

LA POLISEMIA NEL VOCABOLARIO DI BASE DELL'ITALIANO

FEDERICA CASADEI
UNIVERSITÀ DELLA TUSCIA

Abstract – The existence of a relationship between frequency and polysemy in words is well known in statistical linguistics. Nevertheless, few data are available on the amount of polysemy in high-frequency words across languages, and no systematic investigation of the Italian lexicon has been carried out so far. This paper presents a detailed analysis of the number and types of senses of the about 7,000 highest frequency lexemes that constitute the basic vocabulary in Italian (vocabolario di base, VDB). Data confirm that frequency and polysemy are strongly related: the percentage of polysemic words in VDB (89%) is much greater than in the overall lexicon (19%); and within VDB, too, changes in frequency correlate with growth in polysemy. Verbs are found to be more polysemic than nouns, in the sense that there are less monosemic verbs than nouns; however, polysemic verbs and nouns give rise to a similar amount of senses (average number of senses is 7.3 in verbs and 6.6 in nouns). The crucial difference between verbs and nouns is to be found in the type, more than in the number, of their senses. In both classes the largest number of senses belongs to the area of basic/common vocabulary, but this area covers 77% of verb senses vs. 55% of noun senses. On the contrary, in nouns 34% of senses (vs. 8% in verbs) belong to the area of technical-scientific vocabulary, and 67% of nouns have at least one technical-scientific sense.

Keywords: polysemy; senses; word frequency; statistical linguistics; vocabolario di base.

1. Quantificare la polisemia

1.1. La diffusione della polisemia nelle lingue

Uno dei pochi punti di convergenza dei molti e variegati approcci alla semantica lessicale è il riconoscimento della centralità della polisemia nel funzionamento del linguaggio verbale. Che la si consideri più ‘in negativo’, come fonte di ambiguità lessicale (e dei conseguenti problemi di disambiguazione in cui si imbattono gli elaboratori non umani), o più ‘in positivo’, come prodotto della proprietà semiotica cruciale delle lingue naturali – l’indeterminatezza semantica – o addirittura come un principio di organizzazione cognitiva che agisce non solo nel lessico ma anche nella morfologia e nella sintassi, non c’è dubbio che la capacità dei segni linguistici di articolare il significato in gruppi di sensi tra loro correlati costituisca uno degli aspetti più salienti della semantica delle lingue verbali.

Ma quanto è diffusa la polisemia dal punto di vista quantitativo? Quante sono, in una lingua, le parole polisemiche? La difficoltà nel rispondere a questa domanda sembra dipendere da un motivo di natura pratica, cioè dal fatto che per nessuna lingua disponiamo di un repertorio lessicale completo e che attestiamo definitivamente quanti dei lessemi in essa esistenti siano monosemici e quanti polisemici, e, nel secondo caso, quali e quante accezioni abbiano. Questa difficoltà pratica, però, non è che l’espressione di due questioni ben più insormontabili, sintetizzabili nel fatto che né i confini del lessico né quelli del significato dei lessemi sono (in linea teorica, non solo empirica) tracciabili.

L’impossibilità di circoscrivere il lessico entro un insieme definito è conseguenza

della sua intrinseca creatività, in virtù della quale la massa lessicale è non solo enorme di fatto – nell’ordine dei milioni di lessemi – ma in linea di principio illimitata. Da ciò deriva, nelle parole di De Mauro (1999a, p. VIII), che “la rappresentazione in forma di dizionario del lessico di una lingua non può non essere altro che forzatamente parziale”.

Inoltre (e soprattutto, parlando di polisemia) anche per le parole registrate nelle fonti lessicografiche non esiste alcun criterio che consenta di stabilire quali dei sensi che esse manifestano nell’uso linguistico abbiano un’autonomia semantica tale da farli ritenere accezioni distinte di quel lessema e non semplici casi di estensione del significato in relazione al particolare contesto d’enunciazione. Consideriamo ad esempio i diversi sensi che ha l’aggettivo *leggero* in contesti come *un farmaco leggero* (‘con pochi effetti collaterali’), *droga leggera* (‘che dà effetti meno gravi’), *un vino leggero* (‘poco alcolico’), *un colore leggero* (‘tenue’), *un trucco leggero* (‘che si nota poco’), *un sibilo leggero* (‘che si sente poco’); dobbiamo dire che ciascuno di questi sensi rappresenta un’accezione autonoma di *leggero* o li dobbiamo ricomprendere sotto un’unica accezione, semanticamente unitaria, formulabile come ‘che produce un effetto fisico-percettivo meno intenso’? Tra i semanticisti vi è sia chi sostiene la tesi ‘monosemista’ sia chi sostiene quella ‘polisemista’, ma pressoché tutti convergono nel ritenere che non esista un modo per dirimere la questione in modo conclusivo e che si debba postulare un *continuum* che vede a un estremo l’omonimia (più significati irrelati), all’altro estremo la vaghezza (vari sensi correlati a un significato unitario) e nel mezzo la polisemia (più significati correlati), senza però che si possa tracciare un confine netto tra queste tre possibili declinazioni della relazione tra forma e contenuto.

In realtà, dunque, rispondere alla domanda ‘Quante sono in una lingua le parole polisemiche?’ è, più che difficile, impossibile. Al massimo possiamo chiederci quante sono in una lingua le parole che i dizionari di quella lingua registrano come polisemiche. E anche accontentandoci per motivi operativi di far coincidere il lessico di una lingua con le (relativamente poche) parole registrate nei dizionari e nei database lessicali, resta il problema dovuto alla labilità del confine tra accezioni vere e proprie e usi estensivi, che ha tra le altre conseguenze la variabilità dei criteri in base ai quali queste fonti scelgono di definire e registrare la polisemia. Non è raro che una parola che per un dizionario è polisemica sia monosemica per un altro, ad esempio perché non ne viene registrata un’accezione obsoleta o metaforica; così il lessema *agricoltura* è monosemico per la maggior parte dei dizionari italiani ma è polisemico nel Gradit, dove ha una seconda accezione, di basso uso, ‘agraria’. E una differenza di senso che un dizionario decide di trattare come un caso di polisemia può essere trattata, in un altro dizionario, come una sfumatura non meritevole di dare luogo ad accezioni distinte del lessema in questione; ad esempio *agraria* nel Gradit ha due accezioni, una ‘complesso di discipline che studiano l’agricoltura’ e un’altra ‘facoltà universitaria dove si studiano tali discipline’, mentre nel vocabolario Treccani il lemma risulta monosemico ed è definito come ‘Complesso di scienze e di pratiche applicate alla coltivazione della terra; anche come facoltà universitaria, che raggruppa tutte le discipline (...) per la preparazione dei dottori in agraria’.

Più in generale i dizionari variano molto nella quantità di sensi che registrano, fermo restando che nessuno di essi è in grado di registrare tutti i sensi esistenti nell’uso linguistico e che, da questo punto di vista, tutti sottostimano la polisemia, in una percentuale che appare difficile da calcolare. Un dato a riguardo è indicato da Green (1989, trad. it. p. 74), la quale porta, a sostegno della tesi per cui ogni parola è virtualmente polisemica e a dimostrazione della diffusione degli usi estensivi, un’analisi di tre brani estratti da romanzi americani moderni, da cui risulta che una media del 15% dei

nomi, verbi e aggettivi che vi compaiono sono usati in sensi non contemplati dai dizionari. Commentando questo risultato Nagy (1995, p. 2) osserva che “Fifteen percent may not sound like a very large number, but if a new meaning were added to the dictionary for 15% of the words in a substantial amount of text, there would be a massive increase in the total number of meanings represented”; a suo avviso, il fatto che i dizionari non rispecchino la variabilità contestuale del significato è uno dei motivi per cui in essi la polisemia risulta sottostimata, oltre al fatto di includere un gran numero di vocaboli specialistici e/o di bassa frequenza che tendono a essere monosemici.

1.2. La diffusione della polisemia nelle fonti lessicografiche

Pur con queste premesse, e in mancanza di alternative, si può partire dai dizionari per ricavare qualche cifra generale sulla diffusione della polisemia, intesa in prima battuta come rapporto tra il numero di lemmi e il numero di significati registrati nelle fonti lessicografiche. Già questa prima valutazione non è semplice, poiché molti dizionari dichiarano il numero di lessemi registrati ma non il numero di accezioni. Nella Tabella 1 ho riassunto i dati ricavati da una ricognizione su alcuni dizionari di lingue europee, ovviamente prendendo in considerazione solo quelli che dichiarano il numero di accezioni: per l'italiano lo Zingarelli 2014, per il francese il Larousse nell'edizione elettronica, il Petit Robert 2014 e il Grand Robert in edizione elettronica, per lo spagnolo il dizionario della Real Academia Española sia nell'edizione del 1992 che in quella del 2001. Per l'inglese ho usato le statistiche fornite da Word Net, mentre i dati relativi al Webster sono tratti dal calcolo fatto a campione da Nagy (1995); ho incluso infine nella ricognizione anche i dati forniti da Cornetto, un database lessicale olandese costruito secondo gli stessi criteri di WordNet¹.

fonte	entrate	accezioni	rapporto
WordNet	155.287	206.941	1,3
Cornetto	92.000	118.000	1,3
Larousse ed. elettronica	90.000	135.000	1,5
DRAE 1992	83.014	154.480	1,8
DRAE 2001	88.431	161.962	1,8
Webster	267.000	600.000	2,2
Zingarelli 2014	144.000	380.000	2,6
Grand Robert CD-ROM	100.000	350.000	3,5
Petit Robert 2014	60.000	300.000	5

Tabella 1

Rapporto tra numero di entrate e numero di accezioni in varie fonti lessicografiche.

Se è quindi fuor di dubbio che il numero di significati superi di gran lunga quello dei lessemi, non è però chiaro in quale misura ciò avvenga, visto che a seconda della fonte considerata si va da meno del doppio fino a cinque volte tanto. Né è chiaro da cosa dipenda tale variabilità: apparentemente non è legata all'ampiezza del lemmario, dunque

¹ I dati relativi ai dizionari sono reperibili sui siti delle rispettive case editrici. Le statistiche di WordNet si trovano sul sito dell'università di Princeton, <http://wordnet.princeton.edu/>, mentre quelle del database Cornetto si trovano sul sito dell'università Vrije di Amsterdam, <http://www2.let.vu.nl/oz/ctl/cornetto/index.html>.

forse è legata ai criteri di compilazione del lemmario e in particolare alla quantità di termini tecnico-scientifici inclusi; essendo tendenzialmente monosemici, i tecnicismi contribuiscono a diminuire il rapporto tra numero di entrate e numero di accezioni. In ogni caso, escludendo i due valori più bassi e quello più alto, e facendo una media tra i restanti, risulta un rapporto di 2,2 accezioni per entrata – cioè un numero medio di 2,2 accezioni per lessema. Si può tenere fermo questo dato per avere un termine di confronto con quanto emergerà dall'analisi della polisemia nel vocabolario di base.

Dal solo dato relativo al numero di lemmi e di accezioni non è possibile calcolare quanti dei lessemi in questione siano monosemici e quanti polisemici – un'informazione che, a mia conoscenza, non è fornita da nessun dizionario. Un calcolo del genere può essere fatto, per l'inglese, a partire da WordNet. Nella versione 3.0 WordNet contiene 155.287 parole, l'83% delle quali risultano monosemiche. La quota di polisemia si attesta quindi al 17%, come già nelle precedenti versioni del database; Miller (1995, p. 40) riferiva infatti che “approximately 17% of the words in WordNet are polysemous”. La Tabella 2 riassume le percentuali di parole monosemiche e polisemiche presenti in WordNet, divise per categoria lessicale.

	numero	monosemici	polisemici
sostantivi	117.798	101.863 (86,5%)	15.935 (13,5%)
verbi	11.529	6.277 (54,5%)	5.252 (45,5%)
aggettivi	21.479	16.503 (77%)	4.976 (23%)
avverbi	4.481	3.748 (84%)	733 (16%)
tot	155.287	128.391 (83%)	26.896 (17%)

Tabella 2
Parole monosemiche e polisemiche in WordNet.

La percentuale di polisemia è dunque minima nei sostantivi – che risultano addirittura meno polisemici degli avverbi – e massima nei verbi, quasi metà dei quali risultano essere polisemici. E' proprio tra i verbi, inoltre, che si registrano i casi di maggiore polisemia: *cut* e *run* hanno in WordNet 41 accezioni, e *break* arriva a 59 (un dato, quest'ultimo, superato di gran lunga in altre fonti lessicografiche: stando alle statistiche dell'Oxford English Dictionary, nella seconda edizione del 1989 l'entrata più lunga del dizionario è il verbo *set* con “over 430 senses”, <http://public.oed.com/history-of-the-oed/dictionary-facts>).

I dati di WordNet sono sostanzialmente confermati da quelli ricavabili da fonti analoghe disponibili sia per l'inglese che per altre lingue. Per l'italiano, ad esempio, risulta che delle circa 260.000 parole registrate nel Gradit quelle polisemiche sono oltre 50.000, cioè il 19% (De Mauro 1999b, p. 1178).

Dunque se si considera il lessico di una lingua nel suo insieme la polisemia appare quantitativamente minoritaria rispetto alla monosemia. Questi conteggi però non tengono in considerazione la frequenza d'uso delle parole, che è invece la variabile più rilevante rispetto alla quale valutare la polisemia. Da tempo è noto infatti alla statistica linguistica che esiste una correlazione tra la frequenza di una parola e la sua polisemia, in base al principio della versatilità economica delle parole formulato da Zipf (1949) per cui le parole più frequenti risultano essere semanticamente più generiche e dunque più disponibili, rispetto alle parole di minor frequenza, a modularsi in un'ampia gamma di significati. E poiché le parole più frequenti sono quelle che compongono la stragrande maggioranza dei testi scritti e parlati prodotti in una lingua, ne consegue che la polisemia sarà anche minoritaria nel lessico, ma è tuttavia preponderante nell'uso linguistico. Come spiega De Mauro (2009):

Nelle rappresentazioni dizionariistiche dei lessici monolingui delle diverse lingue i lessemi polisemici sono una minoranza. La maggioranza dei lemmi si presenta con una sola accezione. Ma ciò non deve fare considerare la polisemia un fenomeno minoritario. [...] i lessemi polisemici, presentati come tali anche nei dizionari monolingui, minoranza nel lessico potenziale e, quindi, nel dizionario, sono di gran lunga i più frequenti nei testi. Una relazione precisa è nota alla statistica linguistica. Data una lista di frequenza o uso delle parole di una qualsiasi lingua, i lessemi più frequenti hanno in media un maggior numero di accezioni, più esattamente al crescere della frequenza secondo potenze cresce secondo la successione naturale il numero delle accezioni. Se parole con frequenza x hanno una accezione, parole con frequenza x^2 hanno due accezioni, tre accezioni quelle con frequenza x^3 , quattro accezioni quelle con frequenza x^4 e così via. Nei dizionari di frequenza di tutte le lingue un numero ristretto di lessemi, circa duemila, costituisce la testa delle liste di lessemi in ordine di frequenza decrescente e tali lessemi hanno una frequenza di gran lunga maggiore rispetto alle altre migliaia e decine di migliaia. Essi costituiscono ciò che si dice il *vocabolario fondamentale*. I lessemi del vocabolario fondamentale occupano mediamente tra l'80 e il 90% delle occorrenze dei lessemi nei testi. Ne consegue che necessariamente in ogni possibile testo almeno otto o nove parole su dieci sono occorrenze di lessemi polisemici.

La correlazione tra frequenza d'uso e polisemia individuata da Zipf ha trovato conferma nei decenni successivi in tutti gli studi di statistica linguistica, e varie analisi condotte su campioni di parole di alto uso mostrano che, in esse, la quantità di accezioni è molto superiore alla media generale. Per l'inglese, ad esempio, un'analisi di 100 sostantivi di alta frequenza ha mostrato che essi hanno in WordNet una media di 5,15 accezioni ciascuno e che solo 8 risultano monosemici (Leacock, Towell, Voorhees 1993, p. 260). Per l'italiano non è mai stata condotta in modo sistematico una verifica di questo genere, ed è appunto ciò che mi sono riproposta di fare analizzando la quantità e il tipo di polisemia presente nei lessemi del vocabolario di base.

2. Monosemia e polisemia nel vocabolario di base dell'italiano

Com'è noto, il vocabolario di base dell'italiano (d'ora in avanti VDB), presentato nella prima versione in appendice a De Mauro (1980), è costituito da circa 7.000 lessemi ripartiti in tre fasce di frequenza: il vocabolario fondamentale, formato dai circa 2.000 lessemi più usati in assoluto in italiano e la cui occorrenza costituisce il 90% del lessico di tutti i testi scritti e parlati; il vocabolario di alto uso, formato da altri 2.700 lessemi di alta, benché minore, frequenza; e il vocabolario di alta disponibilità, formato da circa 2.300 lessemi relativamente infrequenti nello scritto e nel parlato ma noti ai parlanti perché legati a oggetti e azioni di grande rilevanza nella vita quotidiana.

Il VDB si presenta come una lista e non fornisce indicazioni sul significato e sull'eventuale polisemia dei lessemi inclusi. E' però possibile ricostruire l'articolazione semantica dei lessemi del VDB partendo dal Gradit, poiché il dizionario riporta per ciascun lemma, e per ciascuna delle sue eventuali accezioni, una marca d'uso che ne segnala l'ambito di diffusione. Le marche d'uso usate dal Gradit possono essere ripartite in quattro classi:

- 1) vocabolario di base: marche FO (vocabolario fondamentale), AU (vocabolario di alto uso) e AD (vocabolario di alta disponibilità);
- 2) vocabolario comune (marca CO), costituito da circa 47.000 lessemi estranei al VDB ma che si ritengono noti a chiunque abbia un livello di istruzione medio-superiore;
- 3) vocabolario tecnico-specialistico (TS), costituito da circa 107.000 lessemi usati prevalentemente o solo in ambito scientifico, tecnologico o professionalmente settoriale;

4) ambito d'uso non tecnico-scientifico ma non comune: marche BU (basso uso), OB (obsoleto), LE (uso solo letterario), DI (uso dialettale), RE (uso regionale).

Nel seguito userò le sigle indicate qui sopra per riferirmi agli ambiti d'uso individuati dal Gradit.

Un'interrogazione del Gradit condotta sul CD-ROM 2007 restituisce 2.077 lemmi monorematici² marcati FO, 2.663 marcati AU e 1.988 marcati AD, per un totale di 6.728 lemmi appartenenti al vocabolario di base. Il supporto informatico non consente il recupero in automatico dei lemmi mono- o polisemici, dunque il calcolo deve essere fatto manualmente. La Tabella 3 riassume i risultati di questo conteggio:

	lessemi	monosemici	polisemici
FO	2.077	76 (4%)	2.001 (96%)
AU	2.663	223 (8%)	2.440 (92%)
AD	1.988	426 (21%)	1.562 (79%)
tot VDB	6.728	725 (11%)	6.003 (89%)

Tabella 3
Lessemi monosemici e polisemici nel VDB.

Alla luce di questi dati non è esagerato dire che la quasi totalità del VDB è costituita da lessemi polisemici: questi coprono quasi il 90% del VDB nel suo complesso, arrivando a costituire il 96% del vocabolario fondamentale. Risulta quindi pienamente confermata la relazione tra frequenza e polisemia dei lessemi; e ciò sia nel rapporto tra il VDB e l'insieme del lessico (nel quale, in base ai dati citati sopra, la quota totale di polisemia si attesta intorno al 17-19%), sia nel rapporto tra le tre fasce di frequenza all'interno del VDB stesso, dove la percentuale di lessemi polisemici risulta minima nel vocabolario di alta disponibilità e massima in quello fondamentale.

E c'è da aggiungere che questo conteggio in realtà sottostima la presenza di polisemia nel VDB, per motivi che dipendono sia dal modo in cui il Gradit ha costituito il proprio lemmario, sia dallo stile delle definizioni. Riguardo al primo aspetto, influisce notevolmente sulla quantificazione della mono-/polisemia il fatto che nel Gradit compaiano a lemma i femminili dei sostantivi animati, che spesso non hanno definizione ma un semplice rinvio secco al corrispettivo maschile (del tipo *adulta* → *adulto*) e che risultano perciò monosemici; e poiché in molti casi questi femminili hanno la qualifica d'uso AD, la loro presenza contribuisce ad aumentare di molto la quantità di monosemia nei sostantivi AD, che sono infatti, anche per questo motivo, la maggiore fonte di monosemia nell'intero VDB con 393 casi sui 725 registrati. Se espungessimo da questi 393 casi i circa 100 femminili con rinvio secco al maschile, la quota totale di monosemia nel VDB scenderebbe dall'11 al 9%.

Un caso analogo a quello dei femminili animati, anche se numericamente più trascurabile, è quello in cui il lessema in questione è sì monosemico, ma la sua definizione consiste nel rinvio a un altro lessema (del quale è variante grafica o sinonimo stretto) a sua volta polisemico: la preposizione *fra* risulta monosemica perché è definita semplicemente come variante di *tra*, ma poiché quest'ultima è polisemica dovremmo considerare polisemica, a rigore, anche *fra*; lo stesso per *difatti*, che rinvia a *infatti*, e per *sino*, che

² Il lemmario del Gradit include, oltre a circa 260.000 lemmi monorematici, anche circa 130.000 sottolemmi polirematici, cioè unità lessicali formate da più parole. Di questi ultimi non ho tenuto conto nell'indagine poiché le polirematiche non sono contemplate dal VDB e dunque nessuna di esse compare nel Gradit con la marca FO, AU o AD.

rimanda a *fino*. Assimilabili a questi casi sono quelli di *babbo*, geosinonimo di *padre* e, come quest'ultimo, portatore di un'accezione estensiva 'uomo che adempie alle funzioni di guida, consigliere o maestro' (*E' stato un babbo per me che, da bambino, un babbo non l'avevo mai avuto*³) e di *fifa*, sinonimo colloquiale di *paura* sia nel senso proprio (*avere paura/fifa del parto*) che in quello attenuato 'timore' (*avere paura/fifa che grandini*). Analogamente *sbronzato* è definito nel Gradit 'ubriaco' e di *ubriaco* condivide la polisemia (*sbronzato d'amore, di felicità*); *complotto* è definito 'congiura, cospirazione', entrambi lessemi con sensi estensivi; *camposanto* rinvia a *cimitero* e come quest'ultimo può avere sensi estensivi (*La scienza lavora instancabilmente a quel grande columbarium dei concetti che è il camposanto delle intuizioni* (F. Nietzsche); *Una macchinetta che altrimenti sarebbe già al camposanto dei computer*; *Spesso i play-off sono stati il camposanto dei dominatori della prima fase*); *ammalarsi* è definito 'diventare malato' e, come *malato*, può avere sensi estensivi o figurati (*ammalarsi di psicoterapia, di austerità, di disoccupazione, di Internet, di noia*). In due casi, infine, la definizione data dal Gradit contiene un'indicazione sulla polisemia del lessema pur senza esplicitarla: *successivo* è definito come 'che viene dopo in un ordine spaziale o temporale' (e che i due sensi spaziale e temporale corrispondano ad accezioni diverse lo dimostra il fatto che in *dopo* sono ben distinti) e *finanziario* è definito sia come 'relativo alle finanze' (*problemi finanziari*) sia come 'relativo alla finanza' (*rubrica finanziaria*).

Vi è poi un certo numero di lessemi – una cinquantina – per i quali i quali il Gradit non registra un senso, soprattutto estensivo o metaforico ma anche, più raramente, letterale, pure abbastanza consolidato. Di seguito qualche esempio: per *abbonarsi* non è registrato il senso 'fare l'abitudine, capitare ripetutamente', registrato invece per l'aggettivo *abbonato* (*Il Bologna si è abbonato a San Siro*; *Stephan si è abbonato alle partenze sprint*; *Da quando hanno iniziato il nido [i bambini] si sono abbonati a kg di catarro*); per *affettare* non sono indicati sensi estensivi e figurati, tra cui quello 'annientare, distruggere' registrato invece per *fare a fette* (*Ho l'ansia che mi affetta*; *Petteri è il miglior difensore ma Green lo ha affettato*) e quello iperbolico 'ammazzare' (*Se faccio tardi mia madre mi affetta*); per *convivere* non sono registrati usi estensivi (*Come convivere con la Germania*; *Bisogna abituarsi a convivere con i gabbiani in città*) e i sensi figurati 'avere familiarità, doversi abituare a' (*convivere con la paura, il diabete, il terremoto*) e 'coesistere' (*In lei convivono sentimenti opposti*; *Una musica in cui convivono jazz e pop*; *Le ambizioni di Renzi possono convivere con il governo di Letta?*); per *lessare* non è indicato il senso iperbolico (*Il caldo mi ha lessato il cervello*, *Queste scarpe lessano i piedi*) e quello figurato 'stremare' (*L'antibiotico mi ha lessato*; *Il dopo pranzo mi ha lessato e ho dormito tre ore*); per *segretaria* non è indicata l'accezione, presente nel maschile, 'chi ha la massima carica direttiva in partiti o sindacati' (*Susanna Camusso è segretaria della CGIL*; *Debora Serracchiani è eletta nel 2008 segretaria del PD di Udine*); per *trasfusione* non è indicato il senso figurato 'passaggio, scambio' (*L'ennesima trasfusione di soldi pubblici*; *Una trasfusione continua di emozioni tra l'artista e la natura*; *L'insegnamento è una trasfusione di conoscenze da insegnante ad alunno*); per *variopinto* non è indicato il senso figurato 'vario, vivace' (*Il carattere variopinto del movimento [new age]*; *Un variopinto cicaleggio* (M. Tobino); *Il cantiere [della teologia italiana] si presenta rigoglioso e variopinto*).

Se si considerassero polisemici anche tutti i lessemi elencati sopra, la percentuale di monosemia nelle tre aree del vocabolario di base si ridurrebbe di un altro 1,5-2%. Nel

³ Tutti gli esempi citati nel seguito sono autentici e sono stati ricavati dalla consultazione di varie fonti di italiano scritto.

complesso, dunque, considerati sia questi ultimi casi sia quello dei femminili animati, le percentuali indicate nella Tabella 3 sono presumibilmente ancora più sbilanciate a vantaggio della polisemia nell'ordine del 3-4%.

3. Polisemia interna ed esterna al vocabolario di base

Scorrendo la lista dei lessemi polisemici del VDB si ha l'impressione che non tutti lo siano allo stesso modo, nel senso che alcuni appaiono, per così dire, 'meno polisemici' di altri. Mi riferisco al fatto che in alcuni casi il lessema è sì polisemico, ma una sola delle sue accezioni ricade nel VDB mentre le altre ricadono in ambiti d'uso esterni al VDB. Ad esempio *utilizzare* ha una sola accezione AU e un'altra OB; *coltello* ha una sola accezione FO e un'altra TS; *lucidare* ha un'accezione AD, una TS e una OB; *torre* ha una sola accezione FO, un'altra RE e ben 11 accezioni TS⁴. In questi casi, cioè, la polisemia del lessema è confinata al di fuori degli ambiti d'uso più comuni della lingua, diversamente dai casi in cui vi è una pluralità di accezioni tutte riconducibili al VDB, com'è ad esempio per *luce* (che ha 28 accezioni delle quali 11 sono FO) o per *aprire* (che ha 25 accezioni delle quali 18 sono FO).

Per valutare l'incidenza nel VDB di questi due diversi casi ho analizzato le qualifiche d'uso delle accezioni in cui si articolano i lessemi polisemici individuati. Ne emerge che nella quasi totalità dei casi i lessemi polisemici del VDB hanno almeno due accezioni, e spesso più di due, appartenenti al VDB o al massimo al vocabolario comune. Solo una minima parte di essi, invece, come mostra la Tabella 4, ha un'unica accezione nel VDB e un'altra, o varie altre, esterne alla sfera del vocabolario comune, cioè che ricadono nell'ambito delle terminologie tecnico-specialistiche oppure risultano marcate BU/OB/LE/RE/DI.

	lessemi polisemici con una sola accezione VDB	% sui lessemi polisemici VDB
FO	87	4%
AU	175	7%
AD	181	11%
tot	443	7%

Tabella 4

Lessemi polisemici del VDB che hanno una sola accezione nel VDB.

Dunque non solo il VDB è costituito al 90% da lessemi polisemici, ma il 93% di essi manifesta tale polisemia articolandosi in sensi di alta frequenza; percentuali che sono entrambe addirittura del 96% nella fascia del vocabolario fondamentale.

Mi pare un aspetto interessante anche dal punto di vista teorico, poiché aggiunge al dato generale per cui la polisemia di un lessema aumenta al crescere della sua frequenza d'uso l'ulteriore elemento che tanto è maggiore la frequenza d'uso di un lessema, tanto più il significato di quel lessema tende ad articolarsi in raggruppamenti di sensi che ricadono anch'essi nelle fasce di maggior frequenza d'uso.

⁴ Tralascio di considerare il caso, di natura in parte diversa, in cui la diversa marca d'uso delle accezioni corrisponde a una diversa qualifica grammaticale, come in *automobile* (FO come sostantivo e BU come aggettivo) o *durante* (FO come preposizione e TS come sostantivo).

4. Polisemia nei sostantivi e nei verbi del VDB

4.1. Quantità di sostantivi e verbi mono- e polisemici

Per avere un quadro più completo della polisemia nel VDB ho analizzato in dettaglio le due categorie dei sostantivi e dei verbi, che oltre a essere le più significative sono quelle più rappresentate nel VDB, come nel lessico in generale. I sostantivi costituiscono il 64,5% del VDB⁵ e i verbi il 22,5%, e nell'insieme le due categorie costituiscono l'87% del VDB. Le Tabelle 5 e 6 riassumono i dati relativi alla quantità di sostantivi e di verbi monosemici e polisemici nelle tre fasce del VDB.

	sostantivi	monosemici	polisemici
FO	1.090	47 (4%)	1.043 (96%)
AU	1.693	174 (10%)	1.519 (90%)
AD	1.560	393 (25%)	1.167 (75%)
tot	4.343	614 (14%)	3.729 (86%)

Tabella 5
Sostantivi monosemici e polisemici nel VDB.

	verbi	monosemici	polisemici
FO	612	15 (2%)	597 (98%)
AU	594	19 (3%)	575 (97%)
AD	308	25 (8%)	283 (92%)
tot	1.514	59 (4%)	1.455 (96%)

Tabella 6
Verbi monosemici e polisemici nel VDB.

Il dato più evidente che emerge dai conteggi è la conferma della maggiore polisemia dei verbi rispetto ai sostantivi: questi ultimi sono polisemici in una percentuale complessiva che è inferiore alla media generale dei lessemi del VDB (86% contro l'89% generale), mentre i verbi superano abbondantemente la media generale (96% contro l'89% generale). Lo scarto è meno netto di quello che, stando ai dati di WordNet citati nella Tabella 2, si ha nel lessico complessivo – dove tra la percentuale di verbi e di sostantivi polisemici risulta una differenza di 32 punti percentuali. Ma è comunque consistente, e tanto più lo diventa man mano che ci si allontana dalla fascia del vocabolario fondamentale. In quest'ultima le percentuali di sostantivi e di verbi polisemici sono molto simili, mentre il divario maggiore si ha nella fascia del vocabolario di alta disponibilità.

Si conferma a questo riguardo quanto accennato nel Paragrafo 2 sul fatto che i sostantivi AD sono la maggior fonte di monosemia nel VDB. Ciò dipende sicuramente, come si è detto, da un elemento contingente, cioè l'inclusione tra i sostantivi AD di un centinaio di femminili animati privi di definizione. Ma emerge anche un fattore più

⁵ Questo dato, come tutti quelli relativi ai sostantivi, non corrisponde a quanto risulta dall'interrogazione del CD-ROM del Gradit perché quest'ultima restituisce anche i casi in cui il lemma appartiene al VDB e ha in intestazione la qualifica di sostantivo ma non appartiene al VDB nell'accezione come sostantivo. Si tratta cioè di casi come *integrale* (che è in intestazione aggettivo e sostantivo AD ma nell'accezione come sostantivo ha la qualifica TS) o come *dopo* (che è FO come avverbio e preposizione ma è CO come sostantivo). Poiché questi lessemi non appartengono al VDB in quanto sostantivi, li ho espunti da tutti i conteggi. La questione non si pone per i verbi, poiché nessuno di essi appartiene al VDB in una qualifica grammaticale diversa da quella di verbo.

strettamente semantico, legato al fatto che il vocabolario di alta disponibilità include diversi lessemi che sono anche termini tecnici – e che infatti il GradiT registra con la doppia marca d'uso AD e TS, ad esempio *alluminio, asma, catarro, mestruazione, petalo, respirazione, rosolia, scarlattina, sudorazione, tosse, vertebra, vitamina* – o che sono in qualche misura assimilabili a tecnicismi poiché indicano apparecchi, macchinari, utensili e altri oggetti a vario grado di 'tecnologicità': *accendino, accetta, aratro, ascia, asciugacapelli, aspirapolvere, astronave, autoambulanza, autoblindo, autobotte, autobus, battipanni, cacciavite, cavatappi, centralino, citofono, contachilometri, corazzata, elicottero, furgone, gasolio, metropolitana, missile, peschereccio, petroliera, pneumatico, portachiavi, portacipria, portaerei, portafinestra, portamonete, roulotte, telecamera, tergicristallo, thermos, tornio, tostapane, trapano*; lessemi, dunque, che sono più tipicamente monosemici rispetto ad altri tipi di nominali. Nelle altre fasce del VDB i lessemi di questo genere sono più rari e concorrono in misura minima a produrre casi di monosemia: tra i monosemici di fascia FO ve ne sono solo due (*aeroplano* e *bicicletta*) e sono pochi anche nella fascia AU (*accendisigaro, ascensore, autocarro, automezzo, cannocchiale, lampadina, mitra, rivoltella, sirena, televisore, tram*).

Più in generale, l'analisi semantica dei sostantivi monosemici del VDB sembra confermare l'ipotesi di Zipf che sia la genericità semantica a rendere i lessemi di alta frequenza più disponibili allo sviluppo polisemico. In tutte e tre le fasce del VDB, infatti, una buona parte dei casi di monosemia è dovuta a lessemi il cui significato individua referenti piuttosto specifici, come ad esempio:

- nomi di parti del corpo (*alluce, palpebra, polpastrello*);
- nomi di sport (*motociclismo, pallacanestro, pattinaggio*);
- nomi di unità di misura (*chilogrammo, etto*);
- nomi di mesi e stagioni (*agosto, dicembre, febbraio, gennaio, giugno, luglio, marzo, ottobre, settembre; lunedì, martedì, mercoledì, venerdì; inverno, pomeriggio*);
- nomi di mestieri (*bambinaia, bancario, calciatore, cantante, cantautore, cassiera, chirurgo, crocerossina, cuoca, dentista, domestica, elettricista, falegname, ferroviere, giornalista, grossista, impiegato, insegnante, istruttore, legale, matematico, negoziante, oculista, orologiaio, panettiere, prostituta, psichiatra, pubblicitario, pugile, religioso, rigattiere, sciatore, scienziato, segretaria, tabaccaio, timoniere, torero, trapezista, veterinario*);
- nomi di capi di abbigliamento e simili (*asciugamano, berretto, biancheria, calzatura, collant, impermeabile, mutanda, panciotto, pigiama, plaid, pullover, tovaglia, tovagliolo, tuta, vestaglia*);
- nomi di cibi e bevande (*affettato, analcolico, aperitivo, arrosto, bibita, cioccolatino, cioccolato, formaggino, formaggio, grappa, maionese, pastasciutta, pasticcino, purè, ricotta, sciroppo, sottaceto, spumante, tagliatella, vermouth, whisky*);
- nomi di malattie o condizioni fisiologiche (*ascesso, bronchite, capogiro, catarro, forfora, gravidanza, mestruazione, raffreddore, sbadiglio*);
- nomi di negozi o luoghi di attività professionale (*ambulatorio, cartoleria, locanda, maglificio, motel, panetteria, pizzeria, portineria, rosticceria, autoscuola*); nomi di luoghi specifici o di loro parti (*bettola, camerata, camposanto, frutteto, pianoterra, pineta, podere, rione, sotterraneo; altopiano, antartico, artico*);
- nomi di oggetti materiali vari (*cancellata, cassapanca, cassetto, bacinella, bancarella, battipanni, ciminiera, dentifricio, fiammifero, lampadario, mestolo, pennarello, petardo, pinzetta, pneumatico, portacenere, portamonete, salsiera, salvadanaio,*

scaffale, seggiola, sofà, soprammobile, stoviglia);

- etnici (*algerino, altoatesino, aostano, argentino, asiatico, ateniense, austriaco, centroamericano, colombiano, iugoslavo, londinese, newyorkese, nordamericano, palestinese, peruviano, potentino, praghese, statunitense, sudamericano, sudtirolese, tunisino, venezuelano*).

Naturalmente nulla vieta che anche questi lessemi sviluppino accezioni tecnico-scientifiche o metaforiche, come avviene per molti loro equivalenti, da *aprile* a *scarpa*, da *macellaio* a *forchetta*. Ma non sembra casuale che proprio questo tipo di lessemi il cui significato è molto circostanziato rappresenti una porzione consistente della monosemia nominale nel VDB.

Non mancano peraltro, tra i monosemici, lessemi di significato astratto: nella fascia di alta disponibilità troviamo *misurazione, opposto, ripensamento, sgomento, suggerimento, tornaconto* e in quella di alto uso *aggiunta, antipatia, esitazione, funzionamento, indignazione, malcontento, nostalgia, saggezza, sopruso*. Nel vocabolario fondamentale circa la metà dei nomi monosemici sono o aggettivi sostantivati (*difficile, importante, impossibile, infelice, malato, meglio, nervoso, noto, occidentale, ottimo, peggio, potente, ridicolo, sconosciuto, sfacciato, ultimo, violento*) o usi sostantivali di avverbi e congiunzioni (*davanti, dentro, dietro, dunque, quando, sopra*); gli aggettivi sostantivati abbondano anche tra i monosemici di alto uso (tra gli altri: *adolescente, adulto, ambizioso, ansioso, aristocratico, audace, avaro, bisognoso, colpevole, competente, defunto, disoccupato, egoista, esperto, fesso, furbo, immorale, indispensabile, inevitabile, ingenuo, ingiusto, ingrato, insensato, insolente, ipocrita, lecito, meridionale, minorenne, mite, poetico, prepotente, religioso, schifoso, sciocco, settentrionale, sovversivo, superfluo, timido, ubriaco, visibile*).

Diverso il discorso per i verbi. Tra i monosemici di fascia AD circa la metà hanno caratteristiche semantiche simili a quelle dei nomi visti sopra, nel senso che indicano azioni specifiche e/o concrete (*abbonare, abbonarsi, abbottonare, affettare, allattare, arredare, creparsi, lessare, multare, piallare, pittare, prudere, risciacquare, ristampare, scarcerare, spettinare, tossire*); ma vi sono anche verbi dal significato molto astratto, come *causare* (AU) e *diventare* (FO). Un gruppo consistente di verbi monosemici, presenti in tutte e tre le fasce del VDB, fa riferimento ad azioni o eventi di natura psicologico-emotiva: *accontentare, accorgersi, addolorare, affezionarsi, affiatare, allarmare, annoiarsi, congratularsi, contentare, controllarsi, decidersi, dominarsi, emozionare, entusiasmare, esitare, meravigliarsi, rattristare, sgomentare, tranquillizzare* (e forse anche *disubbidire, contraddirsi, tradirsi*).

4.2. Quantità e tipi di accezioni sviluppate da sostantivi e verbi

In una seconda fase dell'analisi ho calcolato il numero di accezioni sviluppate dai sostantivi e dai verbi polisemici presenti nel VDB, ripartendole in tre sottogruppi: accezioni basico-comuni (cioè che presentano nel Gradit le marche FO, AU, AD o CO), accezioni tecnico-specialistiche (marca TS), accezioni non comuni (marche BU, OB, LE, DI, RE). Per le sole accezioni del primo gruppo ho inoltre registrato l'eventuale indicazione nel Gradit di un uso estensivo o figurato.

Anche in questo caso il conteggio è stato fatto manualmente, non essendo possibile estrarre in modo automatico dal CD-ROM né le accezioni principali (segnalate da un numero progressivo: 1, 2 ecc.) né quelle secondarie (segnalate da una lettera che segue il numero: 1a, 1b ecc.) né tantomeno le subaccezioni (che si trovano all'interno delle accezioni principali o secondarie precedute da una barra verticale o da un punto e virgola).

Per quanto riguarda le accezioni secondarie, le ho conteggiate sempre come accezioni distinte. Per quanto riguarda le subaccezioni, le ho conteggiate come accezioni solo se dopo la barra o il punto e virgola compare un senso diverso da quello principale e non una mera indicazione fraseologica (com'è ad esempio nella prima accezione di *affare*: "... | farsi gli affari propri, occuparsi solo dei propri problemi")⁶. Le Tabelle 7 e 8 riassumono i dati relativi alla quantità e al tipo di accezioni sviluppate dai sostantivi e dai verbi del VDB.

	sostantivi polisemici	totale accezioni	accezioni VDB e CO	accezioni TS	accezioni BU/OB/LE/DI/LE
FO	1.043	9.599	5.498	3.085	1.016
AU	1.519	9.397	5.094	3.290	1.013
AD	1.167	5.805	3.007	2.140	658
tot	3.729	24.801	13.599 (55%)	8.515 (34%)	2.686 (11%)

Tabella 7
Quantità e tipi di accezioni dei sostantivi polisemici del VDB.

	verbi polisemici	totale accezioni	accezioni VDB e CO	accezioni TS	accezioni BU/OB/LE/DI/LE
FO	597	6.168	4.928	426	814
AU	575	3.301	2.388	347	566
AD	283	1.282	917	127	238
tot	1.455	10.751	8.233 (77%)	900 (8%)	1.618 (15%)

Tabella 8
Quantità e tipi di accezioni dei verbi polisemici del VDB

Una prima considerazione che si può ricavare dal dato relativo al totale delle accezioni è che sia per i sostantivi che per i verbi il numero di accezioni è più alto nella fascia del vocabolario di base, andando progressivamente a diminuire nelle altre due; nei verbi, in particolare, il numero di accezioni si riduce quasi della metà già nel vocabolario di alto uso. Questo dato riconferma ancora una volta la relazione tra frequenza e polisemia: così come il numero maggiore di lessemi polisemici si ha nel vocabolario di base e decresce con il diminuire della frequenza, analogamente il numero maggiore di accezioni si ha nei lessemi del vocabolario di base e decresce con il diminuire della frequenza.

Prendendo sempre in considerazione il numero totale di accezioni sviluppate da sostantivi e verbi, si conferma la maggiore tendenza alla polisemia da parte dei verbi ma non nella misura che ci si poteva aspettare: il rapporto tra numero di accezioni e numero di lessemi – cioè il numero medio di accezioni – è 7,3 per i verbi e 6,6 per i sostantivi, con una differenza che non appare molto significativa. Anzi, se si considerano separatamente le tre fasce del VDB, si vede che solo nel vocabolario fondamentale il rapporto tra numero di accezioni e numero di lessemi è più alto nei verbi (10,3 contro il 9,2 nei sostantivi); nel vocabolario di alto uso è invece 5,7 per i verbi e 6,2 per i sostantivi, e nel vocabolario di alta disponibilità è 4,5 per i verbi e 4,9 per i sostantivi. Dunque i verbi ‘battono’ i sostantivi quanto a numero di accezioni solo se sono di ambito FO, mentre nelle altre due fasce di frequenza sono i sostantivi ad avere il maggior numero di accezioni.

⁶ L'aver considerato anche le accezioni secondarie e le subaccezioni rende i miei dati non confrontabili con quelli relativi all'intero Gradit forniti da De Mauro (1999b), che contemplano solo le accezioni principali.

Persino se si considerano i casi di ‘super-polisemia’ sempre citati nella letteratura e che anche qui ho menzionato, cioè quelli dei verbi che hanno un numero altissimo di accezioni, non emerge un divario eclatante rispetto ai sostantivi. Come si vede dalla Tabella 9, in cui sono elencati i verbi e i sostantivi con 30 e più accezioni (tutti del vocabolario fondamentale tranne tre), sicuramente i casi di maggiore polisemia si hanno tra i verbi, ma la differenza tra le due classi non è così significativa, considerato anche che si tratta di poche centinaia di accezioni su un totale di svariate migliaia.

prendere	67	taglio (AU)	43
fare	56	forma	41
passare	52	ordine	40
essere	48	rete (AU)	36
portare (AU)	48	fondo	35
tenere	47	tavola	35
dare	45	punto	33
andare	37	croce	32
chiudere	36	forza	32
stare	36	passaggio	32
uscire	35	potenza	32
correre	34	tempo	32
girare	34	luce	30
avere	32	testa	30
cadere	32	voce	30
lasciare	32		
mettere	32		
tagliare	32		
tirare	31		
volere	31		
venire	30		

Tabella 9
Verbi e sostantivi con 30 e più accezioni.

In sostanza ciò che emerge dal confronto tra i dati presentati nelle Tabelle 7-8 e quelli presentati nelle Tabelle 5-6 è che la categoria dei verbi è *nell'insieme* più polisemica di quella dei sostantivi, mentre se si considerano solo i lessemi polisemici delle due categorie essi risultano ‘generare polisemia’ in misura molto simile. E infatti se si calcola il rapporto tra quantità di accezioni e quantità di lessemi non sul numero di lessemi polisemici ma sul totale dei lessemi (inclusi i monosemici), esso risulta di 7,1 (anziché 7,3) per i verbi e 5,7 (anziché 6,6) per i sostantivi, con uno scarto decisamente più significativo.

Il concetto di maggiore polisemia dei verbi rispetto ai sostantivi va quindi inteso – almeno in riferimento al VDB – nel senso ‘ci sono più lessemi polisemici tra i verbi che tra i sostantivi’, mentre appare meno corretto intenderlo nel senso ‘i verbi generano molte più accezioni dei sostantivi’.

La differenza più rilevante nel comportamento dei verbi e dei sostantivi VDB rispetto alla polisemia mi pare vada cercata non tanto nella quantità di accezioni che i lessemi delle due classi sviluppano, ma nel tipo, o meglio nell’ambito d’uso, di tali accezioni. Come mostrano le Tabelle 7 e 8, sia nel caso dei verbi che in quello dei sostantivi la percentuale maggiore di accezioni ricade nell’ambito del vocabolario di base e comune (dato, questo, coerente con quanto già visto nel Paragrafo 3 sulla polisemia interna ed esterna al VDB). Ma mentre nel caso dei verbi le accezioni di ambito basico-comune arrivano a costituire quasi l’80% del totale, nel caso dei sostantivi ne costituiscono poco più della metà, mentre una quota notevole è occupata dalle accezioni di

ambito tecnico-specialistico. Queste ultime rappresentano il 34% del totale delle accezioni dei sostantivi e solo l'8% di quelle dei verbi; nel caso dei verbi, anzi, le accezioni tecnico-specialistiche sono all'ultimo posto quanto a numerosità, precedute anche, e non di poco, dalle accezioni di ambito non comune.

Che siano i sostantivi, anziché i verbi, la fonte maggiore di accezioni tecnico-specialistiche, non è un risultato inatteso, considerata la predominanza degli elementi nominali nella formazione delle terminologie e la generale tendenza dei linguaggi specialistici e scientifici nella direzione della nominalizzazione e del depotenziamento dell'elemento verbale. E non stupisce neanche che tra i lessemi del VDB possa annidarsi un certo numero di termini tecnici, posto che la rideterminazione semantica di parole della lingua comune è uno dei meccanismi cui ricorrono i linguaggi specialistici, e alcuni più di altri, per la creazione dei loro termini (Sobrero 1993, Casadei 1994, Gualdo e Telve 2011). Colpisce tuttavia l'entità del fenomeno, considerato che stiamo analizzando l'area lessicale di massima frequenza d'uso: oltre un terzo delle accezioni dei sostantivi polisemici del VDB risulta essere di ambito tecnico-scientifico.

Questo dato appare ancora più notevole alla luce di quello relativo alla dispersione delle accezioni tecnico-specialistiche, molto alta nei sostantivi e bassa nei verbi: come mostrano le Tabelle 10 e 11, ha almeno una accezione tecnico-specialistica il 67% dei sostantivi e il 33% dei verbi. Ovvero, due terzi dei sostantivi del VDB sono anche termini tecnici.

	sostantivi VDB polisemici	con almeno 1 accezione TS
FO	1.043	708 (68%)
AU	1.519	1.035 (68%)
AD	1.167	763 (65)
tot	3.729	2.506 (67%)

Tabella 10

Sostantivi VDB con almeno una accezione tecnico-specialistica.

	verbi VDB polisemici	con almeno 1 accezione TS
FO	597	215 (36%)
AU	575	190 (33%)
AD	283	75 (27%)
tot	1.455	480 (33%)

Tabella 11

Verbi VDB con almeno una accezione tecnico-specialistica.

Nella Tabella 12 è elencata la trentina di sostantivi (tutti del vocabolario fondamentale tranne 6 di alto uso e nessuno di alta disponibilità) che sviluppano 15 e più accezioni in ambito tecnico-scientifico. Una ricognizione sulle loro accezioni mostra che sono moltissimi i linguaggi specialistici rappresentati (così come, in generale, nell'insieme dei sostantivi del VDB), da quelli delle scienze dure – fisica, matematica, chimica, biologia – a quelli di scienze più molli – diritto, economia, medicina, filosofia, linguistica – fino ai linguaggi settoriali non scientifici, dall'araldica alla numismatica, dal linguaggio militare a quelli sportivo e giornalistico.

forma	22
riduzione	22
gruppo	21
taglio (AU)	21
ordine	20
rete (AU)	20
colonna (AU)	19
corda	19
croce	19
fondo	19
corona (AU)	18
titolo	18
modulo (AU)	17
ponte	17
serie	17
sistema	17
tavola	17
testa	17
canale	16
catena	16
divisione	16
figura	16
indice (AU)	16
massa	16
campo	15
passaggio	15
posizione	15
quadro	15

Tabella 11
Sostantivi VDB con 15 e più accezioni tecnico-specialistiche.

Quanto ai verbi, nessuno di essi arriva alle dieci accezioni: il verbo più terminogeno è *incassare*, che ha 9 accezioni tecnico-specialistiche, seguito da *caricare*, *legare* e *tagliare* con 8. Non senza una certa generosità, peraltro, nell'attribuzione della marca TS, perché alcune delle accezioni così qualificate sono in realtà di uso abbastanza comune, come ad esempio per *incassare* nel senso calcistico 'subire reti' o per *caricare* nel linguaggio commerciale 'aggiungere al costo' e in quello militare 'eseguire una carica'; e infatti diverse di queste accezioni hanno nel Gradit la doppia marca CO e TS: *legare* è sia CO che TS metallurgia nel senso 'di due o più metalli, formare una lega' e *tagliare* è sia CO che TS sartoria nell'accezione 'confezionare un indumento'.

Un altro fenomeno sempre relativo alle accezioni tecnico-specialistiche rispetto al quale si ha una notevole differenza tra verbi e sostantivi è quello che riguarda i lessemi con più di due accezioni nei quali tutte le accezioni, esclusa la prima, sono di ambito TS; ad esempio *sperone* ha 15 accezioni di cui solo la prima ('arnese per spronare il cavallo') è AD mentre le altre sono tutte TS. Questi casi si trovano tra i sostantivi di tutte le fasce di frequenza, con una minor presenza nel vocabolario fondamentale e massima nel vocabolario di alta disponibilità; tra gli esempi più salienti cito nel vocabolario fondamentale *denaro* (8 accezioni TS su 9), *intero* (5 su 6), *albero* (8 su 10); nel vocabolario di alto uso *reazione* (7 su 8), *passante* e *trave* (6 su 7), *messaggero* (4 su 5); nel vocabolario di alta disponibilità, oltre a *sperone*, *suola* (10 su 11), *radicale* (9 su 10, e anche la decima è sia AD che TS), *scalpello* (6 su 5), *bastarda* e *ventilatore* (5 su 6). Tra i verbi, invece, il fenomeno è eccezionale: non c'è nessun lessema del genere tra i verbi

fondamentali e ce n'è solo un tra quelli di alto uso (*attrezzare* con 3 accezioni TS su 4) e tra quelli di alta disponibilità (*incollare* con 4 su 5).

Un'ultima considerazione, infine, sulla presenza dei sensi estensivi e figurati nell'ambito delle accezioni del gruppo basico-comune. Dato che questo sottoinsieme di accezioni è molto più numeroso nei verbi, e dato che in generale la polisemia verbale è ritenuta più associata a processi metaforici rispetto a quella nominale, mi sarei aspettata che la quantità di accezioni estensive e figurate fosse decisamente maggiore nei verbi. Invece, come mostra la Tabella 12, l'incidenza di queste accezioni sul totale di quelle basico-comuni è molto simile nei sostantivi e nei verbi, benché maggiore nei secondi.

	accezioni VDB e CO sostantivi	accezioni estens./fig.		accezioni VDB e CO verbi	accezioni estens./fig.
FO	5.498	1.051 (19%)		4.928	970 (19%)
AU	5.094	1.244 (24%)		2.388	716 (29%)
AD	3.007	711 (23%)		917	316 (34%)
tot	13.599	3.006 (22%)		8.233	2.002 (25%)

Tabella 12

Usi estensivi e figurati nelle accezioni VDB e CO di sostantivi e verbi.

5. Qualche osservazione conclusiva

Il dato generale più significativo che emerge dall'analisi che ho presentato è la straordinaria polisemia del VDB, sia per quanto riguarda il numero di lessemi polisemici, sia per quanto riguarda il numero di accezioni che essi sviluppano. Il numero medio di accezioni per lessema nel VDB è di 6,6 per i sostantivi e di 7,3 per i verbi, laddove per il lessico complessivo la media desumibile dalle fonti lessicografiche di varie lingue è del 2,2. Inoltre la quantità di lessemi polisemici nel VDB è dell'89% contro il 19% stimato per il lessico italiano nel suo insieme, e raggiunge addirittura il 96% nel vocabolario fondamentale per decrescere progressivamente in quelli di alto uso e di alta disponibilità.

Questo risultato rappresenta una conferma inequivocabile dell'intuizione di Zipf sulla correlazione tra frequenza d'occorrenza e polisemia, e anche, più in generale, di "ciò che la psicolinguistica e la glottodidattica oramai da decenni hanno reso piuttosto esplicito, ossia il fatto che i fenomeni di frequenza giochino un ruolo significativo a ogni livello della produzione e della ricezione linguistica." (Chiari 2008, p. 77). E accanto alla psicolinguistica e alla glottodidattica possiamo aggiungere, negli anni più recenti, la linguistica cognitiva, che è oggi l'approccio nel quale viene sottolineata con più forza l'importanza della frequenza e della familiarità nel funzionamento non solo delle forme e delle strutture linguistiche ma anche in quello dei processi cognitivi sottostanti. La frequenza è ritenuta infatti, in questo approccio, un fattore decisivo nella *performance* cognitiva, che agisce quindi a ogni livello, da quelli più astratti, come la formazione dei prototipi categoriali, a quelli più concreti, come lo sviluppo della polisemia metaforica (e non è un caso, probabilmente, che proprio una linguistica così fortemente interessata alla polisemia e per la quale la metaforicità è un tema cruciale, finisca per ritenere centrale la questione della frequenza).

In particolare vari studi di semantica cognitiva sottolineano la relazione di *feedback* tra frequenza e polisemia dei lessemi, nel senso che se da un lato la maggior frequenza d'uso è il fattore di innesco per lo sviluppo della polisemia, dall'altro l'elevata polisemicità di un lessema contribuisce a favorirne una maggior frequenza d'uso, perché

accresce la quantità di contesti in cui può essere usato e la quantità di espressioni formulaiche e idiomatiche in cui compare (Fenk-Oczlon, Fenk 2010a, 2010b). Su quest'ultimo punto i conteggi che ho presentato non forniscono dati specifici (benché a una valutazione sommaria sia evidente che il numero di polirematiche associate nel Gradit ai lessemi del VDB è altissimo), ma mi pare possa essere legato all'ipotesi del *feedback* tra polisemia e frequenza uno degli elementi emersi, e cioè il fatto che la polisemia dei lessemi del VDB si manifesta soprattutto in sensi anch'essi di alta frequenza. La quasi totalità dei lessemi del VDB (il 93% in media e il 96% di quelli del vocabolario fondamentale) ha più di un'accezione che ricade anch'essa nel VDB o al massimo nel vocabolario comune; e sia per i sostantivi che per i verbi la percentuale maggiore di accezioni (il 55% per i sostantivi e il 77% per i verbi) si sviluppa nell'ambito del vocabolario di base e comune.

Un secondo dato significativo emerso dall'analisi è quello relativo alla presenza di accezioni tecnico-specialistiche nei sostantivi del VDB: queste accezioni rappresentano il 34% del totale per i sostantivi, e l'esame della loro dispersione mostra che il 67% dei sostantivi ne ha almeno una. Benché il fenomeno della sovrapposizione tra linguaggi specialistici e lingua comune non sia inatteso, per i motivi che ho ricordato nel Paragrafo 4.2., appare però sorprendente l'entità in cui si manifesta nel VDB: che due terzi dei sostantivi del VDB siano anche, in una o più delle loro accezioni, termini tecnici, mi pare un dato che supera ogni previsione a riguardo.

Non era negli obiettivi dell'indagine un'analisi semantica della polisemia dei lessemi del VDB, ma qualche indicazione in questo senso comunque emerge. L'esame dei sostantivi monosemici sembra confermare l'ipotesi di Zipf che sia la genericità semantica a rendere i lessemi di alta frequenza più disponibili allo sviluppo polisemico. In tutte e tre le fasce del VDB, infatti, larga parte dei casi di monosemia è dovuta a lessemi il cui significato individua referenti specifici o concreti; non sembra dunque casuale che la fonte maggiore di monosemia nel VDB siano i sostantivi di alta disponibilità, che annoverano molti lessemi di questo tipo e in particolare lessemi che sono anche termini tecnici o che sono assimilabili a tecnicismi. Quanto alla presenza di accezioni estensive e figurate nei sostantivi e nei verbi, i dati mostrano che l'incidenza di queste accezioni sul totale di quelle basico-comuni è molto simile nelle due categorie, benché maggiore nei verbi. Non emerge quindi una netta conferma del dato generale per cui la polisemia verbale è più tipicamente associata a processi metaforici rispetto a quella nominale; ma per valutare meglio questo aspetto servirebbe un'analisi più raffinata delle accezioni che hanno le qualifiche 'estensivo' e 'figurato', troppo generiche probabilmente per cogliere le differenze in gioco e in particolare quella tra polisemia metonimica e polisemia metaforica, sulla quale si basa molta della differenza semantica tra lessemi nominali e verbali.

Riferimenti bibliografici

- Carloni F. 2000, *Le relazioni statistiche tra frequenza e significato delle parole nella lingua italiana*, in "Italice" 77 [4], pp. 523-534.
- Casadei F. 1994, *Il lessico nelle strategie di presentazione dell'informazione scientifica: il caso della fisica*, in De Mauro T. (a cura di), *Studi sul trattamento linguistico dell'informazione scientifica*, Bulzoni, Roma, 1994, pp. 47-69.
- Chiari I. 2008, *La chiave probabilistica delle lingue: teoria linguistica e applicazioni computazionali*, in Fulci L. e Sciubba E. (a cura di), *Linguaggio, mente e società*, EuRoma, Roma, pp. 55-79.
- De Mauro T. 1980, *Guida all'uso delle parole*, Editori Riuniti, Roma.
- De Mauro T. (1999a), *Introduzione* al Gradit, pp. VII-XLII.
- De Mauro T. (1999b), *Postfazione* al Gradit, vol. 6, pp. 1163-1183.
- De Mauro T. (2009), *Basi di conoscenze e banche dati lessicali XXI secolo*, in *Enciclopedia Treccani*, Istituto dell'Enciclopedia Italiana, Roma. www.treccani.it/enciclopedia/basi-di-conoscenze-e-banche-dati-lessicali_XXI_Secolo (10.2.1014)
- De Mauro T. e Chiari I. (a cura di) 2005, *Parole e numeri. Analisi quantitative dei fatti di lingua*, Aracne, Roma.
- Fenk-Oczlon G. and Fenk A. 2010a, *The association between word frequency and polysemy: a chicken and egg problem?*, in Solovyev V. and Polyakov V. (eds.), *Proceedings of the XIIth International Conference "Cognitive Modeling in Linguistics"*, Kazan, Kazan State University Press, pp. 167-170.
- Fenk-Oczlon G. and Fenk A. 2010b, *Frequency effects on the emergence of polysemy and homophony*, in "International Journal of Information Technologies and Knowledge" 4 [2], pp. 103-109.
- Gradit = *Grande Dizionario Italiano dell'Uso* ideato e diretto da Tullio De Mauro, 6 voll., UTET, Torino, 1999 (2a ed. 8 voll., *ivi*, 2007).
- Green G.M. 1989, *Pragmatics and natural language understanding*, Erlbaum, Hillsdale; trad. it. di Castelnovo W. 1990, *Pragmatica*, Muzzio, Padova.
- Gualdo R. e Telve S. 2011, *Linguaggi specialistici dell'italiano*, Carocci, Roma.
- Leacock C., Towell G. and Voorhees E. 1993, *Corpus-based statistical sense resolution*, in *HLT '93 (Proceedings of the Workshop on Human Language Technology)*, Association for Computational Linguistics, Stroudsburg, pp. 260-265.
- Miller G.A. 1995, *WordNet: a lexical database for English*, in "Communications of the ACM" 38 [11], pp. 39-41.
- Nagy W.E. 1995, *On the role of context in first- and second language vocabulary learning*, Technical Report 627, Center for the study of reading, College of Education, University of Illinois at Urbana-Champaign.
- Sobrero A.A. 1993, *Lingue speciali*, in Sobrero A.A. (a cura di), *Introduzione all'italiano contemporaneo. La variazione e gli usi*, Laterza, Roma-Bari, pp. 237-277.
- Zipf G.K. 1949, *Human behaviour and the principle of least effort. An introduction to human ecology*, Addison-Wesley Press, Cambridge.