

La variazione semasiologica di sottocodice

Un esempio di analisi lessicale corpus-based

Leo Budinich
Università di Torino

This paper aims to illustrate how corpus linguistics techniques allow to study the phenomenon of semasiological variation from a sociolinguistic perspective. Semasiological variation can be described as that kind of lexical variation that implies the presence of two or more meaning-variants for a single form, where the variation between semantic structures can correlate with the variation of social variables. Adopting a prototypically structured model of lexical meaning and using collocation analysis as a means to provide a representation of the usage features exhibited by a lemma in a corpus, assumed as a representative sample of a given sociolinguistic variety, it becomes possible to characterize the different semantic configurations displayed by a lexeme in different language varieties. Four case studies of Italian lexemes illustrate a possible implementation of this kind of methodology with regard to the sociolinguistic dimension of field variation, using data extracted from the RIDIRE web-corpus.

Keywords: semasiological variation, sociolinguistics, prototype semantics, collocations

1. Inquadramento teorico

1.1 Introduzione

Se confrontiamo i due frammenti di testo che seguono, raccolti rispettivamente da un articolo del quotidiano *La Repubblica* del marzo 2015 e da un articolo de *La Gazzetta dello Sport* del marzo 2013, cioè da due fonti il cui uso della lingua non dovrebbe risentire di particolari influenze locali, l'aggettivo *opportunist*a pare assumere significati differenti in base ai contesti in cui è inserito:

- (1) Il mio personaggio parte che è un mediocre, viscido opportunist

- (2) È stata una prestazione di grande spessore, quella dell'attaccante Davide Cais nella gara in cui l'Atalanta si è imposta sul Sassuolo per 3-1. Generoso, intelligente, opportunista, si è mosso molto bene dettando diverse azioni pericolose.

Mentre nella prima frase di (1) *opportunist*a è coordinato ad aggettivi dal valore negativo come *mediocre* e *viscido*, ed è a questa triade di attributi che sembra riferirsi l'espressione *lato oscuro* presente nella frase seguente, in (2) l'aggettivo si accompagna a *generoso* e *intelligente*, e pare riferito ad uno sportivo di cui si cita la *prestazione di grande spessore* e di cui si dice che *si è mosso molto bene*. Nei due frammenti, cioè, *opportunist*a assume valori connotativi opposti: negativo in (1), positivo in (2). Il dizionario *Nuovo De Mauro* descrive un *opportunist*a come “chi si comporta con opportunismo”, mentre, *opportunismo* viene descritto così¹:

1. Atteggiamento di chi, adattandosi di volta in volta alle circostanze del momento, agisce pensando esclusivamente al proprio tornaconto: *fare qcs. per opportunismo, opportunismo politico*. 2. Spec. nelle competizioni sportive, capacità di saper approfittare delle minime occasioni favorevoli che si presentano per segnare un punto, per vincere, ecc.

Il *Nuovo De Mauro*, cioè, conferma il fatto che *opportunist*a può assumere sia un significato negativo, descritto nell'accezione 1, sia un significato più positivo, descritto nell'accezione 2. È interessante notare, però, che il significato positivo viene assunto da *opportunist*a “specialmente nelle competizioni sportive”: sembrerebbe, dunque, che nel lessico dello sport l'aggettivo in questione possa assumere un significato diverso rispetto a quello esibito nella lingua comune. Da un punto di vista linguistico, di fronte ad una *variazione del significato* come questa ci si possono porre diverse domande: (a) come inquadrare questo tipo di fenomeno da un punto di vista sociolinguistico? (b) come può essere concepita la presenza di un'oscillazione del significato sul piano semantico? (c) come si può studiare questo tipo di fenomeno senza basarsi solamente sulle proprie impressioni soggettive? Il resto della presente sezione è diviso in tre parti, ognuna delle quali risponde grosso modo a uno degli interrogativi formulati poco sopra. La prima, intitolata “la variazione lessicale semasiologica”, è dedicata a inquadrare in una prospettiva sociolinguistica ciò che abbiamo appena chiamato *variazione del significato*; la seconda, “quale teoria semantica?”, illustra la possibilità di analizzare la variazione semasiologica all'interno di un quadro teorico cognitivista; la terza,

¹ L'edizione consultata è quella ospitata online sul sito del settimanale Internazionale (<http://dizionario.internazionale.it/>).

intitolata “le tecniche di misurazione”, spiega come l’analisi delle collocazioni possa essere utilizzata per studiare empiricamente i fenomeni di variazione semasiologica. La seconda sezione, intitolata “implementazione dell’analisi”, è divisa in tre parti: nella prima, “modelli analitici di riferimento”, si indicano le scelte metodologiche su cui si è basata l’analisi empirica; nella seconda, “il corpus RIDIRE”, le caratteristiche del corpus scelto come base di dati; nella terza, “selezione dei lessemi e procedura d’analisi”, si illustrano le procedure operative attraverso cui si è svolta l’analisi. La terza sezione, “casi di studio”, illustra i risultati ottenuti dell’analisi di quattro lessemi dell’italiano corrente. Seguono infine le “conclusioni”, in cui vengono presentate alcune considerazioni generali.

1.2 La variazione lessicale semasiologica

Da un punto di vista sociolinguistico il caso di *opportunist* costituisce un esempio di *variazione lessicale*, cioè quel tipo di variazione per cui «lexical choices a speaker makes may characterise their place in society» (Fabiszak 2005: 1737). In generale, la variazione lessicale può essere di tipo *onomasiologico* o *semasiologico*. Si tratta di due prospettive analitiche speculari con cui approcciarsi al dato lessicale, che possono essere descritte così:

the semasiological perspective takes its starting point in the word as a form, and describes what semantic values (as dependent variable) the word (as independent variable) may receive. The onomasiological perspective takes its starting point on the level of semantic values and describes how a particular semantic value (as independent variable) may be variously expressed by means of different words (as dependent variables). (Geeraerts, Grondelaers & Bakema 1994: 5)

La *variazione lessicale onomasiologica*, in altre parole, può essere definita sommariamente come la possibilità di esprimere un medesimo concetto attraverso forme differenti, mentre la *variazione lessicale semasiologica* come la possibilità di esprimere concetti differenti con una medesima forma.

In una prospettiva sociolinguistica tali possibilità di realizzare uno stesso significato attraverso forme differenti o di codificare strutture concettuali differenti attraverso una medesima forma possono veicolare un significato sociale, cioè essere in correlazione con determinati fatti sociali.

La variazione semasiologica non è stata molto esplorata in sociolinguistica (Robinson 2010: 86); ciò è dovuto sostanzialmente al fatto che nella sociolinguistica correlazionale lo strumento analitico fondamentale è costituito dalla *variabile sociolinguistica* (Tagliamonte 2006: 70). Questa è definibile come

ogni insieme di modi alternativi di dire la stessa cosa, di realizzazioni diverse di un'unità o entità del sistema linguistico, in cui le realizzazioni diverse abbiano pertinenza sociale, rechino significato sociale, correlino con variabili sociali (in senso ampio) (Berruto 2006: 132-133, corsivo nostro)

La variazione lessicale semasiologica, che implica una stabilità della forma ed una realizzazione variabile del significato, non può perciò essere concepita in termini di variabili sociolinguistiche, proprio per via dei termini con cui queste ultime sono definite (Cerruti 2011). Non per questo, però, è stato negato l'interesse sociolinguistico di questo tipo di fenomeno: nel definire una variabile come «two alternative ways of saying the same thing», Labov indica come «the converse is also included in the study of variation: situations where there are alternative meanings conveyed by the same form» (Labov 2008: 2). Lo studio di questo genere di variazione, dunque, non è stato molto praticato più che altro a causa dell'impossibilità di usare il quadro di riferimento tradizionale della sociolinguistica laboviana; come ha osservato ancora Labov,

our methods are precise and reliable when our dependent variable is the choice of two forms for the same meaning; when we are dealing with a choice of two meanings for one form it is a different matter (Labov 1978, 8)

Anche se il concetto di variabile sociolinguistica non è applicabile allo studio della variazione semasiologica, quest'ultimo può comunque essere perseguito tramite l'adozione di una concezione più generica di *variabile*. Quest'ultima, infatti, può essere definita in senso ampio come la «proprietà operativizzata di un oggetto» (Corbetta 2003: 2, 33), dove la *proprietà di un oggetto* è ciò che si produce dall'applicazione di un certo concetto (ad esempio, quello di “significato”), ad un oggetto (ad esempio, ad un lessema, come *opportunist*). Secondo la trattazione che ne dà Corbetta, una proprietà può assumere *stati* diversi su diverse *unità di analisi* (ad esempio, il significato di *opportunist* può variare a seconda delle singole occorrenze esaminate), mentre la possibilità di distinguere gli stati assunti da una proprietà su diverse unità d'analisi passa attraverso la cosiddetta *operativizzazione*, che è quell'insieme di operazioni empiriche descritte tramite regole definite attraverso cui una proprietà viene “misurata”, cioè resa osservabile oggettivamente e trattabile statisticamente (*ibidem*, 27-29). È tramite l'operativizzazione, quindi, che «si realizza la traduzione empirica di una teoria» (*ibidem*, 26), e che si può studiare il comportamento di una variabile così definita.

Ora, se adottiamo questa definizione, sembra che per studiare la variazione semasiologica in una prospettiva sociolinguistica, ovvero per osservare la presenza di correlazioni fra fatti sociali e variazione semasiologica, dovremmo aver

bisogno di (a) una teoria del significato lessicale che permetta di spiegare casi come quello di *opportunist* in termini di variazione e (b) di tecniche che permettano di tradurre tale apparato teorico in operazioni di misurazione empirica. Come nota Hasan, è proprio l'assenza di questi due elementi, cioè sia di una teoria semantica adatta a concepire la possibilità di variazione semasiologica che di tecniche tramite cui misurarla, che fa negare a Labov la possibilità di studiare questo tipo di fenomeno in chiave sociolinguistica (Hasan 2009: 50-51). Tuttavia, esistono opzioni alternative all'approccio al significato lessicale prescelto da Labov, mentre le tecniche per analizzare l'uso della lingua hanno subito importanti innovazioni negli ultimi quarant'anni, anche grazie al contributo della linguistica dei corpora.

1.3 Quale teoria semantica?

Il concetto di variazione viene definito da Berruto come la

proprietà di un'entità di assumere diverse manifestazioni, di essere realizzata in modi differenti, di conoscere alternative che non mutano però la natura fondamentale dell'entità stessa. (Berruto 2011: 1547)

In questo senso, la variazione semasiologica dovrebbe implicare che di un lessema sono possibili attualizzazioni semantiche variabili. Una concezione referenziale del significato come quella adottata da Labov (Cerruti 2011: 213), però, non pare permetta di parlare di un'*attualizzazione variabile* del significato. Nella sua disamina su «quale sia esattamente il luogo della variazione del significato lessicale», Berruto osserva che per poter parlare di variazione semasiologica sembra necessario chiamare in causa almeno una terza entità oltre al significato linguistico e al referente, cioè un livello concettuale che medi fra realtà e piano semantico: con la presenza di una terza entità, mentre il significato sistemico del lessema rimane stabile, può variare la

particolare rappresentazione concettuale che viene attivata in relazione a un determinato referente (o una classe di referenti) o un determinato stato di cose (o una classe di stati di cose, *Sachverhalten*) in una certa situazione specifica (Berruto 2010: 16-17)

Di un dato lessema, in altre parole, possono variare i *significati contestuali* o *sensi* attivati in specifici contesti². Seguendo questa linea di pensiero, possiamo osservare che le teorie semantiche adatte per parlare di variazione semasiologica devono (a) assumere l'esistenza di un piano concettuale, (b) adottare una concezione *contestuale* del significato.

Fra le teorie semantiche che permettono di intendere in questi termini il significato lessicale, può essere particolarmente interessante la cosiddetta *teoria dei prototipi* sviluppata all'interno della linguistica cognitivista³, in quanto:

- a. prevede l'esistenza di un piano concettuale che media fra le parole e i referenti, frutto di un processo cognitivo di concettualizzazione tramite cui gli individui elaborano i dati della propria esperienza e li sintetizzano all'interno di rappresentazioni schematiche e idealizzate della realtà (Ježek 2011: 74-75);
- b. per quanto miri soprattutto a rendere conto dell'organizzazione delle strutture semantiche sul piano psicologico (Casadei 2003: 24), permette di intendere le categorie strutturate prototipicamente anche a livello di comunità linguistica (Murphy 2002: 391-392), e quindi di concepire la struttura semantica di un dato lessema in termini sociolinguistici (cfr. Geeraerts 2008);
- c. prevede che la struttura semantica dei lessemi si costruisca dinamicamente attraverso il loro uso in contesto; in altri termini, una struttura prototipica del significato è l'esito di una continua sovrapposizione di atti di classificazione contestuale (Geeraerts 2010: 187-188), e in questo senso adotta una concezione pienamente contestuale del significato linguistico.

Per via di (c), la teoria dei prototipi offre la possibilità di operativizzare il concetto di *significato* di una forma lessicale in termini di *frequenza d'uso contestuale*. Questo passaggio logico è permesso dalla nozione di *entrenchment*. Con questo termine nella teoria cognitivista viene indicato quel processo di radicamento mnemonico di un concetto o di una coppia forma-concetto nella conoscenza linguistica di un parlante, ovvero

² Adottiamo qua la definizione di Casadei di *sensu*: «ogni volta che un parlante usa all'interno del discorso una particolare parola, sceglie di trasmettere solo uno dei possibili significati di quella parola, e a sua volta l'ascoltatore dovrà interpretare ciò che viene detto scegliendo, tra le interpretazioni possibili, quella a suo avviso più congruente con la situazione comunicativa e più rispondente alle intenzioni comunicative dell'interlocutore. Questo significato specifico, individuato in quel particolare contesto comunicativo da quei particolari locutori, è appunto il senso.» (Casadei 2003: 35-36).

³ Per un'introduzione a questa teoria, cfr. Geeraerts (2010), pagg. 182-203.

the storage of concepts and constructions as (variably) routinized items in long-term memory, [that] accounts for the emergence of linguistic items with a high degree of unit-hood, that is, symbolic associations between semantic and phonological structures (Schmid 2007: 121)

L'*entrenchment* intrattiene una relazione molto stretta con la frequenza d'uso: più un significato specifico o funzione di una forma linguistica viene usato frequentemente, più si "radica" (*ibidem*, 119). Se pensiamo agli effetti di questo meccanismo a livello di comunità parlante, intesa come insieme delle conoscenze linguistiche dei singoli individui, possiamo allora dire che più il senso di un lessema è usato socialmente, più è associato stabilmente a tale forma a livello di uso linguistico (Glynn 2014: 13-14). La struttura prototipica di un lessema, cioè, può venire analizzata empiricamente in termini di frequenze d'occorrenza degli usi contestuali: più un senso/uso contestuale è frequente, più sarà radicato, e quindi strutturalmente importante nel determinare la struttura semantica complessiva del lessema (*ibidem*, 27). Tramite la nozione di *entrenchment*, cioè, diventa possibile "misurare" la struttura prototipica assunta da un certo lessema in un dato contesto sociale, e dunque studiare il rapporto di correlazione esistente fra le varianti semantiche manifestate da un lessema e le variabili extra-linguistiche che definiscono contesti sociali differenti.

1.4 Le tecniche di misurazione

Anche se il senso assunto da un lessema in un dato contesto comunicativo non viene determinato solamente dal suo contesto linguistico d'occorrenza, è ragionevole sostenere che quest'ultimo può dirci più di qualcosa sull'interpretazione da assegnare ad una parola in date circostanze. Se prendiamo una parola come *gatto*, ad esempio, da una frase come

(3) L'umano e il gatto la videro sollevare la testa con gli occhi chiusi.

possiamo inferire che, probabilmente, (a) *gatto* in questo contesto indica un'entità animata, in quanto soggetto del verbo *vedere*, (b) *gatto* indica un'entità non-umana, in quanto la parola che denota l'entità è coordinata con *umano*. Certo, non possiamo sapere molto di più sull'entità di cui si sta parlando, cioè individuarne altre caratteristiche concettuali, ma quantomeno sappiamo che il referente di *gatto* è, ragionevolmente, un'entità vivente che possiede degli organi visivi. Il contesto linguistico d'occorrenza, dunque, può essere considerato almeno un indicatore parziale del senso assunto da un elemento lessicale in un certo contesto comunicativo. Se, dato un lessema, raccogliamo quindi un insieme dei suoi esempi d'uso

in contesto e ne analizziamo i rapporti sintagmatici più frequenti, cioè le *caratteristiche distribuzionali*, dovremmo poter considerare i risultati di questa analisi come degli indici delle sue caratteristiche concettuali. Questa intuizione è alla base di quella che è stata chiamata l'*ipotesi distribuzionale*: l'idea che «la distribuzione statistica delle parole nei contesti linguistici giochi un ruolo chiave nel caratterizzarne il comportamento semantico» (Lenci 2010: 57).

Questa ipotesi è stata avanzata da diversi studiosi nel corso degli anni⁴, ma la sua formulazione più nota e icastica resta quella proposta da John Firth con le parole «[y]ou shall know a word by the company it keeps» (Firth 1957: 11) e tramite la nozione di *collocazione*, intesa teoricamente come «an abstraction at the syntagmatic level» (McEnery & Hardie 2012: 123) e praticamente come «a combination of two words that exhibit a tendency to occur near each other in natural language, i. e. to co-occur» (Evert 2009: 1214). Questa nozione costituisce la chiave di volta della possibilità di “misurare il significato”, in quanto una collocazione può essere intesa sia come (a) un fenomeno empirico, cioè un evento direttamente osservabile nei dati linguistici, che come (b) un epifenomeno, cioè il risultato visibile dell'azione di fattori linguistici non direttamente osservabili (*ibidem*, 1218-1219). Infatti, se assumiamo che i corpora costituiscano dei campioni statistici di date varietà di lingua, possiamo costruire una rappresentazione empirica dell'uso esibito da un dato lemma in una data varietà esaminando i suoi collocati in un corpus rappresentativo. A sua volta, questa rappresentazione può essere assunta come un indice della struttura semasiologica del lessema in tale varietà, e di conseguenza i differenti insiemi di collocati di un medesimo lemma in corpora rappresentativi di varietà differenti possono essere interpretati come rappresentazioni delle varianti semantiche manifestate dal lessema in tali varietà.

Tramite l'adozione di una semantica lessicale prototipica e l'uso di corpora e relative tecniche di analisi dei collocati, dunque, disponiamo sia di una teoria che di una tecnica di misurazione adatte a studiare scientificamente cosa ci possa dire su un parlante o un gruppo di parlanti il fatto che usi una data parola con un dato senso in un dato contesto.

Alcuni esempi di questo tipo di approccio allo studio della variazione semasiologica sono costituiti dagli studi di Wulff, Stefanowitsch e Gries (2007), Peirsman, Heylen e Geeraerts (2010), Glynn (2013).

⁴ Per una breve rassegna di alcuni linguisti che hanno affrontato il tema, cfr. Gries 2012, 57-58.

2. Implementazione dell'analisi

2.1 Modelli analitici di riferimento

2.1.1 *Un modello di variazione delle strutture prototipiche*

Il modello semantico delle strutture prototipiche porta a ipotizzare la presenza di differenti effetti di variazione sulle strutture semantiche di un lessema; seguendo il modello proposto da Geeraerts 1997 per la variazione diacronica, abbiamo ipotizzato di poter individuare tramite l'approccio testuale quantitativo tre differenti modalità con cui si può manifestare la variazione della struttura semasiologica di un lessema:

- a. variazione della tipicità delle accezioni polisemiche;
- b. presenza o assenza di sotto-accezioni periferiche⁵;
- c. presenza di sensi particolarmente frequenti (Geeraerts 1997: 23-26).

In pratica, ci aspettiamo di poter trovare casi in cui:

- A. un'accezione è più frequente (e quindi più radicata) in una data varietà di lingua;
- B. un'accezione specializzata è presente o assente solamente in date varietà di lingua;
- C. un senso specifico è usato con grande frequenza solamente in date varietà di lingua.

2.1.2 *Un modello di analisi delle collocazioni*

Per quanto riguarda le modalità con cui si è operativizzato il concetto di collocazione, abbiamo adottato un approccio ispirato a due modelli empirici di analisi dei significati lessicali, cioè quello neo-firthiano dell'*extended unit of meaning* (Stubbs 2002; Sinclair 2004) e quello cognitivista dei *behavioral profiles* (Gries 2012; Glynn 2013). Nello specifico, abbiamo adottato:

⁵ Per quanto riguarda la polisemia nella teoria dei prototipi, oltre a Geeraerts (2010), cfr. Taylor (1995: 99-121) e Lewandowska-Tomaszczyk (2007); si tenga conto che in tale paradigma teorico le accezioni in cui si articola un lessema polisemico vanno intese come aggregazioni di usi contestuali che si collocano all'interno della regione concettuale delimitata dal significato di un lessema e che possiedono «differenti gradi di salienza e coesione» (Croft & Cruse 2010: 138); intenderemo qui la polisemia di un lessema soggetto a variazione semasiologica come una «polisemia sociale», che concerne cioè la molteplicità delle accezioni, più o meno recenti, più o meno diffuse, che un lessema ha assunto e può assumere nel lessico di una lingua, e non nel lessico mentale di un singolo parlante (cfr. Nerlich 2003: 61).

- a. un contesto di co-occorrenza sintattico, in modo da catturare anche le relazioni di co-occorrenza esterne ad uno *span* arbitrario;
- b. una procedura di identificazione delle collocazioni condotta esplorativamente tramite misure di associazione statistica, e nel dettaglio tramite spoglio manuale delle concordanze;
- c. una definizione di nodo basata sui lemmi, al fine di rendere conto della struttura semasiologica dei lessemi e non solamente delle singole forme flesse;
- d. le nozioni neo-firthiane di *collocazione*, *colligazione* e *preferenza semantica* (per cui cfr. Stubbs 2002: 64-65) di cui l'ultima operativizzata in primo luogo secondo quanto proposto da Gries (2012; cfr. *ibidem*, 60-62), e secondariamente attraverso la nozione di similarità semantica di Stubbs;
- e. una modalità basilare di rappresentazione delle caratteristiche d'uso contestuale di un lessema, basata sulla semplice indicazione, per ogni fattore analitico, delle frequenze di occorrenza relativa, senza avvalerci di tecniche di statistica esplorativa.

2.2 Il corpus RIDIRE

2.2.1 Caratteristiche e struttura di RIDIRE

Il corpus da cui sono stati raccolti i dati per l'analisi dei casi empirici è il web-corpus RIDIRE, che consiste in una risorsa di oltre 1 miliardo e 300 milioni di parole estratte da siti web fra il 2009 e il 2013 (Cresti & Panunzi 2013: 154; Moneglia & Paladini 2010; Cresti & Moneglia 2012; Moneglia, Colombo & Panunzi 2014; Panunzi, Cresti & Gregori 2014). RIDIRE è il frutto di una procedura di *web-crawling* "mirato", che ha previsto l'applicazione al processo di *crawling* di criteri di bilanciamento modellati su quelli adottati per il British National Corpus (Cresti & Panunzi 2013: 63). Il risultato è che il corpus è suddiviso in dodici sub-corpora corrispondenti ciascuno a un dominio funzionale (in cui la lingua è caratterizzata dalla funzione con cui viene usata) o semantico (in cui la lingua è caratterizzata dall'insieme degli argomenti per cui viene usata) dell'italiano, selezionati in modo da rappresentare «sufficientemente l'italiano nell'insieme delle sue varietà d'uso» (Moneglia & Paladini 2010: 121). In questo modo la risorsa mira a porsi come una sorta di *reference corpus* per questa lingua. Certo, formalmente il suo essere rappresentativo solamente di quella particolare varietà diamesica detta "italiano del web" non permetterebbe di considerarlo un campione statistico rappresentativo dell'italiano scritto tradizionale (cfr. Tavasani 2011: 39, 45), ma in pratica, dati (a) il suo *corpus design* bilanciato e (b) la generale comparabilità lessicale che anche i *web corpora* non bilanciati hanno dimostrato rispetto alle

risorse di riferimento dello scritto tradizionale (Cresti & Panunzi 2013: 63), assumeremo che RIDIRE sia sostanzialmente equiparabile ad un *reference corpus*.

Tabella 1. Numero di siti e pagine web oggetto del *crawling* e numero di parole di ogni sub-corpus di RIDIRE (Panunzi, Cresti & Gregori 2014: 449).

DOMINI	Siti Web	Pagine	Parole
Amministrazione e Legislazione	39	199.756	399.313.920
Economia e Affari	123	226.535	161.377.152
Informazione	27	550.169	186.577.769
<i>Funzionali (totale)</i>	189	976.460	747.268.841
Architettura e Design	142	136.725	81.235.939
Arti Visive	118	70.601	48.929.903
Cinema	25	122.850	44.370.692
Cucina	20	123.376	45.523.096
Letteratura e Teatro	113	61.935	73.204.712
Moda	103	74.584	21.690.140
Musica	195	113.015	106.287.283
Religione	51	66.053	62.291.806
Sport	49	138.235	82.695.548
<i>Semantici (totale)</i>	816	907.374	566.229.119
<i>Totale</i>	2.010	3.767.668	1.313.497.960

La scelta è ricaduta su questa risorsa per:

- a. le sue dimensioni, sia complessive che relative ai singoli sub-corpus, che permettono di avere a disposizione un numero elevato di occorrenze anche per lemmi di media frequenza;
- b. il suo essere un insieme di sub-corpora costruiti secondo uno stesso *design*, il che permette di evitare le complicazioni che comporterebbe confrontare i collocati di un medesimo lemma in corpora con *design* differenti (cfr. Gries 2006; Baker 2010: 91-96);
- c. l'origine recente dei materiali che lo compongono, che permette di ipotizzare una minor presenza di effetti diacronici sulla variazione lessicale;
- d. il criterio tramite cui è avvenuta la divisione in sub-corpora, in quanto permette di considerare i singoli domini d'uso come campioni rappresentativi di determinate varietà di lingua.

2.2.2 RIDIRE e varietà di lingua

Da una prospettiva sociolinguistica, l'ultima osservazione necessita di qualche parola in più. Poiché i testi che compongono RIDIRE sono classificati solamente

in base all'argomento che li accomuna o alla funzione per cui sono stati redatti, l'unico asse di variazione su cui possiamo collocare i sub-corpora di RIDIRE è l'asse diafasico, in particolare all'interno della cosiddetta *variazione di sottocodice*. Fra le varietà di lingua definite in base alla variazione di sottocodice, dette anche *lingue speciali* o *linguaggi settoriali*, seguendo Berruto possiamo distinguere:

- a. le *lingue speciali in senso stretto*, «cioè i sottocodici veri e propri, forniti e contrassegnati da un proprio lessico particolare ed eventualmente da tratti di morfosintassi e testualità caratteristica», come possono essere la *lingua della chimica* o la *lingua del nuoto*;
- b. le *lingue speciali in senso lato*, che «non hanno propriamente un lessico specialistico ma sono comunque strettamente legate a determinate aree di impiego, e sono caratterizzate da scelte lessicali e da formule sintattiche e testuali», come possono essere la *lingua delle riviste di moda* o la *lingua degli oroscopi* (Berruto 2012: 178).

Queste due classi possono poi essere considerate come due zone più o meno centrali all'interno di un continuum, ad uno degli estremi del quale (in una posizione più prossima all'area delle lingue speciali in senso lato), si collocano le *modalità d'uso*, cioè quei

conglomerati di stili, formule, tipi di testi, varietà speciali plurime frammischiate, ecc. che vengono utilizzati nei grandi domini della società, e che non possono più essere riconosciuti come varietà di lingua in senso sia pure latissimo (*ibidem*, 181-182)

come può essere la *lingua della pubblicità*. Ora, se prendiamo in considerazione i dieci domini semantici di RIDIRE, definiti dal macro-argomento da cui prendono il nome, possiamo certamente escludere la possibilità di considerarli come campioni di lingue speciali in senso stretto, dato che non paiono caratterizzati dalla presenza di un vero e proprio lessico specialistico; rimane allora la possibilità di considerarli rappresentativi, per lo meno in linea di principio, di lingue speciali in senso ampio, tenendo conto, però, che siamo al confine con le modalità d'uso. Rispetto ai tre domini funzionali, invece, definiti non dall'argomento del discorso ma piuttosto dall'uso funzionale della lingua che ne accomuna i testi, la situazione è parzialmente diversa: mentre i testi di *Amministrazione e Legislazione* e di *Economia e affari* paiono poter essere considerati abbastanza ragionevolmente campioni più o meno rappresentativi rispettivamente del *linguaggio*

giuridico-amministrativo e della *lingua dell'economia*, cioè di due varietà linguistiche considerabili come linguaggi settoriali (Cortelazzo 2010: 588; Proietti 2010: 411), quelli presenti in *Informazione*, invece, possono essere visti come campioni della *lingua dei giornali*, che viene generalmente considerata una modalità d'uso (Bonomi 2010). In definitiva, possiamo dire che per 11 sub-corpora su 12 è possibile (sebbene un po' forzosamente) parlare di "campioni rappresentativi di varietà di sottocodice"; ciò non è fattibile per il dominio d'uso *Informazione*, ma ai materiali di questo sub-corpus assegniamo un ruolo diverso nel processo d'analisi.

2.3 Selezione dei lessemi e procedura d'analisi

Per quanto riguarda la scelta dei lessemi da sottoporre all'analisi empirica si sono presi in considerazione soprattutto termini appartenenti al *vocabolario di base*⁶, in modo da evitare che fossero stilisticamente, regionalmente o settorialmente marcati e per essere sicuri che fossero sufficientemente attestati nell'uso comune. Ci si è limitati alle classi dei sostantivi e degli aggettivi; i verbi sono stati evitati per via dell'elevato grado di polisemia (e quindi complessità a livello di struttura semantica) che possono presentare e per l'alto numero di caratteristiche d'uso di cui era necessario tenere conto; gli avverbi perché poco presenti nel vocabolario di base. Nello specifico, i sostantivi scelti sono *disinteresse* e *clandestinità*, gli aggettivi *responsabile* e *spaventoso*. Fra i sostantivi e gli aggettivi disponibili, la scelta è ricaduta su quelli indicati perché parevano mostrare un grado di polisemia moderato (e quindi una struttura prototipica non eccessivamente articolata), ma nel contempo una tendenza a poter assumere accezioni abbastanza distinte all'interno di contesti d'uso diversi.

Per ogni lessema sono stati selezionati tre domini di RIDIRE in cui confrontare l'uso del termine, di cui due variabili e uno fisso, costituito da *Informazione*. La scelta dei domini si è basata su:

- a. l'intuizione personale, che ha portato a selezionare alcuni domini dove la differenza d'uso fra lessemi pareva essere più marcata;
- b. la disamina delle definizioni fornite per i vari lemmi da parte di alcuni dizionari di consultazione, in modo da verificare se alcune accezioni fossero definite come più tipiche di date varietà di lingua;

⁶ Si è usata la lista presente in De Mauro (1980); l'unica eccezione è stata costituita da *clandestinità*, che è stato scelto per intuizione personale a seguito dello spoglio di alcuni quotidiani e riviste.

- c. la frequenza di occorrenza assoluta e standardizzata di ogni lemma in un dato dominio, in modo da escludere i domini in cui un lemma presentava un numero di occorrenze troppo basso, e da tenere conto anche di quanto un dato lemma fosse effettivamente usato nel lessico di un certo dominio, privilegiando i domini in cui la frequenza d'uso relativa era sufficientemente elevata.

Tabella 2. Frequenze d'occorrenza assolute (colonna sinistra) e standardizzate in occorrenze per milione (colonna destra) dei casi di studio nei diversi domini d'uso di RIDIRE. I domini in cui si è confrontato l'uso di un lemma sono evidenziati in grigio.

DOMINI	Clandestinità		Disinteresse		Responsabile		Spaventoso	
Amministrazione	1,145	2,87	987	2,47	36,245	90,77	75	0,19
Economia	18	0,11	146	0,90	17,815	110,40	47	0,29
Informazione	630	3,38	576	3,09	11,462	61,43	1,109	5,94
Architettura	60	0,74	347	4,27	4,972	61,20	281	3,46
Arti Figurative	25	0,51	149	3,04	928	18,97	547	11,18
Cinema	162	3,65	132	2,97	838	18,89	980	22,09
Cucina	2	0,04	15	0,33	393	8,63	60	1,32
Letteratura	195	2,66	363	4,96	1,404	19,18	1,334	18,22
Moda	3	0,14	5	0,23	692	31,90	71	3,27
Musica	67	0,63	447	4,20	848	7,98	775	7,29
Religione	94	1,51	446	7,16	3,501	56,20	406	6,52
Sport	6	0,07	87	1,05	1,908	23,07	343	4,15
<i>Totali</i>	<i>2,407</i>	<i>1,83</i>	<i>3,718</i>	<i>2,83</i>	<i>85,405</i>	<i>65,02</i>	<i>6,041</i>	<i>4,60</i>

Per ogni lessema è stato valutato l'uso contestuale anche nel dominio *Informazione* in quanto, anche se non è considerabile come campione di una varietà di lingua, si è ritenuto che l'analisi dell'uso dei lessemi nella *lingua dei giornali* potesse essere d'interesse, visto che «language in the mass-media is itself highly influential, news-papers are the most widely read long texts, and collocations from this source are therefore important» (Stubbs 2002: 68); inoltre, si è ipotizzato che il dominio *Informazione* potesse svolgere un ruolo di intermediazione nel diffondere un certo uso di un lessema da un dominio specialistico verso altri domini, svolgendo in qualche maniera il ruolo della lingua comune in quei processi di osmosi fra sottocodici e lingua comune che caratterizzano la variazione di sottocodice (Rovere 2011: 40).

Per ognuno dei lessemi selezionati sono stati estratti i primi venti collocati statistici relativi ai sub-corpora scelti per l'analisi, attraverso l'apposita sezione disponibile all'interno dell'interfaccia web del corpus. I parametri adottati per l'estrazione sono stati:

- a. span di +/- 4 parole;
- b. soglia minima di frequenza di occorrenza dei collocati nel corpus di 5 casi;
- c. soglia minima di frequenza di co-occorrenza con il nodo di 5 casi;
- d. definizione del nodo ed estrazione dei collocati in termini di lemma;
- e. adozione della misura di significatività statistica Dice⁷.

I collocati statistici così identificati sono stati considerati come una traccia preliminare per valutare le possibili differenze d'uso dei lessemi nei diversi domini. Dopo aver estratto i collocati relativi ad ogni dominio d'uso per i lessemi esaminati, si è proceduto a raccogliere, per ognuno dei domini e per ciascuno dei lemmi prescelti, le prime 100 concordanze non ordinate restituite per il lemma e il dominio selezionati. Successivamente, per ogni lessema si sono sintetizzate sulla base di alcuni dizionari di consultazione le accezioni principali che il termine può assumere⁸. Il numero delle accezioni sintetizzate per ogni lessema è stato ridotto ad un numero massimo di tre, escludendo gli usi specialistici propri di determinati sottocodici. In base alle definizioni di lavoro così ottenute, ognuna delle occorrenze dei lessemi all'interno delle concordanze estratte è stata classificata manualmente interpretandone l'accezione assunta in contesto. I dati relativi ad ogni lemma e relativi domini d'uso prescelti sono stati poi trasformati in tabelle che sintetizzassero quale accezione fosse quantitativamente prevalente in un dato dominio d'uso. Le frequenze di occorrenza delle differenti accezioni risultanti da questa procedura di classificazione, ispirata da Gries 2012 e da Robinson 2010, sono state considerate indicative della presenza di una variazione semasiologica del tipo riguardante il differente peso strutturale assunto da date accezioni nella struttura polisemica di un lessema in differenti varietà di lingua.

Dato che il procedimento di classificazione delle occorrenze secondo un numero ristretto di accezioni poteva (a) comportare la classificazione in accezioni distinte di occorrenze presentanti pattern d'uso sovrapponibili o viceversa, e (b)

⁷ L'interfaccia di interrogazione di RIDIRE indica che fra le misure di associazione disponibili è presente il log-Dice (cfr. Rychlý 2008); tuttavia, va notato che i valori numerici restituiti dal calcolo statistico selezionando tale opzione sono estremamente bassi (sempre al di sotto di 1), mentre quelli del log-Dice, in caso di associazioni positive, dovrebbero superare abbondantemente questa soglia, arrivando fino a un valore massimo di 14; si è ritenuto, quindi, che la misura calcolata effettivamente sia in realtà il coefficiente Dice. Va detto d'altro canto, che l'unica differenza esistente fra Dice e log-Dice riguarda la dimensione numerica dei coefficienti restituiti, e in nessun modo l'ordine o la differente significatività dei collocati identificati (cfr. *ibidem*, 8-9).

⁸ Le opere consultate sono: il *Devoto-Oli 2014*; il *Gradit 2007*; il *Sabatini-Coletti 2008*; il *Trecani on-line 2015*; lo *Zanichelli 2015*.

oscurare la presenza di altre modalità di variazione semasiologica, in seguito a tale classificazione si è proceduto anche ad un'analisi manuale delle caratteristiche d'uso frequenti individuabili in ognuno dei sottoinsiemi di occorrenze definiti dalla classificazione in accezioni. Le caratteristiche d'uso di cui si è tenuto conto sono state scelte di volta in volta osservando i pattern d'uso più salienti riscontrabili analizzando le concordanze, e in generale hanno riguardato il comportamento sintattico e le caratteristiche semantiche degli elementi linguistici che entravano in relazione sintattica con il lessema considerato. La presentazione dei risultati di quest'analisi ha preso la forma di un semplice conteggio delle frequenze relative (sul tipo di “*disinteresse* nell'accezione A è modificato da un aggettivo in 74 occorrenze su 220”) sintetizzato in tabelle. Successivamente all'elencazione delle caratteristiche d'uso preminenti per ogni sottoinsieme di accezioni, si è provveduto a valutare se alcune occorrenze classificate secondo una certa accezione in un dato dominio d'uso potessero manifestare dei pattern d'uso distintivi rispetto alla medesima accezione in altri domini. Nei casi in cui ciò si è verificato, si sono considerati questi casi come esempi di variazione semasiologica determinata dalla presenza di sensi salienti o di sotto-accezioni specifiche di una certa varietà di lingua.

3. Casi di studio

3.1 *Clandestinità*

Le definizioni di lavoro adottate per l'analisi del lemma *clandestinità* sono le seguenti:

- A. “Condizione di chi o di ciò che si trova in una situazione di segretezza a causa del suo infrangere norme giuridiche o sociali”
- B. “Condizione di chi è entrato in un Paese senza avere i relativi permessi, cioè di chi è un immigrato clandestino”

I domini d'uso di RIDIRE esaminati, invece, sono:

- a. *Letteratura e Teatro*
- b. *Amministrazione e Legislazione*
- c. *Informazione*

L'analisi dei collocati statistici estratti per *clandestinità* non è particolarmente informativa, in quanto le occorrenze della parola nel dominio *Letteratura e Teatro*

sono troppo esigue per restituire valori significativi (solamente i primi quattro lemmi, infatti, presentano una seppur minima forza associativa).

Letteratura e Teatro		Amministrazione		Informazione		
LEMMA	Dice	LEMMA	Dice	LEMMA	Dice	
1	lotta	0,0002	bocciare	0,0192	ronda	0,0048
2	costringere	0,0002	aggravante	0,0170	aggravante	0,0036
3	vivere	0,0001	bocciatura	0,0148	bocciare	0,0022
4	reato	0,0001	commistione	0,0124	reato	0,0019
5			incostituzionale	0,0097	aborto	0,0015
6			trapassato	0,0071	aggravare	0,0007
7			turco	0,0046	clandestino	0,0007
8			aggravare	0,0044	introduzione	0,0005
9			Ue	0,0042	contrastare	0,0005
10			coniugale	0,0034	combattere	0,0005
11			volontariamente	0,0034	punire	0,0005
12			illegalmente	0,0030	pacchetto	0,0004
13			introdotto	0,0027	introdurre	0,0003
14			testare	0,0024	immigrazione	0,0002
15			immigrazione	0,0023	lotta	0,0002
16			reato	0,0019	cessare	0,0002
17			reo	0,0019	contestare	0,0002
18			mafia	0,0018	spostare	0,0002
19			espulsione	0,0017	presunto	0,0002
20			contagio	0,0017	Berlusconi	0,0001

Tabella 3. Collocati statistici di *clandestinità*.

In ogni caso, si può notare come nelle prime tre posizioni di *Letteratura e Teatro*, compaiano due parole che paiono richiamare un'accezione "militante" del lemma, mentre nella quarta è presente il termine *reato*, comune anche agli altri due domini, che richiama invece l'ambito giudiziario. In *Amministrazione e Legislazione* e in *Informazione*, nelle prime quattro posizioni si osservano due lemmi comuni, *bocciare* e *aggravante*, che possono essere assunti come probabili segni del fatto che in questi due domini il lessema è utilizzato in maniera simile; anche nelle posizioni successive troviamo altri lemmi condivisi come *aggravare*, *immigrazione* e *introdurre* / *introdotto* / *introduzione*.

La distribuzione delle due accezioni nei tre domini presenta delle proporzioni quasi opposte fra *Letteratura e Teatro* da un lato e *Amministrazione e Legislazione* e *Informazione* dall'altro. Pare che l'accezione primaria di *clandestinità* in

Letteratura e Teatro sia A, mentre in *Amministrazione e Legislazione* e in *Informazione* sia B, indice abbastanza chiaro della presenza di una variazione semasiologica sul livello strutturale dell'articolazione in accezioni.

Tabella 4. Classificazione in accezioni di *clandestinità*

	Letteratura e teatro	Amministrazione	Informazione
A 'segretezza'	73	15	23
B 'immigrazione'	16	77	70
? (casi dubbi)	11	8	7

Per quanto riguarda le caratteristiche d'uso delle 274 concordanze non dubbie, sia per A che per B i pattern individuati si presentano quasi unicamente nel contesto sinistro del nodo: *clandestinità*, in altre parole, è raramente modificata da un aggettivo o da un sintagma preposizionale nelle concordanze esaminate. Peraltro, sembra emergere una differenza abbastanza chiara fra il comportamento sintattico e semantico delle due accezioni: *clandestinità* A viene trattata praticamente in un'occorrenza su due come un'entità quasi spaziale (cfr. le espressioni ricorrenti *entrare in clandestinità* e *uscire dalla clandestinità*), e co-occorre sintatticamente con una serie di verbi e sostantivi accomunati dalla dimensione semantica della "scelta"; *clandestinità* B, invece, ricorre sistematicamente come complemento di specificazione di un numero ristretto di termini giuridici (cfr. le ben 37 occorrenze di *reato di clandestinità*, interpretabili come un indice della salienza di un singolo senso) o di termini indicanti uno "stato" (cfr. le 21 occorrenze di *condizione di clandestinità*), e co-occorre blandamente con verbi o sostantivi accomunati dalla dimensione semantica del conflitto o della contrapposizione. Per quanto B sia più attestata, inoltre, pare presentare una minor varietà lessicale per quanto riguarda i sostantivi che vi co-occorrono sintatticamente, segno del fatto che il termine è spesso usato in espressioni semi-fisse. Per l'accezione B si può inoltre notare la presenza di alcune differenze d'uso fra *Amministrazione e Legislazione* e *Informazione* per quanto riguarda la co-occorrenza con alcuni sostantivi; sul piano semantico, tuttavia, non paiono essere abbastanza marcate da spingere ad ipotizzare l'esistenza di sotto-famiglie di sensi distinte all'interno di B, ma sembrano piuttosto riconducibili alle differenti caratteristiche discorsive dei due domini: è plausibile, ad esempio, che nei testi di *Informazione* non si parli molto di *stato di clandestinità* perché si tratta di una locuzione più tipica del lessico giuridico.

Tabella 5. Caratteristiche d'uso di *clandestinità*

	A 'segretezza'	B 'immigrazione'
colligazione con preposizioni:		
<i>in / nella clandestinità:</i>	42 su 111 (pos. -2: <i>entrare</i> (5), <i>vivere</i> (3), <i>operare</i> (2), <i>incontrarsi</i> , <i>sopravvivere</i> , <i>agire</i> , <i>scrivere</i> , <i>barcamenarsi</i> , <i>perdersi</i> , <i>ritrovarsi</i> , <i>rifugiarsi</i> , <i>documento pubblicato</i> , <i>scrittura praticata</i>)	/
altre preposizioni spaziali in pos. -1:	8 su 111 (pos. -2: <i>uscire dalla</i> (4), <i>passaggio alla</i> (2), <i>emergere dalla</i> , <i>togliere dalla</i>)	/
<i>di / della clandestinità</i>	/	96 su 163 (pos. -2 in 'informazione': <i>reato</i> (37), <i>aggravante</i> (4), <i>condizione</i> , <i>stato</i> ; pos. -2 in 'amministrazione': <i>condizione</i> (21), <i>stato</i> (6), <i>aggravante</i> (6), <i>situazione</i> (5), <i>accertamento</i> (2))
preferenze semantiche manifestate dai sostantivi / verbi di cui <i>clandestinità</i> è oggetto:		
'scelta' / 'non scelta'	11 su 111 (<i>scegliere la / la scelta della</i> (6), <i>costringere alla</i> (4), <i>condannare alla</i>)	/
'contrapposizione'	/	10 su 163 ('informazione': <i>contrastare la / contrasto alla</i> (4), <i>combattere la</i> (2), <i>lotta alla</i> (2); 'amministrazione': <i>lotta alla</i> (2))

Si può ritenere che B rappresenti l'accezione dominante del lessema in *Amministrazione e Legislazione* e in *Informazione* principalmente per via dell'impatto mediatico e normativo suscitato dall'istituzione del cosiddetto "reato di clandestinità" (risalente all'agosto 2009) nel periodo in cui sono stati raccolti i testi di RIDIRE. Sembrerebbe che questa innovazione lessicale di origine normativa sia stata immediatamente e largamente recepita nei due domini che, per le proprie caratteristiche, sono più propensi a essere influenzati dal lessico giuridico e della

politica, ma solo in misura minore da quello di un dominio d'uso meno sensibile ai temi dell'attualità come *Letteratura e Teatro*.

3.2 *Disinteresse*

Le definizioni di lavoro adottate per l'analisi di *disinteresse* sono le seguenti:

- A. "Mancanza di interesse, impegno, cura per qualcuno o qualcosa"
- B. "Noncuranza del proprio utile morale o materiale"

I tre domini d'uso di RIDIRE presi in considerazione, invece, sono:

- a. *Architettura e Design*
- b. *Religione*
- c. *Informazione*

Tabella 6. Collocati statistici di *disinteresse*.

Architettura e Design		Religione		Informazione		
LEMMA	Dice	LEMMA	Dice	LEMMA	Dice	
1	apparente	0,0012	abnegazione	0,0072	totale	0,0006
2	totale	0,0003	ammirevole	0,0034	verso	0,0003
3	diffuso	0,0003	indifferenza	0,0030	assoluto	0,0003
4	popolo	0,0002	zelo	0,0025	mostrare	0,0002
5	architetto	0,0002	dedizione	0,0022	completo	0,0002
6	abuso	0,0002	evangelico	0,0007	sostanziale	0,0002
7	generare	0,0002	generoso	0,0005	dimostrare	0,0002
8	verso	0,0001	sacrificio	0,0005	politico	0,0002
9	completo	0,0001	distacco	0,0005	mancanza	0,0001
10	dimostrare	0,0001	altrui	0,0003	generale	0,0001
11	politica	0,0001	spirito	0,0002	impegno	0,0001
12	riguardo	0,0001	atteggiamento	0,0002	passato	0,0001
13	confronto	0,0001	pratico	0,0002	personale	0,0001
14	attuale	0,0001	agire	0,0001	solito	0,0001
15	grave	0,0001	cristiano	0,0001	gente	0,0001
16	mancanza	0,0001	totale	0,0001	media	0,0001
17			impegno	0,0001	nessun	0,0001
18			amore	0,0001	problema	0,0001
19			Dio	0,0001	governo	0,0001
20			vostro	0,0001	argomento	0,0001

Per quanto riguarda i collocati statistici di disinteresse si può osservare che, mentre fra *Architettura e Design* e *Informazione* compaiono alcuni lemmi in comune, la lista generata per *Religione* pare offrire un panorama differente. Fra i primi

venti collocati di *Architettura e Design* e *Informazione* compaiono in entrambe le liste *totale, completo, dimostrare*, e si individuano diversi lemmi riconducibili agli ambiti dell'opinione pubblica e della politica, come *popolo, politica, politico, gente, media, governo*; anche in *Religione* compare l'aggettivo *totale*, ma gli altri lemmi paiono richiamare soprattutto gli ambiti dell'etica e della morale, con collocati dalle connotazioni positive come *abnegazione, zelo, ammirevole, dedizione, generoso*.

Osservando i risultati della classificazione in accezioni, invece, si nota una netta differenza nella frequenza d'uso delle due accezioni del lessema fra *Architettura e Design* e *Informazione* da un lato e *Religione* dall'altro, indice di una variazione semasiologica attiva al livello dell'articolazione in famiglie di sensi.

Tabella 7. Classificazione in accezioni di *disinteresse*

	Architettura	Religione	Informazione
A 'indifferenza'	93	37	90
B 'generosità'	4	57	9
? (casi dubbi)	3	6	1

Per quanto riguarda l'analisi delle 290 concordanze non-dubbe, non sono emerse particolari differenze fra le caratteristiche d'uso assunte dall'accezione A in *Architettura e Design* e in *Informazione*, e per questo non sono state evidenziate le frequenze d'occorrenza di date caratteristiche d'uso relative ai singoli domini. È emersa, peraltro, una distinzione abbastanza netta fra le due accezioni assunte dal lemma. In breve: l'accezione A è definita da una tendenza generale sia ad esprimere chi manifesti assenza di interesse, sia cosa ne sia oggetto, e ad essere modificata da aggettivi, principalmente in posizione postnominale; l'accezione B, invece, tende ad essere usata in maniera "assoluta", cioè senza modificazione aggettivale e senza espressione di soggetti o oggetti, ma si trova molto frequentemente ad essere coordinata con altri sostantivi astratti. Per caratterizzare un po' meglio le due accezioni, è interessante notare che: (a) i soggetti di *disinteresse* A nella maggioranza dei casi si riferiscono a entità appartenenti agli ambiti lessicali della politica o della burocrazia; (b) nel campo della modificazione aggettivale, per quanto essa sia poco frequente per B, vi è una contrapposizione sulla dimensione semantica della "diffusione": mentre per A si parla di *disinteresse generale / diffuso / collettivo / comune*, per B esso è soprattutto *personale*, cioè riguarda il singolo e non una frazione della collettività; (c) le preposizioni che occorrono anche con bassa frequenza in posizione -1 sembrano indicare che, mentre *disinteresse* A può essere una causa, *disinteresse* B è piuttosto una maniera dell'azione.

Tabella 8. Caratteristiche d'uso di *disinteresse*.

	A 'indifferenza'	B 'generosità'
espressione del soggetto:	54 su 220	4 su 70
<i>disinteresse da parte di / de*</i>	10 su 54	/
<i>disinteresse di / de*</i>	40 su 54	4 su 4
<i>disinteresse mostrato / dimostrato da*</i>	4 su 54	/
sostantivi attinenti alla sfera lessicale della politica / vita civile:	29 su 54 (es.: amministrazioni (5), governo (3), media (3), politici (3), amministratori (2), partiti (2), autorità, giunta, potere)	/
espressione dell'oggetto:	91 su 220	2 su 70
<i>disinteresse per</i>	52 su 91	1 su 2
<i>disinteresse verso</i>	20 su 91	/
<i>disinteresse nei confronti di</i>	11 su 91	/
<i>disinteresse nei riguardi di</i>	3 su 91	1 su 2
<i>disinteresse a</i>	5 su 91	/
modificazione aggettivale:	74 su 220	8 su 70
preinomiale:	52 su 74	4 su 8
a) specificazione 'intensità'	<i>totale (20), assoluto (6), completo (4), certo (3), crescente, forte, profondo, estremo</i>	<i>totale (2), profondo</i>
b) specificazione 'diffusione'	<i>generale (2), diffuso, largo</i>	<i>personale</i>
c) specificazione 'visibilità'	<i>palese (3), apparente (2), aperto, evidente</i>	/
d) specificazione 'valutazione'	<i>abituale, disarmante, imbarazzante, inaccettabile</i>	/
postnominale:	22 su 74	4 su 8
a) specificazione 'intensità'	<i>totale (7), assoluto (2), profondo</i>	/
b) specificazione 'diffusione'	<i>generale (4) collettivo (3), comune, diffuso</i>	<i>personale (4)</i>
c) specificazione 'visibilità'	<i>evidente</i>	/
d) specificazione 'valutazione'	<i>mortificante, sorprendente</i>	/
colligazione con preposizioni:		
<i>per disinteresse</i>	7 su 220	/
<i>con disinteresse</i>	1 su 220	17 su 70
coordinazione con altri nomi astratti (es.: offrire lealtà, rispetto, cordialità e disinteresse)	26 su 220	49 su 70

In sintesi, la variazione semasiologica relativa a questo sostantivo prende la forma soprattutto di un differente peso strutturale assunto dalle accezioni A e B fra le varietà di lingua di cui i domini *Religione* e *Architettura e Design* possono essere assunti come campioni.

Data l'assenza di comportamenti difformi all'interno dell'accezione A in *Architettura e Design* e *Informazione*, possiamo ritenere che il lessema possieda una struttura semasiologica molto simile nei due domini: non è emersa, cioè, una divisione in sotto-accezioni o la presenza di sensi salienti specifici di una data varietà di sottocodice.

A livello più generale può essere interessante notare che la “assenza di cura” espressa dall'accezione A pare propria soprattutto dei personaggi e delle istituzioni della vita politica e sociale, e che al suo interno spicca per frequenza d'occorrenza il sintagma *totale disinteresse / disinteresse totale*, il che potrebbe essere ricondotto ad alcune caratteristiche del discorso e delle pratiche politiche proprie della società italiana contemporanea. La forte dominanza dell'accezione A in *Informazione* e in *Architettura e Design* potrebbe essere riconducibile proprio al peso che i discorsi tipici della lingua dell'architettura e della lingua dei giornali danno a tematiche come l'assenza di interesse da parte degli enti amministrativi e pubblici nei confronti di qualcosa; il discorso religioso, invece, per sua natura più propenso a ruotare attorno a tematiche relative all'etica e alla morale del singolo, tende a privilegiare l'uso dell'accezione B.

3.3 *Responsabile*

Le definizioni di lavoro adottate per l'analisi dell'aggettivo *responsabile* sono le seguenti:

- A. “Che risponde delle proprie azioni e dei propri comportamenti”
- B. “Che, nel comportamento, rivela le qualità psichiche e l'equilibrio propri di chi è in grado di rispondere consapevolmente delle proprie azioni”
- C. “Che, chi si è reso colpevole di un'azione, o dà origine a un processo negativo”

I domini di RIDIRE su cui si è svolta l'analisi, invece, sono:

- (a) *Economia e affari*
- (b) *Religione*
- (c) *Informazione*

Esaminando i collocati statistici di *responsabile* possiamo notare che emerge un solo lemma comune a tutti e tre i domini, cioè *socialmente*, il che potrebbe farci pensare che l'uso del lessema sia piuttosto differenziato. Fra *Economia e affari* e *Informazione* sono condivisi anche due termini che si possono ricondurre all'accezione A, *funzionario* e *dirigente*.

Tabella 9. Collocati statistici di *responsabile*.

Religione		Economia e Affari		Informazione	
LEMMA	Dice	LEMMA	Dice	LEMMA	Dice
1 procreazione	0,0141	socialmente	0,0657	coerente	0,0282
2 laicato	0,0075	archiviazione	0,0225	conferimento	0,0235
3 cosciente	0,0050	collocamento	0,0183	chiaro	0,0087
4 maturo	0,0048	dirigente	0,0137	redazione	0,0069
5 socialmente	0,0047	indebitato	0,0119	disponibile	0,0068
6 moralmente	0,0044	funzionario	0,0102	Ds	0,0064
7 generoso	0,0034	e-mail	0,0100	dirigente	0,0059
8 obbedienza	0,0027	longevo	0,0071	gene	0,0055
9 pastore	0,0025	Stefania	0,0066	virus	0,0047
10 edificazione	0,0019	interessato	0,0063	attentato	0,0046
11 parrocchia	0,0018	qualunque	0,0052	marketing	0,0043
12 battezzato	0,0018	struttura	0,0050	funzionario	0,0041
13 coniugale	0,0018	organismo	0,0045	Margherita	0,0037
14 enciclica	0,0017	procedimento	0,0039	uccisione	0,0035
15 laico	0,0017	civilmente	0,0036	batterio	0,0032
16 consapevole	0,0016	esecuzione	0,0035	welfare	0,0030
17 onesto	0,0016	mediatore	0,0035	socialmente	0,0029
18 solidale	0,0016	Euromobiliare	0,0035	presso	0,0029
19 libero	0,0015	Imi	0,0034	vice	0,0028
20 evangelizzazione	0,0015	turismo	0,0032	dicastero	0,0028

Scorrendo gli altri collocati, possiamo notare come in *Economia e affari* siano presenti quattro termini che sembrano collegati ad un uso giuridico del lessema (*procedimento*, *civilmente*, *struttura* e *organismo*) e due nomi di organizzazioni/aziende, (*Euromobiliare* e *Imi*); in *Religione*, invece, una buona parte dei collocati fa parte del lessico proprio di tale varietà di sottocodice (*laicato*, *pastore*, *parrocchia*, *battezzato*, *enciclica*, *laico* e *evangelizzazione*), il che farebbe pensare che *responsabile* faccia parte del vocabolario tipico di *Religione*, mentre un'altra parte è costituita da aggettivi di qualità con valore positivo (*cosciente*, *generoso*, *maturo*, *consapevole*, *onesto*, *solidale* e *libero*); in *Informazione*, invece, fra i collocati troviamo quattro lemmi relativi alla vita politica (*Ds*, *Margherita*, *welfare*

e *dicastero*) e cinque altri collocati che paiono indicare cause o eventi negativi (*gene, virus, attentato, uccisione e batterio*).

Dalla classificazione in accezioni emerge come A sia di gran lunga la più attestata in *Economia e affari*, mentre sia mediamente frequente in *Religione* e addirittura scarsa in *Informazione*, il che smentisce parzialmente quanto emerso dallo scorrimento dei collocati statistici. L'accezione B, invece, è molto frequente in *Religione* e mediamente attestata in *Informazione*, il che di nuovo non pareva prevedibile dall'analisi dei collocati. Per quanto riguarda C, infine, si tratta dell'accezione dominante in *Informazione*. In generale, dunque, sembra che *responsabile* presenti una chiara variazione semasiologica nella frequenza d'uso delle diverse accezioni nei tre domini esaminati, dato che in ognuno di essi l'accezione più frequente è diversa.

Tabella 10. Classificazione in accezioni di *responsabile*.

	Economia e Affari	Religione	Informazione
A 'chi risponde'	71	30	19
B 'maturo'	19	62	34
C 'colpevole'	6	5	43
? (casi dubbi)	4	3	4

Dall'analisi delle 289 concordanze non-dubbe possiamo notare che ognuna delle tre accezioni pare definita da alcune tendenze distintive: per quanto riguarda A si può citare quella ad esprimere esplicitamente l'oggetto della responsabilità, l'uso in proporzioni quasi uguali o in funzione predicativa o di modificazione postnominale, la tendenza a modificare sostantivi che si riferiscono principalmente ad umani, ma anche a entità collettive o astratte, la frequenza bassa ma significativa con cui l'aggettivo è modificato da avverbi di maniera o dominio; per B, invece, notiamo l'assenza completa di espressione dell'oggetto, l'uso aggettivale soprattutto in posizione postnominale, ma anche, caso unico, prenominale, la forte tendenza a modificare sostantivi astratti, l'uso in locuzioni con valore avverbiale, la modificazione sia attraverso avverbi di maniera o dominio che attraverso avverbi di quantità (che fa ritenere che in questa accezione *responsabile* possa essere inteso anche come una proprietà scalare), la frequente coordinazione con altri aggettivi qualificativi; per l'accezione C, infine, possiamo notare la tendenza dominante all'espressione dell'oggetto, l'uso in funzione predicativa o di modificazione postnominale in proporzioni analoghe a quelle di A, l'assenza di casi di modificazione avverbiale e la tendenza dominante a modificare sostantivi che si riferiscono ad entità connotate negativamente.

Tabella 11. Caratteristiche d'uso di responsabile. *concordanze in cui era individuabile il sostantivo modificato; **sono i 9 casi di modificazione prenominale.

	A 'che risponde'	B 'maturo'	C 'colpevole'
espressione dell'oggetto:	99 su 120	/	49 su 54
<i>responsabile di / de*</i>	85 su 99	/	46 su 49
<i>responsabile per</i>	11 su 99	/	3 su 49
<i>responsabile verso</i>	2 su 99	/	/
<i>responsabile nei confronti di</i>	1 su 99	/	/
uso aggettivale:			
predicativo	55 su 120	14 su 115	23 su 54
preinomiale	/	9 su 115	/
postnominale	65 su 120	92 su 115	31 su 54
caratteristiche semantiche del sostantivo modificato:			
sostantivi rintracciati*:	94 su 120	89 su 115	43 su 54
umani	46 su 94	9 su 89	21 su 43
entità collettive	24 su 94	18 su 89	8 su 43
astratti	24 su 94	9** + 53 su 89	14 su 43
connotazione negativa del sostantivo modificato:	/	/	36 su 43 (es.: <i>danni, clima di terrore, attentati</i>)
uso in locuzioni avverbiali:	/	12 su 115 (es.: <i>in modo r.</i>)	/
modificazione dell'aggettivo con avverbio in -mente:	15 su 120	12 su 115	/
tipologia degli avverbi:	di maniera o dominio (es.: <i>solidalmente, personalmente</i>)	di maniera o dominio (9 occ. di <i>socialmente r.</i> solo in 'economia'), di quantità (<i>maggiormente r., sommamente r., particolarmente r.</i> ; + 10 occ. di <i>più r.</i>)	/
coordinazione con altri aggettivi:	/	39 su 115 (es.: <i>un impegno responsabile e convergente, un confronto vero e responsabile</i>)	/

Per quanto riguarda la presenza di fenomeni di variazione semasiologica al livello delle sotto-accezioni o dei singoli sensi, l'analisi delle concordanze non ha evidenziato casi significativi, a parte l'occorrenza frequente della locuzione *socialmente responsabile* nel solo dominio *Economia e affari*, il che spingerebbe a caratterizzare questo uso specifico come tipico della lingua dell'economia. A livello sociolinguistico, quindi, possiamo ritenere che l'accezione A sia più tipica della lingua dell'economia, quella B della lingua della religione, mentre l'accezione C caratterizzi lessicalmente la modalità d'uso della lingua dei giornali.

Vale la pena di osservare, infine, come l'analisi delle caratteristiche d'uso ci possa restituire un'immagine abbozzata dell'organizzazione della struttura prototipica del lessema⁹: gli usi più caratteristici di A e C, da un lato, e di B, dall'altro, sono piuttosto differenti, il che potrebbe essere interpretato, a livello di teoria dei prototipi, come il segno del fatto che tali accezioni si collocano ad una certa distanza concettuale le une dalle altre; fra A e C, invece, le differenze d'uso sono molto ridotte (consistono solo nell'assenza di modificazione avverbiale e nell'uso preponderante con sostantivi negativi per C), il che può essere interpretato in termini di una bassa distanza concettuale fra le due accezioni, o addirittura come segno del fatto che C corrisponde ad una sotto-accezione saliente di A non pienamente autonoma.

3.4 *Spaventoso*

Le definizioni di lavoro adottate per l'analisi di *spaventoso* sono le seguenti:

- A. "Che incute timore, spavento, in quanto avvertito come minaccioso, pericoloso o dannoso, che suscita istintivamente paura"
- B. "Che desta profonda impressione per la sua tragicità e gravità"
- C. "Eccezionale, straordinario, al di là d'ogni immaginazione"

I domini di RIDIRE presi in considerazione, invece, sono:

- a. *Letteratura e Teatro*
- b. *Sport*
- c. *Informazione*

Per quanto riguarda l'analisi dei collocati statistici di *spaventoso*, possiamo notare che la lista di *Sport* si ferma alla decima posizione, e che fra i lemmi presenti i

⁹ Sulla possibilità di interpretare le caratteristiche d'uso in questi termini, cfr. Gries 2012, 65; Glynn 2013, 4.

valori Dice sono piuttosto bassi: il fatto è dovuto sostanzialmente allo scarso numero di occorrenze totali del lessema in questo dominio. Tenuto conto di ciò, possiamo osservare come non ci sia alcun lemma comune a tutti e tre i domini: solamente *incidente* compare in più di una lista, cioè in *Sport* e in *Informazione*. Si può ipotizzare, quindi, che l'uso del lessema sia piuttosto differente nei tre sub-corpora esaminati.

Tabella 12. Collocati statistici di *spaventoso*.

Letteratura e Teatro		Sport		Informazione	
LEMMA	Dice	LEMMA	Dice	LEMMA	Dice
1 chiaroveggenza	0,0101	incidente	0,0008	massacro	0,0039
2 favilla	0,0044	caduta	0,0006	sparatoria	0,0019
3 selva	0,0033	talento	0,0002	mole	0,0017
4 spaventoso	0,0027	cifra	0,0002	terroristico	0,0013
5 osceno	0,0020	adesso	0,0001	cocaina	0,0013
6 orribile	0,0013	davvero	0,0001	shock	0,0012
7 abisso	0,0012	popolo	0,0001	incidente	0,0010
8 catastrofe	0,0012	pagare	0,0001	attentato	0,0009
9 cupo	0,0012	succedere	0,0001	cifra	0,0008
10 urlo	0,0010	giocatore	0,0001	disastro	0,0008
11 miseria	0,0010			incendio	0,0007
12 incomprensibile	0,0010			tragedia	0,0006
13 tempesta	0,0010			fame	0,0006
14 solitudine	0,0009			esplosione	0,0005
15 terribile	0,0009			scoppiare	0,0005
16 bestia	0,0009			montagna	0,0004
17 mostro	0,0008			provocare	0,0004
18 incubo	0,0008			violenza	0,0003
19 ignoto	0,0007			vuoto	0,0002
20 inquietante	0,0007			scenario	0,0002

Se osserviamo i collocati presenti in *Letteratura e Teatro*, infatti, sono presenti sostantivi come *selva*, *abisso*, *urlo*, *mostro* e *incubo*, e aggettivi come *osceno*, *orribile*, *cupo*, *terribile*, *ignoto* e *inquietante*, che paiono richiamare l'accezione A. In *Sport*, invece, troviamo sostantivi come *incidente* e *caduta*, che paiono richiamare B, e *talento*, *cifra* e *giocatore*, che forse possono essere ricondotti a C. In *Informazione*, infine, troviamo in maggioranza sostantivi e aggettivi che paiono richiamare l'accezione B, come *massacro*, *sparatoria*, *terroristico*, *incidente*, alcuni sostantivi legati probabilmente a C, come *mole*, *cifra*, *fame*, e anche alcuni lemmi che potrebbero richiamare A, come *violenza* e *scenario*.

Passando all'analisi delle concordanze, le frequenze d'uso delle tre accezioni risultate dalla classificazione manuale delle occorrenze ci dicono che in *Letteratura e Teatro* l'accezione dominante è A, in *Sport* è C, mentre in *Informazione* è B, il che è abbastanza in linea con quanto si poteva ipotizzare dalla lettura dei collocati statistici. La proporzione fra le frequenze d'uso delle diverse accezioni, tuttavia, non coincide con l'immagine restituita dai collocati statistici: la proporzione di A in *Letteratura e Teatro* e quella di B in *Informazione* non sono così elevate come poteva apparire dal numero di lemmi riconducibili a queste accezioni presenti nelle liste dei collocati, mentre la proporzione di C in *Sport*, al contrario, è decisamente più alta di quanto si poteva immaginare. Dalle frequenze d'occorrenza pare comunque che *spaventoso* presenti una sensibile variazione semasiologica sul piano delle accezioni dominanti.

Tabella 13. Classificazione in accezioni di *spaventoso*.

	Letteratura	Sport	Informazione
A 'che fa paura'	46	6	19
B 'tragico'	22	30	45
C 'straordinario'	26	54	30
? (casi dubbi)	6	10	6

Per quanto riguarda l'analisi delle caratteristiche d'uso emerse dallo scorrimento delle 278 concordanze non-dubbe sembra che ciascuna delle tre accezioni presenti alcuni pattern d'uso, a livello sintattico o semantico, che la distinguono dalle altre due. Per A possiamo notare l'uso abbastanza frequente in funzione predicativa, la generale tendenza ad occorrere in posizione postnominale, la tendenza a comparire in coordinazione con altri aggettivi di qualità, la tendenza ad essere modificata attraverso avverbi di quantità. Per B, invece, è determinante la decisa tendenza a occorrere in posizione prenominal e la forte preferenza semantica che accomuna i sostantivi modificati, di cui due su tre indicano eventi o proprietà fortemente negative. Per C, infine, possiamo notare la preferenza per la posizione postnominale e la forte tendenza a modificare sostantivi che indicano entità o proprietà misurabili, che ha luogo praticamente in un caso su due. Non sono state individuate caratteristiche d'uso di una singola accezione che presentino una variazione significativa rispetto al dominio d'uso, quindi si può ritenere che la variazione semasiologica a cui è soggetto *spaventoso* consista soprattutto nella differente frequenza con cui viene utilizzata un'accezione in una data varietà di sottocodice.

Tabella 14. Caratteristiche d'uso di *spaventoso*.

	A 'che fa paura'	B 'tragico'	C 'straordinario'
uso aggettivale:			
sostantivi rintracciati	67 su 71	95 su 97	108 su 110
predicativo	17 su 67	5 su 95	7 su 108
preinomiale	16 su 67	57 su 95	28 su 108
postnominale	29 su 67	33 su 95	73 su 108
altro	5 su 67 (5 occ. della locuzione <i>qualcosa di spaventoso</i>)	/	/
coordinazione con altri aggettivi:	17 su 71 (es.: <i>la morte non sembrò più così spaventosa e terribile</i>)	1 su 97	1 su 110
modificazione avverbiale di quantità o giudizio: (es.: <i>molto spaventoso; certamente spaventoso</i>)	17 su 71	5 su 97	6 su 110
caratteristiche semantiche dei sostantivi modificati: (totale rispetto ai sostantivi rintracciati)	assenza di pattern rilevanti	entità astratte o inanimate connotate negativamente: 62 su 95 (es.: <i>atrocità, battaglia, carneficina, attentato, strage</i>)	entità o proprietà misurabili: 52 su 108 (es.: <i>debito, calo demografico, temperature, velocità, tempo</i>), fra cui 15 di sostantivi denotanti 'numeri': <i>numero</i> (7), <i>cifra</i> (6), <i>somme, percentuali</i>

In quest'ottica, si può ritenere che la preminenza di B in *Informazione*, così come quella di C in *Sport* abbiano a che fare con le caratteristiche dei discorsi tipici di queste due varietà/modalità d'uso: data l'importanza dei temi d'attualità e il ruolo saliente che hanno gli eventi luttuosi o violenti nei media scritti, è comprensibile che B risulti l'accezione più diffusa nella lingua dei giornali; è caratteristica degli sport agonistici, invece, l'attenzione per il dato quantitativo tramite cui si misurano le performance, e quindi, data la tendenza di C a modificare sostantivi "misurabili", anche in questo caso sono le caratteristiche intrinseche della varietà di lingua a spiegare la dominanza di questa accezione nel relativo dominio d'uso.

4. Conclusioni

Basandoci sull'ipotesi che le caratteristiche d'uso contestuale di una parola possono essere assunte come indici della sua struttura concettuale, e tenendo conto della possibilità di intendere i sub-corpora di RIDIRE come campioni rappresentativi di lingue speciali in senso ampio, abbiamo considerato la manifestazione di caratteristiche d'uso differenti da parte di *clandestinità*, *disinteresse*, *responsabile* e *spaventoso* all'interno di sub-corpora distinti come l'indicazione della presenza di variazioni semasiologiche di sottocodice. Mettendo in relazione la frequenza con cui un dato lessema veniva usato secondo un'accezione o un'altra in un dato sub-corpus con le altre caratteristiche d'uso frequenti che caratterizzavano le sue occorrenze in quel dominio d'uso, abbiamo potuto esaminare differenti livelli organizzativi della struttura semantica di un lessema: sia il piano dell'articolazione in accezioni, che quello dei singoli sensi. Nel complesso, il tipo di variazione più evidente e frequente che abbiamo osservato ha riguardato il diverso tasso di frequenza con cui un'accezione è usata in una data varietà di sottocodice; gli altri due tipi di variazione previsti dal modello di Geeraerts 1997 sono stati individuati solo sporadicamente. Va detto, tuttavia, che la maggior diffusione di questo tipo di variazione potrebbe anche essere dovuta alla particolare procedura utilizzata per individuarla, differente da quella adottata per gli altri due, e quindi non è affatto detto che la sua maggior diffusione ed evidenza in questo studio costituisca una caratteristica propria della variazione semasiologica di sottocodice o della variazione semasiologica *tout court*.

L'analisi dei pattern semantici e sintattici che caratterizzano l'uso di un certo lessema ha permesso di mettere a fuoco in maniera abbastanza chiara le caratteristiche che contraddistinguono date accezioni e dati sotto-gruppi di sensi all'interno di una medesima accezione; inoltre, ha permesso di individuare comportamenti sintattici o semantici sistematici non pienamente prevedibili a priori: nel caso di *spaventoso* nell'accezione C ("straordinario"), ad esempio, abbiamo potuto notare come sia presente un nutrito gruppo di occorrenze in cui l'aggettivo modifica sostantivi indicanti entità misurabili. Da un altro punto di vista, invece, l'analisi dei pattern d'uso sembra permettere anche un'intuitiva visione generale dell'organizzazione della struttura polisemica dei lessemi esaminati, come abbiamo visto nel caso di *responsabile*.

I casi di studio esaminati illustrano in quale modo la struttura e gli attributi concettuali che definiscono il significato di un lessema siano suscettibili di variare a seconda dell'ambito discorsivo in cui questo è usato. Questo genere di variazione, rispetto a lessemi appartenenti al vocabolario corrente come quelli considerati, sembra poter possedere un carattere sottile e probabilistico che non è detto

sia percepibile introspektivamente, ma che diventa osservabile in maniera più distinta con l'adozione di un metodo analitico quantitativo basato sull'uso di corpora. Ci pare, dunque, che questo tipo di approccio possa costituire un mezzo promettente per iniziare ad esplorare la distribuzione delle disomogeneità semantiche che caratterizzano una comunità linguistica.

Riferimenti

- Baker, P. 2010. *Sociolinguistics and Corpus Linguistics*. Edinburgh: Edinburgh University Press.
- Berruto, G. 2006. *Fondamenti di sociolinguistica*. Roma: Laterza.
- Berruto, G. 2010. Semantica strutturale e variazione, una rivisitazione. *Incontri Linguistici* 33: 11-31.
- Berruto, G. 2011. Variazione linguistica. In R. Simone (ed.), *Enciclopedia dell'italiano*. Roma: Istituto dell'Enciclopedia italiana, 1547-1550.
- Berruto, G. 2012. *Sociolinguistica dell'italiano contemporaneo. Nuova edizione*. Roma: Carocci.
- Bonomi, I. 2010. Giornali, lingua dei. In R. Simone (ed.), *Enciclopedia dell'italiano*. Roma: Istituto dell'Enciclopedia italiana, 580-583.
- Casadei, F. 2003. *Lessico e semantica*. Roma: Carocci.
- Cerruti, M. 2011. Il concetto di variabile sociolinguistica a livello del lessico. *Studi Italiani di Linguistica Teorica e Applicata* 40(2): 211-31.
- Corbetta, P. 2003. *La ricerca sociale: metodologia e tecniche. Volume 2: Le tecniche quantitative*. Bologna: il Mulino.
- Cortelazzo, M. 2010. Giuridico-amministrativo, linguaggio. In R. Simone (ed.), *Enciclopedia dell'italiano*. Roma: Istituto dell'Enciclopedia italiana, 588-589.
- Cresti, E. & Moneglia, M. 2012. Risorse di rete per l'insegnamento dell'italiano. In P. Bianchi, N. De Blasi, C. De Caprio & F. Montuori (eds), *La variazione nell'italiano e nella sua storia. Varietà e varianti linguistiche e testuali*. Firenze: Cesati, 597-613.
- Cresti, E. & Panunzi, A. 2013. *Introduzione ai corpora dell'italiano*. Bologna: il Mulino.
- Croft, W. & Cruse, A.D. 2010. *Linguistica cognitiva*. Roma: Carocci.
- De Mauro, T. 1980. *Guida all'uso delle parole. Parlare e scrivere semplice e preciso per capire e farsi capire*. Roma: Editori Riuniti.
- Evert, S. 2009. Corpora and collocations. In A. Lüdeling & M. Kytö (eds), *Corpus linguistics: an international handbook. Volume 2*. Berlin/New York: De Gruyter, 1212-1248.
- Fabiszak, M. 2005. Semantic and Lexical Change. In U. Ammon, N. Dittmar, K.J. Mattheier & P. Trudgill (eds), *Sociolinguistics. An International Handbook of the Science of Language and Society. Vol. II*. Berlin/New York: De Gruyter.
- Firth, J.R. 1957. *Papers in Linguistics 1934-1951*. London: Oxford University Press.
- Geeraerts, D. 1997. *Diachronic Prototype Semantics. A contribution to historical lexicology*. Oxford: Clarendon Press.
- Geeraerts, D. 2008. Prototypes, stereotypes, and semantic norms. In G. Kristiansen & R. Dirven (eds), *Cognitive Sociolinguistics. Language Variation, Cultural Models, Social Systems*. Berlin/New York: Mouton de Gruyter, 21-44.

- Geeraerts, D. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Geeraerts, D., Grondelaers, S. & Bakema, P. 1994. *The Structure of Lexical Variation: Meaning, Naming, and Context*. Berlin/New York: Mouton de Gruyter.
- Glynn, D. 2013. The semantics of sociolinguistic variation. A quantitative study of dialect effects on the polysemy of *annoy*.
http://www.dsglynn.univ-paris8.fr/articles/Glynn_2013_Semantics_Socioling_Variation.pdf
 (accessed June 27, 2016).
- Glynn, D. 2014. Polysemy and synonymy. Cognitive theory and corpus method. In D. Glynn & J. Robinson (eds), *Corpus Methods for Semantics*. Amsterdam/Philadelphia: Benjamins, 7-38.
- Gries, S.T. 2006. Exploring variability within and between corpora: some methodological considerations. *Corpora* 1(2): 109-151.
- Gries, S.T. 2012. Behavioral profiles. A fine-grained and quantitative approach in corpus-based lexical semantics. In G. Libben, G. Jarema & C. Westbury (eds), *Methodological and Analytic Frontiers in Lexical Research*. Amsterdam/Philadelphia: John Benjamins, 57-80.
- Hasan, R. 2009. On semantic variation. In J.J. Webster (ed.), *The collected works of Ruqaiya Hasan. Vol. II. Semantic variation. Meaning in society and in sociolinguistics*. London-Oakville: Equinox, 41-71.
- Ježek, E. 2011. *Lessico: classi di parole, strutture, combinazioni*. Bologna: il Mulino.
- Labov, W. 1978. Where does the sociolinguistic variable stop? A response to Beatriz Lavendera. In *Working Papers in Sociolinguistics*. Austin: Southwest Educational Development Laboratories.
- Labov, W. 2008. Quantitative Reasoning in Linguistics.
<http://www.ling.upenn.edu/~wlabov/Papers/QRL.pdf> (accessed June 27, 2016).
- Lenci, A. 2010. Modelli distribuzionali del lessico: metodi computazionali per l'analisi semantica. *Informatica Umanistica* 3: 57-69.
- Lewandowska-Tomaszczyk, B. 2007. Polysemy, Prototypes, and Radial Categories. In D. Geeraerts & H. Cuyckens (eds), *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, 139-169.
- McEnery, T. & Hardie, A. 2012. *Corpus Linguistics*. Cambridge: Cambridge University Press.
- Moneglia, M., Colombo, S., Panunzi, A. 2014. Una Guida alle funzioni di ricerca linguistica nel *web corpus RIDIRE*. In E. Garavelli & E. Suomela-Härmä (eds), *Dal manoscritto al web: canali e modalità di trasmissione dell'italiano. Tecniche, materiali e usi nella storia della lingua*. Firenze: Cesati, 481-502.
- Moneglia, M. & Paladini, S. 2010. Le risorse di rete dell'italiano. Presentazione del progetto "RIDIRE.it". In E. Cresti, I. Korzen (eds), *Language, Cognition and Identity. Extension of the Endocentric/Esocentric Typology*. Firenze: Firenze University Press, 111-128.
- Murphy, G.L. 2002. *The Big Book of Concepts*. Cambridge, MA: MIT Press.
- Nerlich, B. 2003. Polysemy: past and present. In B. Nerlich, Z. Todd, V. Herman & D.D. Clarke (eds), *Polysemy. Flexible Patterns of Meaning in Mind and Language*. Berlin/New York: Mouton de Gruyter, 49-76.
- Panunzi, A., Cresti, E. & Gregori, L. 2014. RIDIRE. Corpus and Tools for the Acquisition of Italian L2. In A. Abel, C. Vettori & N. Ralli (eds), *Proceedings of the XVI EURALEX International Congress: The User in Focus*. Bolzano: EURAC research, 447-462.
- Peirsman, Y., Heylen, K. & Geeraerts, D. 2010. Applying word space models to sociolinguistics. Religion names before and after 9/11. In D. Geeraerts, G. Kristiansen & Y. Peirsman

- (eds), *Advances in Cognitive Sociolinguistics*. Berlin/New York: De Gruyter Mouton, 111-137.
- Proietti, D. 2010. Economia, lingua dell'. In R. Simone (ed.), *Enciclopedia dell'italiano*. Roma: Istituto dell'Enciclopedia italiana, 411-413.
- RIDIRE. <http://www.ridire.it/it.drwolf.ridire/home.seam> (accessed July 27, 2016).
- Robinson, J. 2010. Awesome insights into semantic variation. In D. Geeraerts, G. Kristiansen & Y. Peirsman (eds), *Advances in Cognitive Sociolinguistics*. Berlin/New York: De Gruyter Mouton, 86-109.
- Rovere, G. 2011. Variazione di sottocodice. In M. Cerruti, E. Corino & O. Cristina (eds), *Formale e informale. La variazione di registro nella comunicazione elettronica*. Roma: Carocci, 36-56.
- Rychlý, P. 2008. A Lexicographer-Friendly Association Score. In P. Sojka & A. Horák (eds), *RASLAN 2008: Recent Advances in Slavonic Natural Language Processing*. Brno: Masaryk University, 6-9.
- Schmid, H.J. 2007. The Notion of *Entrenchment* and *Saliency* in Cognitive Linguistics. In D. Geeraerts & H. Cuyckens (eds), *The Oxford Handbook of Cognitive Linguistics*. Oxford: Oxford University Press, 118-138.
- Sinclair, J. 2004. *Trust the Text. Language, corpus and the discourse*. London: Routledge.
- Stubbs, M. 2002. *Words and phrases: corpus studies of lexical semantics*. Oxford/Malden: Blackwell.
- Tagliamonte, S. 2006. *Analysing Sociolinguistic Variation*. Cambridge: Cambridge University Press.
- Tavosanis, M. 2011. *L'italiano del web*. Roma: Carocci.
- Taylor, J.R. 1995. *Linguistic Categorization. Prototypes in Linguistic Theory*. Oxford: Clarendon Press.
- Wulff, S., Stefanowitsch, A. & Gries, S.T. 2007. Brutal Brits and persuasive Americans. Variety-specific meaning construction in the into-causative. In G. Radden, K.M. Köpcke, T. Berg & P. Siemund (eds), *Aspects of Meaning Construction*. Amsterdam/Philadelphia: John Benjamins.