



Laura Giacomini

Un dizionario elettronico
delle collocazioni come
rete di relazioni lessicali

Studio sul campo
semantico della paura



Introduzione

Kennen Sie eigentlich Momo? Momo lernt Deutsch als Fremdsprache. Er kann schon alle Sprachen der Erde, nur Deutsch lernt er gerade. Heute morgen ist er aufgestanden und hat sein Bett *gebaut*. Nein, nicht gebaut, gebaut ist militärisch, er hat sein Bett *gemacht*. Dann hat er sich die Zähne *gewaschen*. Nein, nicht gewaschen, gewaschen ist Französisch (se laver les dents), er hat sich die Zähne *geputzt*. Dann hat er eine Dusche gemacht. Nein, nicht gemacht, machen ist Italienisch (fare la doccia). Er ist unter die Dusche *gegangen* oder hat geduscht. Kann man sagen: er hat eine Dusche *genommen*? Momo wird sich erkundigen (Hausmann 2003: 309).

Le collocazioni sono uno dei fenomeni linguistici forse meno evidenti all'occhio del parlante madrelingua eppure tra i più imprescindibili per l'atto comunicativo. Un fenomeno che tocca trasversalmente tutte le dimensioni di variazione di una lingua. Dalla svolta firthiana negli anni '50 si è assistito al proliferare di studi linguistici che hanno scandagliato le collocazioni da diverse prospettive, quali la semantica lessicale, la sintassi, l'estrazione da corpora elettronici. Nondimeno, se le ricerche lessicologiche hanno contribuito ad approfondire la conoscenza di questo fenomeno moltiplicando al contempo gli approcci scientifici attraverso cui affrontarlo, nel panorama lessicografico italiano si assiste in genere ad una applicazione timida e circoscritta delle conoscenze raggiunte, già visibile nell'evidente mancanza di un riconoscimento dello status di collocazione nei metatesti quali le avvertenze per la consultazione. Negli ultimi anni si è avuto di fatto un lento miglioramento nella rappresentazione delle combinazioni lessicali all'interno dei dizionari semasiologici di lingua comune, siano essi monolingui o bilingui. Tuttavia la descrizione delle combinazioni di tipo collocativo è a livello macrostrutturale, microstrutturale e mediostrutturale ancora molto lontana da risultati soddisfacenti, mancando spesso di fondamenti basilari quali la coerenza dei dati e l'attenzione alle necessità dell'utente.

La ricerca qui esposta ha come fine l'elaborazione di un modello lessicografico per un dizionario elettronico delle collocazioni italiane. La lingua considerata è quella comune, anche se le collocazioni di tipo specialistico non vengono escluse dalla trattazione. Il dizionario speciale ha funzione complementare rispetto al dizionario generale monolingue o bilingue, di cui va ad integrare i dati. I presupposti di partenza, che concorrono a determinare l'architettura del dizionario, sono da un lato l'utilità di individuare delle reti semantiche tra parole, sperimentata sulla base del lessema *paura* e di sostantivi ad esso semanticamente affini, e dall'altro lato la constatazione che le collocazioni sono sì un fenomeno particolarmente significativo dal punto di vista semantico, ma che esse dovrebbero anche essere inquadrare in precisi schemi sintattici che ne facilitino il reperimento da parte dell'utente. La ricerca nasce dalla considerazione che per la lin-

gua italiana è a tutt'oggi assente un dizionario di questo tipo, a cui ascriviamo a priori un notevole interesse per un'utenza che includa traduttori professionisti così come apprendenti avanzati della lingua italiana.

L'esposizione delle varie fasi della ricerca si snoda attraverso cinque capitoli, che ne forniscono una suddivisione sia cronologica che tematica. Il capitolo 1 "Le teorie sulle emozioni tra scienze cognitive e semantica lessicale" costituisce la parte enciclopedica del lavoro, e presenta una panoramica delle maggiori teorie cognitive sulla natura delle emozioni e sul linguaggio a queste connesso, descrivendo al contempo il procedimento di selezione dei sostantivi ai quali sono stati successivamente applicati l'analisi lessicologica ed il modello lessicografico. Il capitolo 2 "Corpus, cooccorrenze ed analisi statistica" descrive l'approccio metodologico in rapporto al corpus utilizzato, la definizione di collocazione approntata ai fini specifici della ricerca, l'estrazione automatica dei dati e l'iniziale valutazione quantitativa e propriamente statistica delle collocazioni candidate. La fase successiva, descritta al capitolo 3 "L'analisi prelessicografica: dalla descrizione sintagmatica alla selezione delle collocazioni", si apre con una analisi delle strutture sintagmatiche in cui i sostantivi studiati occorrono, per poi proseguire con la prima categorizzazione sintattica e semantica delle collocazioni, selezionate attraverso una imprescindibile procedura di valutazione non automatizzata.

Conclusa la parte prelessicografica della ricerca, il capitolo 4 "Modellazione di una banca dati contenente le collocazioni italiane" è incentrato sull'inserimento del materiale collocativo in una banca dati lessicale da cui il dizionario possa attingere automaticamente i lemmi, le collocazioni ed il sistema di rinvii sintattici e semantici che connettono sia singoli lessemi che combinazioni di parole. Il capitolo 5 "Modellazione di un dizionario delle collocazioni italiane" traduce in un modello lessicografico concreto i risultati ottenuti nei diversi stadi di analisi precedenti; una entrata lessicografica rappresentativa è presente in allegato come lista sorgente. Il capitolo parte da una esposizione delle caratteristiche funzionali del dizionario per descriverne successivamente nel dettaglio la struttura. Fulcro tematico dell'intero lavoro è la sezione sulla microstruttura del dizionario, nella quale prende corpo la classificazione finale delle collocazioni secondo un complesso sistema di interconnessioni a diversi livelli linguistici (sintattico, ontologico-semantico, pragmatico). Il capitolo si conclude con una sezione dedicata alle modalità di consultazione del dizionario e con delle riflessioni su una possibile futura rielaborazione del materiale collocativo, con un suo adattamento a finalità lessicografiche ancora più specifiche ed una sua applicazione ad una porzione di lessico più estesa, che superi i confini dello studio esemplare qui condotto.

L'intenzione di presentare in una sola opera aspetti semantici e sintattici comuni a più lessemi porta a prendere in considerazione in questa Introduzione le opere già sul mercato che si avvicinano anche solo in parte a questa prospettiva, allo scopo di valutarne vantaggi e svantaggi. La produzione nella lessicografia straniera (soprattutto inglese) di dizionari speciali dedicati ai vari tipi di forme fraseologiche o alle valenze sintattiche dimostra sia la possibilità concreta di realizzazione che l'utilità di queste opere, che godono già da alcuni decenni di grande popolarità tra disparate categorie di utenza. Nella panoramica che segue, si distingueranno alcuni modelli lessicografici, distinguibili sulla base del supporto utilizzato, dell'impostazione semasiologica o onomasiologica, come pure del tipo di dati lessicografici forniti. Verranno inoltre riportati esempi di come questi modelli elaborino il lemma italiano *paura* e, dove questo non sia fattibile, equivalenti in lingue diverse (ingl. *fear*, fr. *peur*). Non si entrerà nel merito della definizione di collocazione adottata per le singole opere lessicografiche¹.

Modello 1: dizionario generale semasiologico con inserimento di collocazioni/sezione fraseologica, focus sulle indicazioni semantiche

Esistono per la lingua italiana pochi dizionari generali che inseriscono in modo sistematico collocazioni e valenze sintattiche all'interno della glossa. Si tratta di opere relativamente recenti. Il loro intento è dettato dalla crescente attenzione a questi aspetti linguistici in ambito lessicografico, ed è senz'altro apprezzabile da questo punto di vista. Tuttavia, le informazioni sul lemma che queste opere puntano ad offrire sono talora molto generiche e poco gestibili da parte dell'utente. Questi dizionari sono disponibili in forma sia cartacea che elettronica, anche se quest'ultima è una semplice trasposizione in formato digitale della prima, senza rielaborazioni del materiale lessicografico in termini di rinvii o di sistemi di ricerca articolati. Di seguito qualche esempio.

SC (2007) offre parzialmente delle indicazioni valenziali (sintattiche) sui lemmi (in particolare il numero di argomenti propri di un verbo) ed una consistente sezione fraseologica, nella quale tuttavia compaiono in prevalenza espressioni idiomatiche. Le collocazioni sono relegate al rango di esempi senza che vi sia una separazione tra combinazioni lessicali libere e non. La valenza sintattica di *paura*, certamente desumibile dagli esempi generali da parte di un utente relativamente competente, non viene isolata con sistematicità. La tipologia degli esempi forniti può inoltre portare facilmente a dei dubbi sul carattere libero o

1 Allo scopo di rendere trasparente l'utilizzo nelle prossime pagine dei termini collocazione e fraseologia, si rimanda al Quadro A in appendice, con una panoramica teorica dedicata a questi fenomeni.

meno di certe combinazioni (*matta e tremenda* si accompagnano tipicamente a *paura*, e se sì, sono gli unici aggettivi a farlo?) e sulla correttezza di costruzioni non riportate anche se estremamente frequenti nella lingua comune (è corretto dire *essere assalito dalla paura di morire?*). Si nota per di più come non siano registrate combinazioni intuitivamente significative come *incutere paura* o *tremare per la paura*, solo per citarne alcune, e come siano quasi del tutto assenti indicazioni sulle differenze di registro (solo di *farsela sotto dalla paura* si segnala l'appartenenza al lessico popolare).

Giacoma/Kolb (2009) è un dizionario generale bilingue che si presenta dichiaratamente come opera innovativa per l'inserimento sistematico di *collocatori* all'interno della glossa (figura 0.1). La versione elettronica consente una più agevole consultazione dell'opera, anche perché ciascuna parola della glossa è un elemento attivo che produce un rinvio al corrispondente lemma. Nella prima parte dell'articolo la valenza sintattica del sostantivo viene indicata esplicitamente solo nel primo esempio (integrato da due collocatori) e nel relativo traduce, mentre negli esempi successivi, estesi, è ottenibile tramite astrazione. La parte fraseologica, che si apre con un pallino, contiene forme polilessicali estremamente eterogenee: queste vanno dalla combinazione pressoché libera (*fare/mettere paura a qu*) all'espressione idiomatica fino al proverbio, sono presentate in forma impersonale e concisa (*vivere nella paura*) e seguite da forme personali ed estese (*vive nella paura di ammalarsi*), e non sono suddivise in base alla loro costruzione sintattica. Alcuni problemi legati a questa impostazione sono ad esempio la mancata indicazione del criterio alla base della differenziazione tra le combinazioni libere e fraseologiche, come anche di quello alla base dell'attribuzione al ruolo di collocatori di certe parole piuttosto che di altre (perché *guerra* e *solitudine*?)². Una collocazione molto comune quale *incutere paura* si trova ad esempio registrata solamente sotto la forma verbale.

2 Non viene in realtà operata una distinzione tra basi e collocatori nel senso di Hausmann (1984), per cui non si può nemmeno affermare che si sia scelto di registrare le collocazioni di preferenza sotto le une o sotto gli altri: le parole che si accompagnano in modo tipico ad un lemma vengono sempre indicate come collocatori.