n.			x														42
ultima n.	x	x pl	x pl	x pl	x pl	x pl	x pl	x pl	x pl	x pl			x pl	x	x		254
prime notizie									x								92
n. accreditata		x															1
n. angosciata			x														2
n. anticipata			x														11
n. bomba (inv)	x	x											x	x			32

n. certa	x		x	x pl	x pl	x pl	x						x	x			e pl	45	
n. confidenziale			x																4
n. culturale	x													x					3
n. economica	x									x pl				x	x				16
n. esclusiva	x																		-
n. esplosiva			x																3
n. flash																x			8
n. frita/rifrita													x						-
n. improvvisa																	e		7
n. inventata	x													x	x				7
n. lampo					x					x						x			1
n. pesante					x														1
n. politica	x				x									x	x				16
n. riservata	x		x											x	x				63
n. scandalistica	x													x			e		-
n. shock																x			8
n. scontata																	x		2
n. segreta	x				x									x				x	12
n. spaventosa					x													e	2
n. sportiva	x									x pl		x pl	x pl	x	x		x pl		17
n. stupenda					x														2
n. telefonica												x							-
n. telegrafica												x							-
n. top secret					x														1
n. trionfale																			1
n. ufficiale	x		x		x									x	x	x	x		32
n. ufficiosa	x		x pl											x	x	x	x		7
notizie bibliografiche					x	x sing	x	x pl	x pl			x						x	9
n. esagerate										x									8
<b>Verbo+Complemento</b>	<b>Ddc DCI</b>	<b>DCL</b>	<b>MdD</b>	<b>DO</b>	<b>GraDI</b>	<b>PF</b>	<b>SC</b>	<b>NTr</b>	<b>Zin</b>	<b>GDIH</b>	<b>SCAN</b>	<b>ThTr</b>	<b>DAU</b>	<b>ELDI</b>	<b>CORIS</b>				
abbondare di notizie	x																		-
accertare	x													x					5
accogliere	x																e		65
amplificare	x													x					1
annunciare	x														x				22
(pre)annunziare																			-
apprendere	x																		-
appurare																			-
ascoltare le notizie		x																	43
aspettare	x																		52
assumere																			9
attendere	x																		35
attingere (alle) notizie																			22
(non) avere di	x																		686
avvalorare	x																		-
avvertire con																			-
avvisare con																			-
cantare																			-
censurare	x																		7
chiedere notizie	x																		274
commentare	x																		104
comunicare	x																		102
confermare	x																		88
congetturare																			-
conoscere																			9
controllare	x																		11
dare (proprie)	x																		586
dare rilievo a																			2
dare risalto a	x																		12
diffondere	x	x																	280
diramare	x																		13
divulgare	x																		73
domandare																			10
essere confortato/ecc. da	x																		34

essere senza notizie										x										3	
falsificare												x								-	
far avere						x pl		x						x						8	
fare notizia	x	x	x	x	x	x	x	x	x					x	x	x				191	
fornire	x			x											x					163	
gonfiare	x		x					x							x					5	
ingigantire	x														x					2	
informare con															x					2	
intercettare															x					2	
inventare															x	e				6	
irradiare															x					-	
lasciare senza									x											1	
leggere	x	x			x		x	x	x						x					158	
mandare proprie notizie								x	x									x		21	
manipolare			x												x					7	
mettere in circolazione	x														x					1	
mettere in giro	x														x					4	
montare			x					x							x					2	
oscurare	x																			1	
partecipare							x					x								-	
pescare							x					x								3	
pompate			x												x					2	
portare			x					x		x										135	
preannunciare														x						-	
prendere																		e		21	
preparare a (cattiva n.)																		e		1	
propagare	x		x																	2	
propalare			x					x	x					x						2	
possedere				x di			x			x								x		6	
pubblicare	x	x	x			x		x	x									x		181	
raccolgere (le) notizie	x	x					x	x	x				x	x	x	x				110	
rallegrarsi per	x																			-	
rattristarsi per	x																			-	
recare	x	x						x												17	
rettificare	x	x														x				2	
ricavare notizie da									x									x		22	
ricevere (da)	x			x	x	x				x			x	x±				x		250	
richiedere			x																	23	
riferire (per doverosa n.)	x	x	x+				x	x					x					e		66	
riportare	x	x	x					x					x			x		e		183	
rivelare	x	x	x													x				31	
sapere														x	x		e			71	
sentire (alla radio/televisione)	x			x		x pl		x								x				87	
sgonfiare			x													x				-	
smentire	x	x	x	x				x								x				76	
smontare	x	x														x				1	
spargere								x					x			x				21	
spacciare			x																	2	
sobbalzare a	x																			-	
strillare				x																4	
tacere	x															x				11	
tenere nascosta	x																			18	
tenere per sé	x																	x		2	
trasmettere	x	x				x		x	x	x			x			x				62	
udire																x				6	
vendere	x															x				6	
venire a conoscenza														x		x				7	
verificare	x																	e		29	
vivere				x																	
assodata			x																	-	
<b>Soggetto+Verbo</b>	<b>Ddc</b>	<b>DCI</b>	<b>DCL</b>	<b>MdD</b>	<b>DO</b>	<b>GraDit</b>	<b>PF</b>	<b>SC</b>	<b>NTr</b>	<b>Zin</b>	<b>GDH</b>	<b>SCAN</b>	<b>ThTr</b>	<b>DAU</b>	<b>ELDI</b>	<b>CORIS</b>					
affiora			x pl					x												2	
allama	x													x						1	
arriva	x																			624	



avverte											x					-
avvisa											x					1
circola	x		x pl				x	x			x		x		c	72
coincidono															c	1
colpisce															c	14
conferma															c	9
corre/correa	x		x				x	x	x		x	x	x			17
crea scalpore/ecc.*	x														c	53
danneggia	x															-
desta scalpore/ecc*															c	40
è apparsa sul giornale			x													49
è uscita sul giornale			x													28
fa (molto) scalpore/ecc*			x								x+					764
filtra	x		x													49
giunge (nuova)	x		x +				x	x	x		x			x a		329
ha fatto (il giro del mondo)			x													48
ha scioccato			x													1
informa							x				x	x				1
interessa												x				16
lascia senza fiato/parole	x															3
mancano notizie			x					x								15
è pervenuta (perviene)							x				x					18
riempie di gioia/angoscia/ecc															c	5
rimbalza (in)	x		x+											x		56
scarseggiano								x								-
sconvolge	x														c	16
scotta														x		-
scuote	x	x														10
si diffonde														x	x	48
si rivela															c	3
si sparge														x		59
sorprende	x															12
suscita scalpore/ecc*														x		61
si susseguono			x					x								7
si incalzano			x													-
tranquillizza	x															5
trapela	x		x	x far			x	x	x lasciar		x	x lasciar	x			73
vola														x		-
* scalpore, allarme, clamore, panico, rumore, sgomento																

#### 4. Ricognizione pattern collocazionali nei repertori in ordine di frequenza delle co-occorrenze restituite dal corpus

Aggettivi	Dde DCI	DCL	MID	DO	GrADit	PF	SC	NTr	Zin	GDIH	SCAN	ThTr	DAU	ELDT	CORIS
brutta	x		x				x	x				x	x	c	236
falsa	x	x pl		x pl				x pl			x		x	c	221
utile	x	x pl													128
vera	x	x		x pl			x	x pl			x			c	118
importante	x		x		x	x					x			c	114
interessante		x pl								x pl				x	65
positiva	x														60
precisa						x					x				55
recente	x	x pl	x			x		x pl			x	x pl		x	53
terribile	x					x					x			x	34
fresca	x	x pl				x			x		x		x	c	31
dettata	x														29
confortante	x	x									x		x		28
triste	x	x										x		c	28
sensazionale	x	x						x				x	x	x	27
tragica	x	x										x			27
curiosa	x														26
clamorosa	x	x	x												24
infondata	x	x pl	x		x	x	x		x			x	x		24
preoccupante	x	x pl										x			22
rassicurante	x	x										x			22
sconvolgente	x	x	x												22
allarmante	x	x pl										x	x		20
rilevante													x		20
lieta	x	x		x				x							19
negativa	x													c	19
sicura		x									x		x		17
attendibile	x	x					x		x			x	x		16
attesa	x											x			16
confusa		x									x				16
frammentaria								x						x	16
inattesa	x	x										x			16
incoraggiante	x	x pl													16
tendenziosa	x	x pl		x pl			x					x	x	x	16
contrastanti		x													14
drammatica	x														14
inesatte		x													14
scarna	x	x pl													14
scarse	x		x agg+n		x agg+n		x agg+n			x agg+n		x agg+n		xagg+n	13
splendida	x											x			13
vecchia	x	x pl						x			x				12
contraddittoria	x										x	x			11
ghiotta	x												x		11
meravigliosa											x				11
incerta	x	x pl	x			x					x				10
incredibile	x		x												10
sconcertante	x												x		10
vaga								x						x	10
allarmistica		x													9
inedita	x	x pl											x		9
inquietante	x														9
inaspettata	x	x	x									x		x	8
catastrofica	x	x pl											x		7
dolorosa	x	x						x						c	7
grave	x					x					x				7
parziale	x														7
ferale		x		x											6
specifica	x													x	6
tranquillizzanti		x											x		6
tremenda	x											x		c	6
credibile	x	x											x		5
documentata	x														5
esatta						x					x			c	5
sconfortante	x														5
agghiacciante	x		x											c	4
calunniosa													x		4
consolante	x	x									x				4

fondata	x	x					x				x	x			4	
funesta	x	x						x pl				x pl			4	
generica	x														4	
gradita	x	x													4	
inattendibile	x	x pl					x						x		4	
piacevole	x	x					x					x			4	
sbalorditiva												x			4	
significativa	x														4	
singolare	x														4	
sorprendente	x		x												4	
spiacevole	x	x					x					x	x		4	
veritiera	x														4	
deprimente	x														3	
dubbia		x pl										x			3	
eclatante		x													3	
fuorviante	x														3	
inverosimili		x pl													3	
strabiliante		x													3	
stupefacente		x										x			3	
esauriente										x pl					2	
felice	x														2	
impressionante												x			2	
incontrollabili		x											x		2	
infamante													x		2	
magnifica												x			2	
scioccante		x													2	
sgradita		x													2	
abbondanti								x							1	
circolante													x		1	
diffamante													x		1	
incontrollate		x													1	
infausta		x										x			1	
nefasta	x	x													1	
scottante	x	x										x			1	
avvilente	x														-	
calda													x		-	
comunicabile													x		-	
cruda	x														-	
dirompente	x														-	
elementare								x pl						x	-	
fausta												x			-	
faziosa													x		-	
imprecisate		x	x sing												-	
inconsistente													x		-	
ingiuriosa													x		-	
luttuosa		x													-	
malsicura													x		-	
raserenante		x													-	
rincuorante	x														-	
stantia		x pl					x					x			-	
stringata	x														-	
<b>Sequenze obbligate</b>	<b>Ddc</b>	<b>DCI</b>	<b>DCL</b>	<b>MdD</b>	<b>DO</b>	<b>Gradit</b>	<b>PF</b>	<b>SC</b>	<b>NTr</b>	<b>Zin</b>	<b>GDH</b>	<b>SCAN</b>	<b>ThTr</b>	<b>DAU</b>	<b>ELDIT</b>	<b>CORIS</b>
buona n.	x		x	x	x		x	x	x pl		x	x			x	1067
ultima n.	x	x pl	x pl	x pl	x pl	x pl		x pl	x pl	x pl	x pl		x pl	x	x	254
cattiva n.	x						x pl	x				x			x	249
bella n.	x		x		x			x					x	x	x pl	149
prime notizie										x						92
n. riservata	x	x											x	x		63
n. certa	x	x	x pl	x pl	x pl	x						x	x		c pl	45
ottima n.			x													42
n. bomba (inv)	x	x											x	x		32
n. ufficiale	x	x	x									x	x	x	x	32
n. sportiva	x							x pl		x pl	x pl		x	x	x pl	17
n. economica	x							x pl					x	x		16
n. politica	x		x										x	x		16
n. segreta	x		x										x		x	12
n. anticipata			x													11

notizie bibliografiche			x	x sing	x	x	x		x					x	9
n. esagerate									x						8
n. flash													x		8
n. shock													x		8
n. improvvisa														c	7
n. inventata	x											x	x		7
n. ufficiosa	x		x pl								x	x	x	x	7
n. confidenziale			x												4
n. culturale	x											x			3
n. esplosiva			x												3
n. angosciosa				x											2
n. scontata														x	2
n. spaventosa				x										c	2
n. stupenda				x											2
n. accreditata			x												2
gioiosa n.			x												1
n. lampo				x					x				x		1
n. pesante				x											1
n. top secret			x												1
n. trionfale				x											1
n. esclusiva	x														-
n. fritta/rifritta											x				-
n. scandalistica	x											x		c	-
n. telefonica									x						-
n. telegrafica									x						-
<b>Verbo+Complemento</b>	<b>Ddc DCI</b>	<b>DCL</b>	<b>MdD</b>	<b>DO</b>	<b>GraDt</b>	<b>PF</b>	<b>SC</b>	<b>NTr</b>	<b>Zin</b>	<b>GDIH</b>	<b>SCAN</b>	<b>ThTr</b>	<b>DAU</b>	<b>ELDIT</b>	<b>CORIS</b>
(non) avere n.	x		x ± di	x	x± da/su		x	x di qsa ±x± di qno	x± più			x	x	x±	686
dare (proprie)	x		x una	x		x x	x	x pl +		x			x±	x	586
diffondere	x	x	x	x		x x	x				x		x	x	280
chiedere notizie	x		x			x x					x		x	c	274
ricevere (da)	x		x	x	x			x			x	x±		x	250
fare n.	x	x	x	x	x	x	x	x				x	x	x	191
riportare	x		x	x			x				x		x	c	183
pubblicare	x		x	x		x	x	x						x	181
apprendere	x		x	x		x x					x		x	c	173
fornire	x		x										x		163
leggere	x	x			x	x	x	x					x		158
portare			x				x		x						135
raccogliere (le) notizie	x		x			x	x	x			x	x	x	x	110
commentare	x												x	c	104
comunicare	x		x			x x	x	x			x		x	c	102
confermare	x		x	x	x	x x					x	x	x	c	88
sentire (alla radio/televisione)	x		x		x pl		x						x		87
smentire	x		x	x	x		x						x		76
divulgare	x		x			x x	x				x		x		73
sapere												x	x	c	71
riferire (per do verosa n.)	x		x	x+		x x					x			c	66
accogliere	x		x											c	65
trasmettere	x		x	x		x x	x				x		x	x	62
aspettare	x		x			x					x				52
ascoltare le notizie		x			x	x						x		x	43
attendere	x					x					x				35
essere confortato/ecc. da	x						x								34
rivelare	x		x	x									x		31
verificare	x													c	29
richiedere			x												23
annunciare	x		x									x			22
attingere (alle) notizie			x			x				x+	x				22
ricavare notizie da							x							x	22
mandare proprie notizie							x	x						x	21
prendere														c	21
spargere			x			x					x		x		21
tenere nascosta	x														18
recare	x		x				x								17
diramare	x		x				x						x		13
dare risalto a	x														12

controllare	x													x		11
tacere	x													x		11
domandare						x					x					10
assumere						x					x					9
conoscere															c	9
far avere						x pl		x				x				8
deensurare	x			x												7
manipolare			x											x		7
venire a conoscenza												x	x			7
inventare													x		c	6
possedere				x di		x		x							x	6
udire														x		6
vendere	x													x		6
accertare	x			x										x		5
gonfiare	x		x					x						x		5
mettere in giro	x													x		4
strillare				x												4
essere senza notizie										x						3
pescare						x						x				3
dare rilievo a															c	2
informare con														x		2
ingigantire	x													x		2
intercettare														x		2
montare			x					x						x		2
pompare			x											x		2
propagare	x		x													2
propalare			x			x	x					x				2
rettificare	x		x											x		2
spacciare			x													2
tenere per sé	x															2
appurare			x					x								2
amplificare	x													x		1
lasciare senza										x						1
mettere in circolazione	x													x		1
oscurare	x															1
preparare a (cattiva) n.															c	1
smontare	x		x											x		1
abbondare di notizie	x															-
(pre)annunziare								x±						x		-
avvalorare	x															-
avvertire con														x		-
avvisare con														x		-
cantare				x												-
congetturare								x						x		-
falsificare														x		-
irradiare														x		-
partecipare								x						x		-
preannunciare														x		-
rallegrarsi per	x															-
rattristarsi per	x															-
sgonfiare				x											x	-
sobbalzare a	x															-
vivere															x	-
n. associata				x												-
<b>Soggetto+Verbo</b>	<b>Ddc</b>	<b>DCI</b>	<b>DCL</b>	<b>MdD</b>	<b>DO</b>	<b>GraDh</b>	<b>PF</b>	<b>SC</b>	<b>NTr</b>	<b>Zin</b>	<b>GDIH</b>	<b>SCAN</b>	<b>ThTr</b>	<b>DAU</b>	<b>ELDIT</b>	<b>CORIS</b>
fa (molto) scalpore		x										x+				764
arriva	x															624
giunge (nuova)	x		x+			x	x	x				x		x a		329
trapela	x		x	x far			x	x	x lasciar			x	x lasciar	x		73
circola	x		x pl				x	x				x		x	c	72
suscita														x		61
si sparge														x		59
rimbalza (in)	x		x+											x		56
crea scalpore/ecc.*	x														c	53
è apparsa sul giornale		x														49

filtra	x		x																	49	
ha fatto (il giro del mondo)		x																		48	
si diffonde													x							x	48
desta scalpore/ecc																				c	40
è uscita sul giornale		x																			28
è pervenuta									x				x								18
corre/correa	x		x					x	x	x			x	x	x						17
interessa														x							16
sconvolge	x																			c	16
mancano notizie			x						x												15
colpisce																				c	14
sorprende	x																				12
scuote	x	x																			10
conferma																				c	9
si susseguono			x					x													7
riempie di gioia/angoscia/ecc																				c	5
tranquillizza	x																				5
lascia senza fiato/parole	x																				3
si rivela																				c	3
affiora			x pl					x													2
allarma	x																			x	1
avvisa														x							1
coincidono																				c	1
ha scioccato		x																			1
informa								x					x	x							1
avverte														x							-
danneggia	x																				-
scarseggiano									x												-
scotta																				x	-
si incalzano			x																		-
vola																				x	-
* scalpore, allarme, clamore, panico, rumore, sgomento																					

## 5. Verifica e descrizione dei pattern collocazionali recensiti nei repertori sulla base dei risultati dello spoglio del corpus

Aggettivi	Co-occ CORIS		Span								Repertori		
			-4	-3	-2	-1	+1	+2	+3	+4			
abbondanti	min. s.	-											1
	pl.	1							1				
	TOT	<b>1</b>											
agghiacciante	min. s.	1; 2					1	1					DdC+1+c
	pl.	1					1						
	TOT	<b>4</b>											
allarmante	min. s.	1; 11					1	5	5	1			DdC+3
	pl.	5; 1					5		1				
	Ma . s.	-											
	pl.	2						2					
	TOT	<b>20</b>											
allarmistiche	min. s.	-											1
	pl.	8						4	2	1	1		
	Ma . s.	1						1					
	pl.												
	TOT	<b>9</b>											
attendibile	min. s.	1; 5				1			2	2	1 (+5)		DdC+5
<i>n. di borsa: 1 co-occ (span +5)</i>	pl.	10						7	1	3			
<i>1 co-occ (span +5)</i>	TOT	<b>16</b>											
attesa	min. s.	1; 12					1	7	4	1			DdC+1
	pl.	1; 1				1			1				
	Ma . s.	-											
<i>N. più precise per quanto a. : 1 co-occ a (span +5)</i>	pl.	1									1 (+5)		
	TOT	<b>16</b>											
avvilente	TOT	-											DdC
brutta	min. s.	110; 4					110		1	2	1		DdC+5+c
	pl.	86; 4					86		2	1	1 (+5)		
	Ma . s.	-											
	pl.	31; 1					31			1			
	TOT	<b>236</b>											
calda	TOT	-											1
calunniosa	min. s.	2				1	1						1
	pl.	2								2			
	TOT	<b>4</b>											
catastrofica	min. s.	1						1					DdC+2
	pl.	2; 4					2	3		1			
	TOT	<b>7</b>											
circolante (pl)	min. s.	-											1
	pl.	1								1			
	TOT	<b>1</b>											

clamorosa	min. s. 7;	12			7	8	4				DdC + 2
	pl. 1;	2		1		2					
	Ma . s.	-									
	pl. 1;	1		1			1				
	TOT	<b>24</b>									
comunicabile	TOT	-									1
confortante	min. s. 2;	4		2		1	3				DdC + 3
<i>n. e info utili e c.:</i> 1 co-occ (span +5)	pl. 2;	14		2		6	4	3	1 (+5)		
	Ma . s.	3				3					
	pl.	3				2		1			
	TOT	<b>28</b>									
confusa	min. s.	1					1				2
	pl. 3;	12		3		6	2	3	1		
	TOT	<b>16</b>									
consolante	min. s. 2			2							DdC + 2
	pl. 1;	1		1		1					
	TOT	<b>4</b>									
contraddittoria	min. s.	-									DdC + 2
<i>n. confuse e c.:</i> 4 co-occ	pl.	11				4	3	2	2		
	TOT	<b>11</b>									
contrastanti	min. s.	-									1
	pl.	13				10	1	2			
	Ma . s.	-									
	pl.	1				1					
	TOT	<b>14</b>									
credibile	min. s. 1			1							DdC + 2
	pl. 1;	3		1			1	2			
	TOT	<b>5</b>									
cruda	TOT	-									DdC
curiosa	min. s. 3;	12	1	2		8	1	3			DdC
	pl. 2;	8		2		4	2	2			
	Ma . s.	-									
	pl.	1				1					
	TOT	<b>26</b>									
deprimente	min. s.	1				1					DdC
	pl.	2				1	1				
	TOT	<b>3</b>									
dettagliata	min. s. 1			1							DdC
<i>più dettagliata/e</i>	pl. 3;	22	1	2		5	16	1			
	Ma . s.	-									
	pl.	3					3				
	TOT	<b>29</b>									



diffamante	min. s.	1				1					1
	pl.	-									
	TOT	<b>1</b>									
dirompente	TOT	-									DdC
documentata (pl)	min. s.	-									DdC
	pl.	5					3	2			
	TOT	<b>5</b>									
dolorosa	min. s. 3;	3			3	2	1				DdC+2+c
	pl. 1				1						
	TOT	<b>7</b>									
drammatica	min. s. 3;	4			3	4					DdC
	pl. 4;	3			4	2		1			
	TOT	<b>14</b>									
dubbia (pl)	min. s.	-									2
	pl.	3				1		1	1		
	TOT	<b>3</b>									
eclatante	min. s.	2					1	1			1
	pl.	1				1					
	TOT	<b>3</b>									
elementare	TOT	-									2
esatta	min. s.	2					2*				2+c
<i>*o no/se è</i>	pl. 1;	2			1**	1	1				
<i>**più o meno</i>	TOT	<b>5</b>									
esaurienti	min. s.	-									1
	pl. 1				1						
	Ma. s.	-									
	pl. 1				1						
	TOT	<b>2</b>									
falsa	min. s. 37;	33	1		36	16	12	3	2		DdC+5+c
<i>f. e calunniosa n.</i>	pl. 32;	115			2	96	4	6	9		
<i>n. del suo suicidio era f.</i>	Ma. s.	-									
<i>f e contraddittorie</i>	pl. 2;	2			2	2					
<i>f e tendenziose</i>	TOT	<b>221</b>									
fausta	TOT	-									1
faziosa	TOT	-									1
felice (sing.)	min. s. 2				2						DdC
	pl.	-									
	TOT	<b>2</b>									
ferale (sing.)	min. s.	6			6						2
	pl.	-									
	TOT	<b>6</b>									

fondata	min. s. 3	1 (-5)	1	1					DdC + 4	
	pl. 1					1				
	TOT 4									
frammentaria (pl.)	min. s. -								2	
	pl. 3;	13	1		2	9	4			
	TOT 16									
fresca	min. s. 2;	6	1*	1		3	2	1	DdC+5+c	
<i>*è fresca fresca, infatti, la n.</i>	pl. 22					19	3			
	Ma. s. 1					1				
	pl. -									
	TOT 31									
funesta	min. s. 1				1				DdC + 3	
	pl. 3					3				
	TOT 4									
fuorviante (pl)	min. s. -								DdC	
<i>*n. errate/parziali e spesso f. 2 co-occ a (span +5)</i>	pl. 3						1	2*		
	TOT 3									
generica (pl)	min. s. -								DdC	
<i>*n. storiche</i>	pl. 1;	3	1			2		1*		
	TOT 4									
ghiotta	min. s. 1;	5			1	1	2	2	DdC + 1	
	pl. 1;	3			1	1	1	1		
	Ma. s. 1				1					
	pl. -									
	TOT 11									
gradita (sing.)	min. s. 1;	3			1	1	2		DdC + 1	
	pl. -									
	TOT 4									
grave	min. s. 3					1	2		DdC + 2	
	pl. 2;	2		1	1		1	1		
	TOT 7									
importante	min. s. 4;	50		1	3	25	22	2	1	DdC+4+c
<i>n. economica i. più importante</i>	pl. 8;	51			8	24	22	4	2	
	Ma. s. 1					1				
	pl. -									
	TOT 114									
imprecisate	TOT -								2	
impressionante	min. s. 1						1		1	
	pl. 1						1			
	TOT 2									
inaspettata	min. s. 7					2	5		DdC + 4	
	pl. 1					1				

	TOT	<b>8</b>							
inattendibile	min. s.	1		1					DdC + 3
	pl. 1;	2		1	1		1		
	TOT	<b>4</b>							
inattesa	min. s. 4;	7	2		2	2	3	2	DdC + 2
	pl. 2;	3		1	1	3			
	TOT	<b>16</b>							
incerta	min. s. 1				1				DdC + 4
	pl. 2;	5			2	3	1	1	
	Ma. s.	-							
	pl.	2				2			
	TOT	<b>10</b>							
inconsistente	TOT	-							1
incontrollabili	min. s.	1					1		2
	pl.	1				1			
	TOT	<b>2</b>							
incontrollate	min. s.	-							1
	pl.	1				1			
	TOT	<b>1</b>							
incoraggiante	min. s.	7				6	1		DdC + 1
	pl.	8				3	5		
	Ma. s.	-							
	pl.	1				1			
	TOT	<b>16</b>							
incredibile	min. s. 3;	5		1	2	5			DdC + 1
	pl. 1;	1		1		1			
	TOT	<b>10</b>							
inedita	min. s.	4				2	1	1	DdC + 2
	pl.	5				4		1 (+5)	
	TOT	<b>9</b>							
inesatte	min. s.	5				3	1	1	1
	pl. 1;	8		1		6		1	1
	TOT	<b>14</b>							
infamante	min. s.	1				1			1
	pl.	1				1			
	TOT	<b>2</b>							
infausta	min. s.	1			1				2
	pl.	-							
	TOT	<b>1</b>							
infondata	min. s. 4;	6	1		2	1	1	3	1
<i>n. di illecito sportivo : 2 co-occ (span +2)</i>	pl. 2;	11		2		7	3	1	
<i>n. di reato : 5 co-occ (span +2)</i>	Ma. s. 1			1					

	pl.	-												
	TOT	<b>24</b>												
ingiuriosa	TOT	-												1
inquietante	min. s. 1;	3							2		1			DdC
	pl. 1;	4						3	1					
	TOT	<b>9</b>												
interessante	min. s. 1;	17						11	3	2	1			3
	pl. 8;	38	1	1	2	4		27	5	3	3			
	Ma . s.	-												
	pl.	1						1						
	TOT	<b>65</b>												
inverosimili	min. s.	-												1
	pl. 1;									1	1			
	TOT	<b>3</b>												
lieta	min. s. 12;	2							1		1			DdC+3
<i>Lieta/lieta notizia : 4 co-occ a span -1</i>	pl. 3;	2						1			1			
sinonimo di Vangelo	Ma . s.	1												
	pl.	-												
	TOT	<b>19</b>												
luttuosa	TOT	-												
magnifica	min. s.	2												1
	pl.	-												
	TOT	<b>2</b>												
malsicura	TOT	-												1
meravigliosa	min. s. 3;	6						5	1					1
	pl. 1;	1						1						
	TOT	<b>11</b>												
nefasta	min. s. 1													DdC+1
	pl.	-												
	TOT	<b>1</b>												
negativa	min. s.	2						2						DdC+c
	pl.	17						14	3					
	TOT	<b>19</b>												
parziale	min. s.	1						1						DdC
	pl. 1;	5						4						
	TOT	<b>7</b>												
piacevole	min. s. 1													DdC+3
	pl.	3						2	1					
	TOT	<b>4</b>												
positiva	min. s.	29						21	3	4	1			DdC
	pl. 1;	24						17	3	1	3			

	Ma . s.	2				1	1			
	pl.	4				3	1			
	TOT	<b>60</b>								
precisa	min. s.	8				7	1			2
<i>n. più precisa/e</i>	pl. 1;	46			1	28	15	2	1	
	TOT	<b>55</b>								
preoccupante	min. s.	5				2	1	1	1	DdC+2
	pl. 2;	15			2	9	4	1	1(+5)	
	TOT	<b>22</b>								
rasserenante	TOT	-								1
rassicurante	min. s.	5				3	1	1		DdC+2
	pl. 4;	13			2 2	5	4	3	1(+5)	
	TOT	<b>22</b>								
recente	min. s. 15;	9			6 9	8	1			DdC+7
<i>r. fughe di n.</i>	pl. 11;	18			1 10	5	3			
	TOT	<b>53</b>								
rilevante	min. s.	5				1	3		1	1
<i>n. scientifiche più r.</i>	pl.	15				10	1	4		
	TOT	<b>20</b>								
rincuorante	TOT	-								DdC
sbalorditiva	min. s. 1;	3			1	3				1
	pl.	-								
	TOT	<b>4</b>								
scarna (pl)	min. s.	-								DdC+1
	pl. 11;	2			2 9	2				
	Ma . s.	-								
	pl. 1				1					
	TOT	<b>14</b>								
n. scarse	min. pl 9;	3	2		7	1		1	1	DdC+6
	Ma. pl 1				1					
	TOT	<b>13</b>								
scioccante (pl)	min. s.	-								1
	pl.	2				2				
	TOT	<b>2</b>								
sconcertante	min. s. 2;	4			2	3		1		DdC+1
	pl.	4				4				
	TOT	<b>10</b>								
sconfortante	min. s.	2				2				DdC
	pl.	3				3				
	TOT	<b>5</b>								
sconvolgente	min. s. 5;	12			1 4	8	1	1	2	DdC+2
	pl. 2;	3			2	3				

	TOT	22							
scottante	min. s.	-							DdC + 2
	pl.	1			1				
	TOT	1							
sensazionale	min. s. 2;	16	2	11	3	1	1		DdC + 5
	pl. 2;	7	2	6	1				
	TOT	27							
sgradita	min. s.	1		1					1
	pl.	1		1					
	TOT	2							
sicura	min. s.	5		5					3
	pl. 1;	10	1	5	2	3			
	Ma. s.	-							
	pl.	1			1				
	TOT	17							
significativa	min. s. 1;	2	1		1	1			DdC
	pl.	-							
	Ma. s.	-							
<i>Buone n. tanto più significative</i>	pl.	1				1			
	TOT	4							
singolare	min. s. 1;	2	1	1	1				DdC
	pl.	1		1					
	TOT	4							
sorprendente	min. s. 2		2						DdC + 1
	pl.	2		2					
	TOT	4							
specifica (pl)	min. s.	-							DdC
	pl. 2;	4	2	4					
	TOT	6							
spiacevole	min. s. 1		1						DdC + 4
	pl.	3		3					
	TOT	4							
splendida	min. s. 9;	2	9	2					DdC + 1
	pl.	-							
<i>n. dal campo norvegese erano s.</i>	Ma. s.	-					1 (+5)		
<i>* N. splendide, non buone</i>	pl.	2		1*					
	TOT	13							
stantia	TOT	-							3
strabilante	min. s. 2;	1	2		1				1
	TOT	3							
stringata	TOT	-							DdC

stupefacente	min. s. 1;	1			1	1													2
	pl.	1					1												
	TOT	<b>3</b>																	
tendenziosa (pl)	min. s.	-																	DdC+6
<i>*false,(esagerate),t.</i>	pl. 1;	15			1	2		9*	4*										
	TOT	<b>16</b>																	
terribile	min. s. 15;	14			15	9	5												DdC+3
<i>n. del genocidio erano talmente t. +5</i>	pl. 1;	4			1	3			1										
	TOT	<b>34</b>																	
tragica	min. s. 19;	4			19	1	3												DdC+2
	pl. 2;	1			2		1												
	Ma. s.	-																	
	pl.	1				1													
	TOT	<b>27</b>																	
tranquillizzanti	min. s. 1				1														2
<i>n. da Bruxelles non erano t. : span +5</i>	pl.	4				1	1	1	1(+5)										
	Ma. s.	-																	
	pl.	1				1													
	TOT	<b>6</b>																	
tremenda	min. s. 3;	2			3	1	1												DdC+1+c
	pl.	1				1													
	TOT	<b>6</b>																	
triste	min. s. 14;	5			14	1	3		1										DdC+2+c
	pl. 3;	6			3	3	2	1											
	TOT	<b>28</b>																	
utile	min. s. 3;	14	1		2	10	1	3											DdC+1
<i>* n. e consigli/info/indirizzi/ecc. u. : 9 co-occ</i>	pl. 7;	102	1		6	83	6	11*	2**										
<i>** n. che si ritengono/possono ES u.</i>	Ma. s.	1					1												
	pl.	1				1													
	TOT	<b>128</b>																	
vaga	min. s. 3;	1			3			1											2
	pl. 2;	4			2	1	1	2											
	TOT	<b>10</b>																	
vecchia	min. s. 2;	5		1	1	3	2												DdC+3
<i>n. già vecchie</i>	pl.	5				1	2	1	1										
	TOT	<b>12</b>																	
vera	min. s. 33;	33	1		12	21	13	13	6	1									DdC+5+c
<i>*vere o truccate immagini e n. : span -5</i>	pl. 12;	36	1*	2	4	5	21	10	2	3									
<i>vere e proprie n. : 2 co-occ (span -3)</i>	Ma. s. 1;	2	1				1	1											
<i>n. non vera/e : molte co-occ span +1</i>	pl.	1					1												
	TOT	<b>118</b>																	
veritiera	min. s.	1								1									DdC
	pl.	3						1	2										
	TOT	<b>4</b>																	

Sequenze obbligate	Co-occ CORIS	Span								Repertori
		-4	-3	-2	-1	+1	+2	+3	+4	
bella n.	min. s. 125; 1 pl. 20 Ma. s. 1 pl. 2 TOT <b>149</b>		1	1	123		1			DdC + 7
			1		19	1*	1*			
			1**							
					2					
buona n.	min. s. 599; 31 pl. 269; 35 Ma. s. 5 pl. 127; 1 TOT <b>1067</b>			2	597	22	3	5	1	DdC + 8
* 1 co-occ span (+5): n. non ES quasi mai b.			3		266	8	13	10	4*	
					5					
			2		125	1				
cattiva n.	min. s. 95; 14 pl. 123; 6 Ma. s. - pl. 11 TOT <b>249</b>				95	2	4	1	7	DdC + 5
					123	2		3	1	
						11				
gioiosa n.	min. s. <b>1</b>				1					1
ottima n.	min. s. 28 pl. 4 Ma. s. 5 pl. 5 TOT <b>42</b>			1	27					1
						1				
ultima n.	min. s. 50 pl. 188 Ma. s. 1 pl. 15 TOT <b>254</b>	1		2*	47					DdC + 12
*ultima clamorosa/raccapricciante n.					179					
prime notizie	min pl <b>92</b>			4*	88					1
*prime frammentarie/confuse/parziali n										
n. accreditata	min. s. <b>2</b>					1		1		1
n. angosciosa	min. s. 1 pl. 1 TOT <b>2</b>						1			1
						1				
n. anticipata (diatesi pax)	min. s. 8 pl. 3 TOT <b>11</b>					8				1
						3				
notizia bomba (inv)	min. s. 29 pl. 3 TOT <b>32</b>					29				DdC + 3
						3				
n. certa	min. s. 15 pl. 29 Ma. s. - pl. 1					12	2	1		DdC + 8
n. più certe: co-occ span +2						24	5			
						1				



	TOT	<b>45</b>							
n. confidenziale	min. s.	1							1
	pl.	3							3
	TOT	<b>4</b>							
n. culturale	min. s.	1							DdC + 1
	pl.	2							2
	TOT	<b>3</b>							
n. economica	min. s.	1							DdC + 3
* n. (geografiche), politiche ed eco.	pl.	15							12 2 1*
	TOT	<b>16</b>							
n. esclusiva	TOT	-							DdC
n. esplosiva	min. s.	1							1
	pl.	2							2
	TOT	<b>3</b>							
n. flash	min. s.	2							1
	pl.	6							6
	TOT	<b>8</b>							
n. fritta/rifritta	TOT	-							1
n. improvvisa	min. s.	5							c
all'i. n. dell'elevazione : co-occ span (-1)	pl. 1;	1			1				1
	TOT	<b>7</b>							
n .inventata	min. s.	1							DdC + 2
	pl.	6							5 1
	TOT	<b>7</b>							
n. lampo	min. s.	-							3
	pl.	1							1
	TOT	<b>1</b>							
n. pesante	min. pl	<b>1</b>							1
n. politica	min. s.	1							DdC + 3
	pl.	15							12 2 1*
	TOT	<b>16</b>							
n. riservata	min. s.	8							DdC + 3
	pl.	54							4 1 2 1 47 5 1 1
	Ma . s.	-							
	pl.	1							1
	TOT	<b>63</b>							
n. scandalistica	TOT	-							DdC + 1
n. scontata	min. s.	2							1
	pl.	-							2
	TOT	<b>2</b>							



Verbo + Complemento	Co-occ CORIS	T. sempli- ci	T. compo- sti	N + Part. Passato concor- dato	Part. Passato concor- dato + N	Repertori
abbondare di notizie	-					DdC
accertare	5	1		3	1	DdC + 2
accogliere	65	10	41	2	12	DdC+1+c
amplificare	1	1				DdC + 1
annunciare	22	8	2	1	11	DdC + 2
(pre)annunziare	-					2
apprendere	173	42	73	26	32	DdC+7+c
appurare	-					2
ascoltare	43	35	8			5
aspettare	52	40	12			DdC + 3
assumere	9	4		5		2
attendere	35	18	2	13	2	DdC + 2
attingere (alle) notizie	22	11	1	10		4
(non) avere n.	686	355	259	8	64	DdC + 10
avvalorare	-					DdC
avvertire con	-					1
avvisare con	-					1
cantare	-					1
censurare	7	5	2			DdC + 10
chiedere	274	214	56	1	3	DdC+5+c
commentare	104	80	20	4		DdC+1+c
comunicare	102	65	16	19	2	DdC+7+c
confermare	72	26	9	37		DdC + 8
congetturare	-					2
conoscere	9	6		3		c
controllare	11	7		4		DdC + 1
dare	586	248	152	102	84	DdC + 9
dare rilievo a	2	2				c
dare risalto a	12	5	7			DdC
diffondere	280	180	40	58	2	DdC + 8
diramare	13	3	2	7	1	DdC + 3
divulgare	73	47	8	13	5	DdC + 6
domandare	10	9	1			2
essere confortato/ecc. da	34					DdC + 1
essere senza	3	1	2			1
falsificare	-					1
fare notizia	191	167	24			DdC + 11
fornire	163	101	9	47	6	DdC + 2
gonfiare	5	2		3		DdC + 3
ingigantire	2			2		DdC + 1
informare con	2	2				1
intercettare	2		2			1
inventare	6	4	2			1+c
irradiare	-					1
lasciare senza	1		1			1
leggere	158	99	50	9		DdC + 6
mandare (proprie) notizie	21	17	4			3
manipolare	7	3		4		2
mettere in circolazione	1		1			DdC + 1
mettere in giro	4	2	1	1		DdC + 1



coincidono	<b>1</b>	1				c
colpisce	<b>14</b>	14				c
conferma	<b>9</b>	9				c
corre/correva	<b>17</b>	17				DdC + 7
crea scalpore/ecc.	<b>53</b>	36	17			DdC + c
danneggia	-					DdC
desta scalpore/ecc.	<b>40</b>	16	24			c
è apparsa (sul giornale)	<b>49</b>	16	33			DdC + 1
è uscita (sul giornale)	<b>28</b>	8	24			1
fa (molto) scalpore/ecc.	<b>764</b>	695	69			2
filtra (costr.causativa: 5 co-occ <i>far</i> e 1 <i>lasciar</i> )	<b>49</b>	39	10			1
giunge (nuova)	<b>329</b>	208	121			DdC + 6
ha fatto il giro (del mondo)	<b>48</b>		48			1
ha scioccato	<b>1</b>		1			1
informa	<b>1</b>	1				3
interessa	<b>16</b>	15	1			1
lascia senza <i>fiato</i> / <i>parole</i>	<b>3</b>	1	2			DdC
mancano notizie (su)	<b>15</b>	15				2
perviene	<b>18</b>	9	9			2
riempie di <i>gioia/angoscia/ansia/ecc.</i>	<b>5</b>	4	1			c
rimbalza (in)	<b>56</b>	28	28			DdC + 2
scarseggiano	-					1
sconvolge	<b>16</b>	7	9			DdC + c
scotta	-					1
scuote	<b>10</b>	3	7			DdC + 1
si diffonde	<b>48</b>	4	44			2
si rivela	<b>3</b>	2	1			c
si sparge	<b>59</b>	27	32			1
sorprende	<b>12</b>	12	4			DdC
suscita	<b>61</b>	40	21			1
si susseguono	<b>7</b>		7			2
si incalzano	-					1
tranquillizza	<b>5</b>	4	1			DdC
trapela (costr.causativa: 13 co-occ <i>far</i> e 8 <i>lasciar</i> )	<b>73</b>	37	36			DdC + 8
vola	-					1
<b>N.B.</b> crea, desta,fa,suscita: allarme, clamore, panico, rumore, scalpore, sgomento						

**6. Ampliamento e descrizione dei pattern collocazionali non recensiti nei repertori e ricavati dallo spoglio del corpus**

Aggettivi	Co-occ CORIS		Span								Occorrenze CORIS	
			-4	-3	-2	-1	+1	+2	+3	+4		
prive	min. s.	2; 9	2					2	3	1	3	4.561
di fondamento/interesse/importanza	pl.	3; 11	3					5	1	2	3	
	TOT	25										
Sequenze obbligate	Co-occ CORIS		Span								Occorrenze CORIS	
			-4	-3	-2	-1	+1	+2	+3	+4		
n. scomode	min. s.	2					1	1			753	
	pl.	10					9			1		
	Ma. s.	-										
	pl.	1					1					
	TOT	13										

Verbo + Complemento	Co-occ CORIS	T. semplici	T. composti	N + Part. Passato concordato	Part. Passato concordato + N	Occ Tot CORIS	Span								Occ Tot CORIS
							-4	-3	-2	-1	+1	+2	+3	+4	
reperire	12	12				1.140									
scambiarsi	19	18		1		5.783									
	Co-occ CORIS	Span								Occ Tot CORIS					
		-4	-3	-2	-1	+1	+2	+3	+4						
n. coperte	min. s.	2							2	2.268					
da segreto professionale/istruttorio	pl.	13				12			1						
	TOT	15													
Soggetto + Verbo	Co-occ CORIS	T. semplici	T. composti			Occ Tot CORIS									
compare	20	7	13			8.133									
proviene	21	21				5.074									
raggiunge	23	16	7			34.775									
risale	19	19				8.670									
	Co-occ CORIS	Span								Occ Tot CORIS					
		-4	-3	-2	-1	+1	+2	+3	+4						
n. concernenti	min. s.	2				2				1.249					
	pl.	11				7		2	2						
	TOT	13													
N.B. CONCERNERE 13 co-occorrenze totali e tuttea Partecipio Presente															
n. riguardante	min. s.	12				9	2*		1(+5)	1.032					
*n. diffusa/magari r.	pl.	61				51	1	6	3						
(+5) n. appresa dalla stampa e r.	TOT	73													
N.B. RIGUARDARE 183 co-occorrenze totali: solo tempi semplici, di cui 73 al Partecipio Presente															

## 7. Classifiche dei collocati di *notizia* ottenute dal calcolo delle misure di associazione

Query: "notizia notizie"		Occorrenze tot.	20414
Collocate	Mutual Information	f (node, collocate)	f (collocate)
Celebrità	12.153993	13	23
trapelò	11.714081	10	24
Cattive	11.562078	12	32
attinte	11.299043	10	32
saputa	10.022919	12	93
sensazionali	9.699130	10	97
filtrano	9.416400	10	118
Brutte	9.317640	31	62
rimbalzano	9.089590	10	148
trapela	9.044229	11	168
riportavano	9.013641	10	156
Verissimo	8.962165	12	194
stillicidio	8.872778	10	172
trapelata	8.834501	30	69
scomode	8.441062	10	232
rimbalza	8.174922	10	279
curiose	8.122966	13	376
ricevo	8.009024	10	313
commentando	7.929082	12	397
riportava	7.866983	13	449
apparse	7.827368	12	426
avute	7.764416	12	445
contraddittorie	7.640832	10	404
diffondendo	7.584798	12	504
inattesa	7.556453	12	514
allarmante	7.514954	12	529
filtrare	7.488472	10	449
attendibili	7.369785	12	585
scambiarsi	7.330638	11	551
apprese	7.218194	33	374
Striscia	7.191591	195	681
sparsa	7.100649	32	347
attendibilità	7.053491	10	607
meteo	7.013641	10	624
circola	6.997547	10	631
scarse	6.980877	12	766
Arrivano	6.941491	10	656
smentisce	6.940692	11	722
contrastanti	6.892920	13	882
diffondersi	6.865110	12	830
risalgono	6.825806	13	924
trasmesse	6.785495	11	804
approfondimenti	6.785316	10	731
reperire	6.732643	11	834
bollettino	6.724450	12	915
darne	6.717537	59	522
sicure	6.690172	12	937
insieme alle	6.686175	10	783
decesso	6.626618	10	816
fornendo	6.558614	13	1112
Collocate	Log-Likelihood Ratio	f (node, collocate)	f (collocate)
buona	5885.625000	651	21769
buone	3303.406250	317	5343
dare	2438.000000	343	29839
Buone	1930.945312	131	314
dato	1874.062500	287	33362
avere	1787.187500	314	55871
avuto	1730.250000	275	36043
ultime	1677.953125	203	9863
reato	1669.625000	193	7742
false	1631.902344	150	2017
cattive	1591.736328	134	1119
un	1584.000000	1040	1818792
giornali	1529.828125	188	9734

Query: "notizia"		Occorrenze tot.	11.058
Collocate	Mutual Information	f (node, collocate)	f (collocate)
saputa	10.907387	12	93
trapela	9.791194	10	168
trapelata	9.718969	30	69
rimbalza	9.059390	10	279
darvi	8.943638	13	393
comunicò	8.511086	10	408
riportava	8.510444	11	449
allarmante	8.399422	12	529
telegiornali	8.382612	10	446
inattesa	8.315391	11	514
smentire	8.276621	13	624
Striscia	8.076060	195	681
tg	8.072376	10	553
sparsa	7.985117	32	347
circola	7.882015	10	631
decesso	7.511086	10	816
choc	7.154224	10	1045
uccisione	7.136140	11	1164
darne	7.100668	57	522
accolta	6.974871	12	1420
ebbi	6.973855	12	1421
sconvolto	6.932029	10	1219
annunciata	6.836998	10	1302
cattura	6.787236	13	1752
smentita	6.783824	33	933
riguardante	6.711836	11	1562
Ansa	6.655058	10	1477
imminente	6.652130	13	1924
fornita	6.631803	10	1501
ricevuta	6.623179	12	1812
confermare	6.547189	12	1910
nascosta	6.490561	13	2152
splendida	6.455024	12	2036
darle	6.417326	34	1377
avuta	6.405910	59	1196
raggiunse	6.367546	11	1983
averne	6.319842	12	2236
diffondere	6.285138	35	1700
Times	6.228868	13	2580
indomani	6.225517	13	2586
risale	6.197889	13	2636
comparsa	6.130265	10	2125
fondamento	6.121154	13	2780
darà	6.098703	10	2172
matinata	6.069119	10	2217
apparsa	5.944416	31	1186
riceve	5.938599	13	3155
agenzie	5.869589	12	3055
commenti	5.862053	12	3071
sollievo	5.835689	11	2867
Collocate	Log-Likelihood Ratio	f (node, collocate)	f (collocate)
buona	6.575.906.250	641	21769
dare	1.808.687.500	238	29839
stata	1.769.375.000	299	87924
dato	1.718.312.500	234	33362
morte	1.573.312.500	219	33815
arriva	1.425.906.250	171	14761
diffusa	1.244.570.312	124	4590
brutta	1.231.765.625	115	3052
cattiva	1.208.070.312	112	2864
fa	1.120.000.000	227	110131
ieri	1.081.750.000	168	37698
data	1.055.812.500	154	27963
avuto	1.045.750.000	162	36043
bella	1.027.437.500	138	18614
reato	993.953.125	113	7742
giornali	768.984.375	96	9734
Breve	441.585.938	23	298
scelte	159.046.875	14	8686



Query: "notizie"		Occorrenze tot.	9356
Collocate	Mutual Information	f (node, collocate)	f (collocate)
Celebrità	13.279.588	13	23
Cattive	12.687.673	12	32
attinte	12.424.638	10	32
sensazionali	10.824.726	10	97
filtrano	10.541.995	10	118
Brutte	10.443.235	31	62
rimbalzano	10.215.185	10	148
scomode	9.566.657	10	232
curiose	9.133.084	12	376
apparse	8.952.963	12	426
avute	8.890.011	12	445
contraddittorie	8.766.427	10	404
telegiornali	8.623.738	10	446
attendibili	8.495.380	12	585
scambiarsi	8.456.233	11	551
meteo	8.139.236	10	624
scarse	8.106.472	12	766
Arrivano	8.067.086	10	656
contrastanti	8.018.515	13	882
trasmesse	7.911.090	11	804
approfondimenti	7.910.911	10	731
reperire	7.858.238	11	834
risalgono	7.835.924	12	924
riportano	7.759.302	10	812
bollettino	7.724.514	11	915
sicure	7.690.237	11	937
formendo	7.684.209	13	1112
brutte	7.682.039	90	1004
diffonde	7.653.149	10	874
di fronte alle	7.579.148	11	1012
fughe	7.459.644	58	581
ricevute	7.440.110	13	1317
giungono	7.410.409	60	808
provengono	7.377.934	13	1375
prive	7.368.522	13	1384
tristi	7.281.255	10	1131
news	7.269.820	11	1254
chiedergli	7.249.713	10	1156
ricevono	7.093.721	10	1288
interessano	7.016.212	11	1495
darmi	6.968.751	11	1545
riservate	6.901.202	60	1150
illecito	6.866.914	11	1658
diffuse	6.732.419	38	1970
coperte	6.602.251	13	2354
concernenti	6.576.300	11	2028
gli	6.425.465	167	461197
raccolte	6.346.180	39	2791
quotidiani	6.325.392	13	2852
Collocate	Log-Likelihood Ratio	f (node, collocate)	f (collocate)
buone	3.775.125.000	315	5343
Buone	2.117.437.500	130	314
ultime	1.936.625.000	198	9863
false	1.867.027.344	150	2017
avere	1.820.812.500	264	55871
sulla	1.804.000.000	330	144912
sono	1.745.500.000	468	510416
fuga	1.576.343.750	156	6603
arrivano	1.264.203.125	128	6058
informazioni	1.168.812.500	161	28222
brutte	1.154.687.500	90	1004
utili	1.058.703.125	116	7965
hanno	970.500.000	227	183580
prime	777.687.500	102	14839
giornali	759.796.875	92	9734
fughe	757.041.992	58	581
giungono	746.046.875	60	808
diffusione	741.093.750	86	7602



## Bibliografia

### **Dizionari**

- DAU (2010): *Grande dizionario analogico della lingua italiana*, progettato e diretto da Simone R., Torino, UTET;
- DCI (2013): *Dizionario Combinatorio Italiano*, a cura di Lo Cascio V., John Benjamins Publishing Company;
- DCL (2009): *Dizionario delle Combinazioni Lessicali*, Urzì F., Convivium;
- DdC (2018): *Dizionario delle collocazioni. Le combinazioni delle parole in italiano*, Tiberii P., Bologna, Zanichelli;
- DISC (2013): *Il Sabatini-Coletti, Dizionario della Lingua italiana*, Milano, Rizzoli;
- DO (2018): *Nuovo Devoto-Oli: il vocabolario dell'italiano contemporaneo*, ideato da Devoto G., Oli G.C. e a cura di Serianni L., Trifone M., Firenze, Le Monnier;
- GDIH (2015): *Grande dizionario Hoepli italiano*, A. Gabrielli. Milano, Hoepli;
- GraDIt (2000): *Grande dizionario italiano dell'uso*, ideato da De Mauro T., Torino, UTET;
- MdD (2010): *MdD. Modi di Dire. Lessico Italiano delle Collocazioni*, Russo D., Roma, Aracne;
- NTr (2018): *Il nuovo Treccani. #leparolevalgono*, Roma, Istituto della Enciclopedia Italiana;
- PF (1992): *Dizionario della lingua italiana*, Palazzi F., Folena G., Loescher;
- SCAN (1993): *Dizionario dei sinonimi e dei contrari analogico e nomenclatore*, Gabrielli A., Loescher;
- ThTr (2018): *Thesaurus, Dizionario analogico della lingua italiana*, in Il vocabolario Treccani. Roma, Istituto della Enciclopedia Italiana;
- Zin (2022): *lo Zingarelli 2023, Vocabolario della lingua italiana (versione base)*, di N. Zingarelli, a cura di Cannella M, Lazzarini B, Zaninello A., Bologna, Zanichelli.

### **Sitografia dizionari**

<https://www.treccani.it/vocabolario/>

### **Banche dati**

**ELDIT (dal 2002): Elektronisches Lernerwörterbuch Deutsch-Italienisch, Accademia Europea di Bolzano/SUPSI/Università degli Studi di Trento;**  
<http://eldit.eurac.edu/>

### **Corpus**

Università di Bologna (dal 2001): <https://corpora.ficlit.unibo.it/TCORIS/>

### **Articoli e monografie**

Arduini A. - Fabbri R. (2008): *Che cos'è la linguistica cognitiva*. Roma, Carrocci Editore, collana "Le bussole";

- Benigno V. - Kraif O. (2016): Core vocabulary and core collocations: combining corpus analysis and native speaker judgement to inform selection of collocations in learner dictionaries, in *Defining collocation for lexicographic purposes. From linguistic theory to lexicographic practice*, Orlandi A.- Giacomini L. (eds.). Bern, Peter Lang AG;
- Benson, M. - Benson, E. - Ilson, R. (1986): *The BBI Combinatory Dictionary of English. A guide to word combinations*; Benson, M.-Benson, E.-Ilson, R. (eds.). Amsterdam/Philadelphia, John Benjamins Publishing Co.;
- Casadei, F. (2003): *Lessico e Semantica*. Roma, Carrocci Editore, collana “Le basi della linguistica”;
- Casadei, F. - Basile, G. (2019): *Lessico ed educazione linguistica*. Roma, Carrocci Editore, collana “Studi superiori”;
- Dardano M. - Trifone P. (1997): *La nuova grammatica della lingua italiana*: Bologna, Zanichelli;
- De Mauro T. (2000): *Introduzione a Grande dizionario italiano dell’uso*, ideato da De Mauro T., Torino, UTET;
- Denes, G. (2009): *Parlare con la testa. Le basi neurologiche e la struttura del linguaggio*. Bologna, Zanichelli;
- Giacomini, L. (2012): *Un dizionario delle collocazioni come rete di relazioni lessicali. Studio sul campo semantico della paura*. Frankfurt, Peter Lang GmbH;
- Granger, S. - Paquot, M. (2008): Disentangling the phraseological web, in *Phraseology, An interdisciplinary perspective*, Granger S. - Meunier F. (eds.). Amsterdam/Philadelphia, John Benjamins Publishing Co.;
- Gries S.T. (2008): Phraseology and linguistic theory: A brief survey, in *Phraseology, An interdisciplinary perspective*, Granger S. - Meunier F. (eds.). Amsterdam/Philadelphia, John Benjamins Publishing Co.;
- Firth, J.R. (1957): *Papers in Linguistics 1934-1951*. London, Oxford University Press;
- Halliday, M.A.K. - Hasan, R. (1976): *Cohesion in English*. London and New York, Longman;
- Hoey, M. (2005): *Lexical Priming. A new theory of words and language*. London, Routledge;
- Ježek, E. (2016): *The lexicon, An introduction*. Oxford, Oxford University Press;
- Jones, S. - Sinclair, J.M. (1974): English Lexical Collocations. A Study in computational linguistics, in *Cahiers de lexicologie* 24, pp. 15-61;
- Legallois, D. - Tutin, A. (2013): Vers une extension du domaine de la phraséologie, *Langages*, 2013 n. 189/1, pp. 3-25;
- Lenci A.-Montemagni S.-Pirrelli V. (2005): *Testo e Computer: elementi di linguistica computazionale*, Roma, Carrocci editore;
- López Ferrero, C. - Battaner, P. (2017): Learning Spanish L1 vocabulary in context, in *Collocations and Other Lexical Combinations in Spanish. Theoretical, Lexicographical and Applied Perspectives*; Torner, S - Bernal E. (eds). London and New York, Routledge;
- Mel’čuk I. (1998): Collocations and Lexical Function, in *Phraseology. Theory, Analysis and Applications*, Cowie A.P..Oxford, Clarendon Press;

- Mel'čuk I. (2003): Collocations: définition, rôle et utilité, in *Les Collocations analyse et traitement*. Amsterdam, De Werelt, collana "Travaux et recherches en linguistique appliquée";
- Moon R. (2012): Dictionaries and collocation, in *Phraseology, An interdisciplinary perspective*, Granger S. - Meunier F. (eds.). Amsterdam/Philadelphia, John Benjamins Publishing Co.;
- Nesselhauf, N. (2005): *Collocations in a Learner Corpus*. Amsterdam/Philadelphia, John Benjamins Publishing Co., collana "Studies in Corpus Linguistics"
- Orlandi, A. (2016): Monolingual collocation lexicography: State of art and new perspectives, in *Defining collocation for lexicographic purposes*, Orlandi, A - Giacomini, L. (eds). Bern, Peter Lang AG;
- Marello C. (1996): *Le parole dell'italiano: Lessico e dizionari*. Bologna, Zanichelli;
- Palmer, F.R. (1968): *Selected Papers of J.R. Firth 1952-59*. London - New York, Longmans;
- Partington, A. (1998): *Patterns and Meaning, Using Corpora for English Language Research and Teaching*. Amsterdam/Philadelphia, John Benjamins Publishing Co., collana "Studies in Corpus Linguistics";
- Pawley, A. - Syder, F.H. (1983): Two puzzles for linguistic theory: nativelike selection and nativelike fluency, in *Language and communication*, Richards J.C. & Schmidt R.W. (eds.). London and New York, Longman;
- Rossini Favretti R.-Tamburini F-De Santis C. (2002): CORIS/CODIS: A corpus of written Italian based on a defined and a dynamic model, in *A Rainbow of Corpora: Corpus Linguistics and the Languages of the World*, Wilson A., Rayson P. and McEnery, T. (eds.), Lincom-Europa, Munich. (2002), 27-38;
- Sinclair, J.M. (1991): *Corpus, Concordance, Collocation*. Oxford, Oxford University Press;
- Sinclair, J.M. (2008): The phrase, the whole phrase, and nothing but the phrase, in *Phraseology, An interdisciplinary perspective*, Granger S. - Meunier F. (eds.). Amsterdam/Philadelphia, John Benjamins Publishing Co.;
- Squillante, L. (2016): *Polirematiche e collocazioni dell'Italiano. Uno studio linguistico e computazionale*. Tesi di dottorato. Hildesheim, Universitätsverlag Hildesheim;
- Sosa A.V. - MacFarlane J. (2002): Evidence for frequency-based constituents in the mental lexicon: collocations involving the word of, in *Brain and Language* 83 (2002), pp. 227-336;
- Torner S. – Bernal E. (2017): *Collocations and Other Lexical Combinations in Spanish: Theoretical, Lexicographical and Applied Perspectives*, New York, Routledge;

### **Sitografia articoli e monografie**

Simone R. (2010): *Causativa, costruzione:*

[https://www.treccani.it/enciclopedia/costruzione-causativa\\_%28Enciclopedia-dell%27Italiano%29/](https://www.treccani.it/enciclopedia/costruzione-causativa_%28Enciclopedia-dell%27Italiano%29/)