

*Isabella CHIARI*

## COLLOCAZIONI E POLIREMATICHE NEL LESSICO MUSICALE ITALIANO

Questo contributo mira a iniziare un lavoro di ricerca che affronterà, con metodologie e strumenti innovativi della linguistica computazionale, un duplice obiettivo: proseguire il lavoro, iniziato da alcuni decenni, sulla chiarificazione dei problemi e delle questioni inerenti alla lessicografia e lessicologia musicale ed esaminare e analizzare un tema molto delicato e eterogeneo della linguistica generale e computazionale attuale, ossia il trattamento delle parole complesse o polirematiche.

Il progetto nasce dunque soprattutto dalla necessità di dare una chiarificazione teorica multidisciplinare e innovative prospettive applicative al continuum di fenomeni che si manifestano come unità lessicali complesse che si muovono al confine tra sintassi e lessico, variamente nominate come collocazioni, polirematiche, lessemi complessi, *idiom*, fraseologie, ecc. La questione sarà affrontata cercando di osservare e comprendere il ruolo della collocazioni e polirematiche nei testi di dominio musicale confrontandole con collocazioni e sintagmi regolari. L'area di interesse individuata è centrale sia da un punto di vista teorico-descrittivo che metodologico e applicativo.

La bibliografia internazionale sulle collocazioni e polirematiche è ricchissima e copre diversi aspetti teorici relativamente a numerose lingue con particolare attenzione all'inglese e contestualmente si osserva un crescente bisogno di strumenti e applicazioni computazionali che siano capaci di trattare unità complesse come collocazioni e polirematiche. La lingua italiana non offre tuttavia una copertura descrittiva sufficiente, né specificatamente strumenti e metodi ritagliati sulle caratteristiche della lingua che possano soddisfare tali bisogni.

## 1. Collocazioni e polirematiche nella riflessione linguistica

Un filone significativo della valutazione linguistica sulle caratteristiche di collocazioni e *idiom* è quello britannico iniziato da John Firth (1957) e Zellig Harris e proseguito da Michael Halliday (1966) e John Sinclair, che elabora la nozione di *open-choice* e *idiom principle* (1991; Sinclair e Carter 2004; Sinclair, Jones et al. 2004); e più recentemente da Cowie (1998) e O'Dell and McCarthy (2008). La tradizione anglosassone possiede un termine generico che può fungere da iperonimo utilizzabile per descrivere una gran quantità di fenomeni diversi sia dal punto di vista semantico, che testuale e morfosintattico, ossia il termine *multi-word expression*, solitamente abbreviato in MWE. Tale termine, molto fortunato in linguistica computazionale, consente di raccogliere sotto un'unica etichetta polirematiche, *idiom*, collocazioni, sintagmi ricorrenti, ecc. Tra le proprietà specifiche attribuite alle MWE in questa tradizione vi sono: selezione limitata del lessico e restrizioni combinatorie, istituzionalizzazione, imprevedibilità semantica, precisione semantica, non-composizionalità e opacità semantica, prosodia semantica, possibile violazione di pattern sintattici, alto grado di lessicalizzazione e di convenzionalità, ecc.

Vi sono numerosi riflessi di questa tradizione nella lessicografia internazionale (Benson 1989), soprattutto per influenza del lavoro di John Sinclair, che ha condotto alla ideazione e produzione di numerosi dizionari di collocazioni (BBI, Oxford Collocations Dictionary, LTP, CCEC, per lo spagnolo REDES e DICE, per il francese EDC e LAF).

Relativamente alla lingua italiana vi sono stati negli ultimi anni alcuni contributi molto significativi: da una parte, ad esempio, la chiarificazione teorica e lessicologica di T. De Mauro anche nei lavori con e di M. Voghera (Voghera 1994; De Mauro e Voghera 1996; Voghera 2004; De Mauro 2005) e da un punto di vista legato alla semantica cognitiva e alla psicolinguistica i lavori di F. Casadei (Casadei 1994; Casadei 1995; Casadei 1995; Casadei 1996), C. Cacciari (Cacciari 1989; Cacciari, Vespignani et al. 2008) e altri (Bisetto 2004; Papagno 2008); dall'altra linea di grande interesse, anche per numerose applicazioni computazionali italiane, è quella che sviluppa la teoria lessico-grammaticale sia dal punto di vista teorico-descrittivo che computazionale, rappresentata in Italia soprattutto dai lavori di Annibale Elia e Simona Vietri (Elia, D'Agostino et al. 1985; Vietri 1986; D'Agostino e Elia 1998; Vietri 2004). La discussione relativa alle

polirematiche nella lessicografia italiana si è sistematizzata ed ha avuto un riconoscimento, anche in termini quantitativi, soprattutto nel *Grande Dizionario italiano dell'uso* (De Mauro 1999) che censisce 67.678 polirematiche diverse (con circa 130.000 lemmi con associate polirematiche), cui si aggiungono circa 2.000 locuzioni complesse di origine latina o straniera (su un totale di 260.709 lemmi monorematici).

Dal punto di vista metodologico, di grande interesse teorico e applicativo in linguistica computazionale, la bibliografia sulle misure statistiche per l'identificazione ed estrazione di MWE (utilizzabile anche nel machine learning, statistical, example- or rule-based, or hybrid) è oramai molto vasta, soprattutto per il trattamento di lingue a morfologia relativamente povera. Solo per dare qualche indicazione di riferimento si segnalano i lavori di impianto statistico, soprattutto di S. Evert (Evert e Krenn 2001; Evert 2004; Hoang, Kim et al. 2009), A. Kilgarriff (2006), Stubbs (2002) e in chiave comparativa (Aristomenis, Nikos et al. 2002; Pearce 2002; Petrovic, Snajder et al. 2006).

Quantitativamente, a seconda di ciò che si decida di far rientrare in questa categoria, si possono stimare un numero di MWE pari a quello dei lessemi semplici memorizzati nel lessico di ciascun parlante<sup>1</sup>; questo in genere si osserva anche a livello di lessico collettivo depositato nelle fonti lessicografiche e computazionali, anche se si tratta di stime piuttosto approssimative senza dati reali a supporto attualmente; a livello di lingua parlata si stimano approssimativamente circa 6 espressioni non letterali al minuto<sup>2</sup>.

Dal punto di vista definitorio una definizione di MWE generica, e in questo soddisfacente, della vaghezza dei confini e varietà dei fenomeni, è quella di Calzolari, Fillmore et al. (2002:1934) secondo cui si tratta di “a sequence of words that acts as a single unit at some level of linguistic analysis”. Per quanto riguarda invece il dominio più specifico

---

<sup>1</sup> Si veda a proposito la stima di Ray Jackendoff (1997: 256) condotta grossolanamente sul corpus *Wheel of Fortune* afferma che/che afferma: “There are a vast number of such fixed expressions; these extremely crude estimates suggest that their number is of about the same order of magnitude as the single words of the vocabulary. Thus they are hardly a marginal part of our use of language”. A livello di lessico condiviso, si valuta che le entrate lessicali del WordNet inglese per il 41% sono costituite da MWE (Fellbaum 1998; Fellbaum 2007).

<sup>2</sup> “Calcoli di frequenza abbastanza approssimativi stimano che le persone usino 6 espressioni non letterali per minuto di parlato, in particolare 1,8 metafore nuove e 4,08 metafore congelate, cioè convenzionalizzate dall'uso” (Cacciari, Vespignani et al. 2008: 413).

delle polirematiche De Mauro (De Mauro 1999: xxxii) scrive che a un primo livello troviamo espressioni dotate di uno “specifico sovrappiù semantico, vale a dire la non ricostruibilità del loro significato in base alla semplice somma dei significati dei singoli componenti” e a un secondo livello esistono polirematiche molto frequenti anche nei linguaggi settoriali, ove spesso troviamo parole molto note che, associate ad altre, assumono significati tecnici e fortemente standardizzati, anche in assenza di quel «sovrappiù semantico rispetto ai suoi componenti» caratteristico della prima tipologia.

Lo spazio lessicale entro cui si individuano fenomeni quali collocazioni, polirematiche, espressioni idiomatiche, fraseologia e unità lessicali complesse è un *continuum* non graduato nel quale si possono individuare entità che di volta in volta esibiscono presenza o assenza di proprietà quali la rigidità morfo-sintattica, la non predicibilità o non componenzialità semantica, interrompibilità, convenzionalità e lessicalizzazione. Se nessuna di queste proprietà può essere considerata discriminante per attribuire una etichetta univoca al fenomeno, esiste tuttavia uno specifico comportamento statistico di queste unità lessicali che ne fa emergere alcune specificità. I lessemi che entrano a far parte di collocazioni e polirematiche hanno congiuntamente una frequenza di occorrenza significativamente più alta della frequenza con la quale i lessemi componenti si congiungono in media con gli altri lessemi di un corpus. Su questa base sono state adoperate diverse misure di associazione o di co-occorrenza per estrarre automaticamente dai corpora possibili candidati a costituire collocazioni e polirematiche.

Solo l'integrazione di indicatori qualitativi (grammaticali, semantici, testuali) con indicatori quantitativi (statistico-testuali) potrà fornire un quadro soddisfacente della varietà di fenomeni che si muovono in questo spazio. In questo lavoro si intende iniziare ad esaminare alcuni aspetti caratterizzanti delle unità lessicali complesse esplorandone il comportamento in uno specifico dominio testuale, quello relativo ai testi in cui si parla di musica.

## **2. Le espressioni complesse nel lessico musicale italiano: costruzione della risorsa**

La scelta del dominio musicale è particolarmente utile per la varietà di caratteristiche della sua terminologia che si presta a osservazioni molto interessanti sull'interazione tra lingua comune e

formazione dei linguaggi specialistici. La lessicologia musicale si fa solitamente iniziare con i lavori di Eggebrecht (1972; 1981) e, per quanto riguarda la tradizione italiana, con i lavori che negli ultimi decenni sono stati curati e coordinati da F. Nicolodi e P. Trovato (Nicolodi e Trovato 1994; Muraro 1995; Nicolodi e Trovato 2000), culminati nella pubblicazione del LESMU (Lessico Musicale Italiano), cui si sono affiancati numerosi altri contributi sia di tipo musicologico che linguistico (Pozzi 1995; Tonani 2005). Dal punto di vista più specificatamente lessicografico è doveroso menzionare almeno il lavoro di riferimento di Coover (1958). La terminologia musicale nella maggior parte delle lingue occidentali è derivata primariamente dall'italiano (ad esempio in espressioni come *da capo*, *opera lirica*, *concerto grosso*, *viola da gamba*, *mezzo piano*, *adagio cantabile*, *con brio*, *bel canto*) ed esibisce caratteristiche interessanti da un punto di vista lessicologico e testuale. La maggior parte della letteratura sull'argomento è centrata sullo studio di termini semplici, monorematiche (Eggebrecht 1955; Muraro 1995; Nicolodi e Trovato 1996; Nicolodi e Trovato 2000) e le risorse lessicografiche specialistiche sono molto carenti nella descrizione e lemmatizzazione delle collocazioni e polirematiche del dominio musicale (Basso 1983; Pfister 1993; Boccagna 1999).

Il lavoro complessivo, di cui questa è una prima esplorazione, si pone come obiettivo la creazione di una risorsa costituita da una lista detta *PoliMus*, di circa 1.500 lemmi polirematici appartenenti al dominio musicale, e di aggiungervi un'analisi morfosintattica, semantica ed etimologica e statistico-testuale allo scopo di permettere una riflessione teorico-linguistica sulle caratteristiche del lessico musicale italiano e allo stesso tempo gettar luce sulle caratteristiche tipiche delle collocazioni e polirematiche italiane.

L'analisi della risorsa complessiva si articola su:

- a) tipologia di lessemi complessi dal punto di vista semantico;
- b) analisi grammaticale, della struttura interna e della cristallizzazione morfosintattica della locuzione;
- c) interazione con il vocabolario comune (per esempio in espressioni del dominio musicale che hanno estensioni semantiche nel vocabolario comune come *all'unisono* “eseguendo simultaneamente la stessa melodia” e anche “insieme, contemporaneamente”<sup>3</sup>, *crescendo rossiniano* “episodio musicale che consiste nella ripetizione continua e

<sup>3</sup> Tutte le definizioni dei termini musicali, ove non specificatamente indicato, sono tratte dal *Gradit* (De Mauro 1999).

prolungata di formule melodiche ritmate con passaggio graduale dal pianissimo al fortissimo” e anche “iterazione, in progressione sempre più forte, rapida e insistita, di un tema, un argomento”, *dare il la* “dare l'intonazione giusta” e anche “dare l'avvio a qcs. o a qcn., scatenare”, *nota stonata* “eseguita fuori tono” e anche “comportamento, gesto o discorso inopportuno”, *tenere bordone* “fare l'accompagnamento a una melodia” e “assecondare, affiancare qcn., spec. in azioni non del tutto lecite”; oppure nella direzione inversa espressioni comuni che sviluppano accezioni specifiche nella terminologia musicale come nel caso di *a capriccio* “capricciosamente, in modo impulsivo, senza riflessione” e in musica “ad libitum”, *a piacere* “a volontà, senza limiti o restrizioni, secondo il gusto personale” e “ad libitum”, *da capo* “di nuovo, da principio” e “nelle partiture musicali, didascalia che, posta alla fine di un brano musicale, ne prescrive la ripetizione integrale”; *golfo mistico* “nei teatri lirici, lo spazio riservato all'orchestra, collocato gener. a un livello inferiore rispetto alla platea”);

d) analisi in campi semantici del sapere musicale (tra cui nomi di strumenti musicali, forme musicali, dinamiche, notazione, tecnica, ecc);

e) comportamento sinonimico che privilegia sinonimie anch'esse espresse in forma di polirematiche (*alterazione in chiave = accidente in chiave, canone infinito = canone circolare = canone perpetuo, coro a voci dispari = coro a voci ineguali = coro a voci miste, flauto dolce = flauto a becco = flauto dritto*);

f) origine e ricostruzione etimologica;

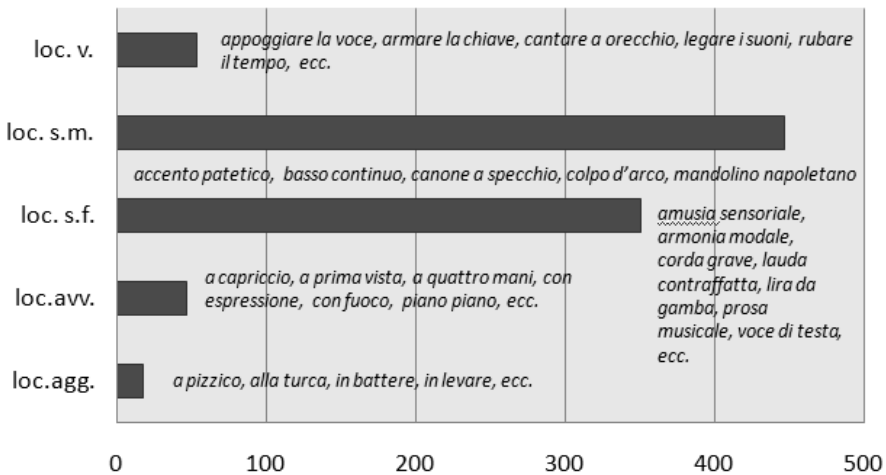
g) omografia testuale, nel caso in cui l'espressione sia passibile di lettura anche in forma non polirematica, rilevante ai fini del trattamento automatico (*basso cantante* “tipo di voce maschile che, pur mantenendo i tratti del basso, si avvicina a quella del baritono | cantante con tale voce”, *cassa chiara* “tamburo militare piccolo e piatto, dal suono acuto e brillante”);

h) frequenza e dispersione misurata in corpora di riferimento e in corpora specialistici.

La lista *PoliMus* attualmente consta di 1.528 lemmi, sottoposti ad analisi morfosintattica e semantica. Il primo nucleo della lista è stato estratto dal *Grande Dizionario Italiano dell'Uso* (De Mauro 1999), la più estesa risorsa lessicografica esistente per la lingua italiana ed è costituito da 496 lemmi polirematici etichettati con marca MUS (su un totale di 2.635 lemmi del dominio musicale).

A questo nucleo sono state fatte numerose integrazioni, sia mediante spoglio manuale a partire da diverse fonti lessicografiche e testuali (Della Corte e Gatti 1930; Limenta 1940; Allorto e Ferrari 1959; Candé 1968; Basso 1983; Calbi 1992; Bertolini e Zannoni 1997; Napoli 2006; Braccini 2007), esplorate non solamente nelle voci a lemma, ma anche all'interno delle singole voci, sia mediante l'integrazione con strumenti di estrazione (semi)automatica di collocazioni (Kilgarriff e Tugwell 2001; Kilgarriff, Rychly et al. 2004).

Dal punto di vista grammaticale la lista di polirematiche attualmente si distribuiscono secondo le parti del discorso indicate in Tabella 1. Come si può vedere, la stragrande maggioranza delle locuzioni rilevate sono locuzioni sostantivali maschili e femminili, mentre le locuzioni avverbiali e aggettivali spesso si sovrappongono includendo espressioni che possono svolgere entrambe le funzioni sintattiche. Ad esempio *in battere* occorre sia in funzione aggettivale in espressioni come *tempo in battere*, *movimento in battere*, *ritmo in battere* e in funzione avverbiale *suonare in battere*, *attaccare in battere*.



**Tabella 1 - Parti del discorso delle polirematiche musicali (PoliMus)**

Bisogna anche aggiungere che soprattutto l'individuazione delle locuzioni verbali è particolarmente insidiosa poiché più raramente queste sono incluse nelle opere lessicografiche, anche specialistiche. Il Gradit (De Mauro 1999), ad esempio, riporta solamente 118 lemmi con accezione MUS tra le mono- e polirematiche (su 2.635 lemmi MUS, ossia

poco più del 4,47%), di cui solo 11 polirematici, ossia il 2,2% delle polirematiche musicali censite (*armare la chiave, dare il la, portare il tempo, portare la voce, rubare il fiato, rubare il respiro, rubare il tempo, stringere il tempo, suonare alla punta, suonare sull'anticipo, tenere bordone*). Se tuttavia si vanno a guardare i 107 lemmi verbali monorematici vi si trovano alcune locuzioni polirematiche associate ed etichettate globalmente come vocabolario comune (CO) e che chiaramente appartengono al dominio tecnico-specialistico musicale (cui andrebbe aggiunta la marca TS MUS), come *accordare gli strumenti*.

In genere se si osservano le distribuzioni delle categorie grammaticali confrontando in percentuale le monorematiche e le polirematiche, non solo del dominio musicale, ma anche di altri settori tecnico specialistici, si scopre che non vi è una omologia. Tale non corrispondenza può dipendere da vari fattori, uno dei quali può essere una certa maggior opacità e individuabilità delle polirematiche verbali che discende anche dal fatto che i verbi usati in accezioni tecniche sono spesso verbi molto comuni. De Mauro stesso (De Mauro 1999: 1177) afferma che:

*Dei procedimenti di arricchimento lessicale questo è di gran lunga il più morbido e, per dir così, insensibile e si capisce che spesso possa essere sfuggito ai lessicografi*

Sempre nel Gradit dei 118 lemmi monorematici verbali del dominio musicale la metà (68 verbi, ossia circa 58%) sono classificati come appartenenti al vocabolario di base (nelle tre fasce fondamentale, alta disponibilità, alto uso) o al vocabolario comune (sono ad esempio FO: *accompagnare, calare, entrare, risolvere, sostenere* AD: *accentare, ingrandire, pizzicare, registrare* AU: *accordare, arrangiare, risuonare, ritardare* CO: *armonizzare, concertare, fraseggiare, glissare, impostare, orchestrare, trasporre, trillare*). Mentre alcuni di questi verbi hanno un ampio spettro di applicazione nel dominio musicale (ad esempio, tra i verbi sopra citati, *calare, risolvere, glissare, fraseggiare*) altri invece si trovano quasi esclusivamente in collocazioni specifiche (è il caso di *pizzicare le corde, sostenere la voce, sostenere una nota, articolare i suoni*).

Più avanti saranno presentati più in dettaglio i dati grammaticali relativi a un sottoinsieme dei sostantivi appartenenti al lessico musicale, che qui si tralasciano.



Un secondo elemento di riflessione è la categorizzazione secondo le marche d'uso proposte dal Gradit. L'attribuzione di una marca è un'operazione complessa poiché richiede considerazioni sia di tipo qualitativo che quantitativo. Vi sono infatti marche a stretta correlazione con indici di frequenza (sono le tre fasce del vocabolario di base e la marca basso uso) e una marca, quella che individua il vocabolario comune, che è anch'essa riferita a stime di frequenza d'uso e di conoscenza da parte dei parlanti; e vi sono marche che invece si riferiscono a varietà diatopiche (regionale, dialettale) o diacroniche (obsoleto), alla provenienza da lingue straniere (esotismo).

Per le 1.032 voci polirematiche aggiunte alla lista base del Gradit si è proceduto dunque all'etichettatura escludendo le etichette del vocabolario di base e quella del vocabolario comune, che semmai potranno essere attribuite a posteriori dopo l'analisi delle frequenze sui corpora. Dunque le etichette applicate sono state essenzialmente TS ed ES (come *air à boire*, *angel lute*, *chanson balladée*, *tonic sol-fa*), cui si aggiungono alcuni casi di OB (*accento parziale*, *canto a nota e parola*, *filare la voce*, *fugha plagale*, *spianare la voce*, *trillo calato*, *voce pecorina*) individuati nello spoglio di fonti testuali e lessicografiche antiche.

<i>Marca d'uso</i>	<i>Lemmi</i>	<i>Esempi</i>
Tecnico-specialistico (TS)	1.357	<i>cembalo d'amore, legare i suoni, organetto automatico</i>
Esotismi (ES)	284	<i>aire de cour, chanson à refrain, steel drum, aman xuur</i>
Obsoleto (OB)	84	<i>spiccare le note, voce pecorina, agilità martellata</i>
Comune (CO)	16	<i>a orecchio, sala da musica, accordo di chitarra, tema musicale</i>
Basso uso (BU)	7	<i>artista orecchiante, oleosità del canto, scala fatale</i>

### 3. I nomi degli strumenti musicali e le loro polirematiche

In questo contributo saranno presentati alcuni dati complessivi e successivamente uno spaccato di analisi condotte su un sottoinsieme dei lemmi selezionati relativo ai 217 lemmi che si riferiscono ai nomi di strumenti musicali (sia mono- che polirematici) italiani e prestiti da lingue straniere (come *adungo*, *armonica a bocca*, *chitarra classica*,

*clarinetto, corno inglese flauto traverso, liuto arabo, oboe da caccia, pianoforte a coda, oud, violino fino a zither).*

L'obiettivo di questa analisi è l'osservazione delle determinazioni aggettivali soprattutto dei nomi di strumento/strumenti musicali e in particolar modo due tipi di casi:

- a) Determinazioni che danno luogo a strutture polirematiche o a collocazioni relative al nome dello strumento;
- b) Determinazioni che rappresentano qualificazioni soggettive degli strumenti e delle famiglie di strumenti musicali, con particolare attenzione alle qualificazioni del suono, del timbro e a tutte le qualificazioni che ci danno un'indicazione sui 'tratti di personalità' che comunemente si attribuiscono agli strumenti nella tradizione occidentale.

Per condurre le analisi si è scelto di usare un corpus che fosse esteso e variato e che potesse contenere testi appartenenti a tipologie diverse.<sup>4</sup> A questo scopo si è scelto di lavorare su ItWac (Baroni e Kilgarriff 2006), corpus di lingua italiana costituito da circa 1.909.535.984 parole (occorrenze). Il corpus è parzialmente bilanciato attraverso una serie di procedure di Web crawling, lemmatizzato con un lemmatizzatore di impianto statistico TreeTagger<sup>5</sup>, interrogato attraverso il Corpus Query System di Sketch Engine (Kilgarriff e Tugwell 2001; Kilgarriff, Rychly et al. 2004).

Per il primo obiettivo si sono individuati i nomi semplici degli strumenti (monorematiche e polirematiche di base), ottenendo una lista di 217 lemmi, questa lista è stata poi proiettata sul corpus ItWac per individuare il numero delle occorrenze per ciascun lemma (e laddove il CQS non riconoscesse la lemmatizzazione, si è proceduto a verificare inserendo anche la forma flessa nell'interrogazione).

La lista finale su cui si è lavorato è costituita dai seguenti 184 lemmi non gravemente ambigui<sup>6</sup>:

<sup>4</sup> Nella continuazione della ricerca si opererà un confronto tra il corpus di riferimento scelto con un corpus specialistico di testi tecnici del dominio musicale.

<sup>5</sup> <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>

<sup>6</sup> Si sono inoltre escluse le forme ambigue, ossia le forme omografe con altre di grande frequenza (ad esempio *lira, organo, piano, piatti, piccolo, piffero, sonagli, tastiera, timpani, tromba, viola*), si sono invece incluse le forme anche ambigue a bassa frequenza e si è proceduto a disambiguazione manuale. Trattandosi di un corpus di circa 2 miliardi di occorrenze solo *piatti*, ad esempio, in ItWac conta 33.525 occorrenze: è dunque impensabile disambiguarle manualmente osservando le

adungo, albero della pioggia, arciliuto, armonica a bocca, armonica a bocca cromatica, armonica a bocca diatonica, armonium, arpa, arpa barocca, arpa birmana, arpa celtica, arpa eolia, arpa eolica, arpa irlandese, arpa kundi, arpicordo, atumpan, aulos, balafon, balalaica, balalaika, bandir, bandurria, banjo, basso tuba, bassotuba, bastone della pioggia, batteria elettronica, berimbau, bifira, bongo, bubboliera, buzuki, campane tubolari, castagnette, caxixi, cetra, chalumeau, chiarina, chitarra, chitarra acustica, chitarra battente, chitarra blues, chitarra classica, chitarra country, chitarra elettrica, chitarra flamenco, chitarra gipsy, chitarra jazz, chitarra rock, cimbasso, cintura zulu, clarinetto, clarinetto alto, clarinetto basso, clarinetto contrabbasso, clarinetto soprano, clavicembalo, clavicordo, contrabbasso, controfagotto, cornamusa, corno di bassetto, corno francese, corno inglese, corrù marino, cromorno, daff, darabukka, didgeridoo, djembé, doum-doum, drum machine, duduk, dulciana, dulcimer, engoma, fisarmonica, flauto, flauto basso, flauto di pan, flauto diritto, flauto dolce, flauto traverso, flicorno, gaida, gandira, ghironda, gong, grancassa, guitar synth, harmonium, heckelphon, hosho, kalimba, kayamba, kazoo, kora, koto, kring, launeddas, liuto, liuto barocco, liuto marocchino, liuto rinascimentale, mandola, mandolino, maracas, marimba, mbira, mejwes, mellotron, müsa, müsa appenninica, nacchere, ngoma-mutandarikwa, oboe, oboe al gaita, oboe da caccia, oboe d'amore, ocarina, onde martenot, organo elettrico, ottavino, ottobasso, oud, palo della pioggia, pandeiro, pianoforte, pianoforte a coda, pianoforte digitale, pianoforte verticale, pifara, pommer, rabab, rakatak, rankett, rauschpfeife, riqq, sarrusofono, sassofono, sassofono baritono, sassofono contralto, sassofono tenore, sax, saxofono, schalmey, sequencer, shamisen, shekere, shofar, sintetizzatore, sistro del mali, sitar, sousafono, synclavier, tamba tamba, tamburello, tamburiza, tamburo, tamburo likuti, tamburo taman, tastiera elettronica, tbilat, tin whistle, tiorba, tubax, ukelele, vibrafono, viola da gamba, viola d'amore, violino, violino elettrico, violino gitano, violino tzigano, violino zigano, violoncello, violoncello barocco, violone, water drum, xalam, xilofono, zampogna, zither.

Già in termini di frequenza grezza dalla lista di partenza solo 22 lemmi hanno occorrenze molto alte (superiori a 1.000), 100 hanno

---

concordanze; stesso dicasi per *lira*, che conta 225.465 occorrenze, le cui principali qualificazioni *vecchia* (10.066 occorrenze), *italiana* (961), *lorda* (272), *turca* (213), *siriana* (21) fanno chiaro riferimento all'unità monetaria e non allo strumento musicale.

occorrenze tra 10 e 999, 35 si collocano tra 1 e 9 (di cui 12 hapax) e 27 non registrano alcuna occorrenza nel corpus (vedi Tabella 2).

<b>Strumenti</b>	<b>Occorrenze in ItWac</b>
Chitarra	41.209
Pianoforte	17.703
Arpa	13.634
Violino	12.596
Flauto	7.913
Tamburo	4.573
Fisarmonica	3.719
Violoncello	3.451
Contrabbasso	3.382
Clarinetto	3.226
Gong	2.311
Mandolino	2.008
Sintetizzatore	1.971
chitarra acustica	1.788
chitarra elettrica	1.719
Sassofono	1.522
Cetra	1.460
Oboe	1.356
Liuto	1.323
Clavicembalo	1.268
Tamburello	1.230

**Tabella 2 - Lista di frequenza lemmatizzata degli strumenti musicali in ItWac**

A questo punto di tutti i lemmi registrati dal corpus si è fatta un'analisi sulle MWE, e in particolar modo sulle sequenze lemma + aggettivo (pre- e post- posto), in modo da osservare le determinazioni aggettivali e quanto tali qualificazioni costituissero lemmi polirematici (presenti o assenti dalla lista). A titolo esemplificativo si riporta in Tabella 3 l'insieme delle determinazioni del lemma *chitarra* che si colloca al rango 1. Tra le determinazioni figurano già polirematiche presenti nella lista di partenza, come *chitarra acustica*, *chitarra elettrica*.

<b>Lemma: Chitarra – Aggettivi frequenti</b>	<b>Occorrenza in ItWac</b>
acustico	2.137
elettrico	2.131
classico	801
solista	343
ritmico	251
secondo	147
primo	122
basso	96
moderno	79
spagnolo	38
folk	37
barocco	33
flamenco	20
brasiliano	17
manouche	14
hawaiano	12
romantico	12

**Tabella 3 - Determinazioni aggettivali del lemma 'chitarra' in ItWac**

Se osserviamo l'esempio possiamo notare che molte delle 17 determinazioni aggettivali (da cui sono state escluse quelle 'soggettive' di cui si parlerà più avanti, come *secco*, *noioso*, *stridente*, *languido*, *vagabondo*, *immancabile*, *cristallino*, *melodico*, *amato*) sono delle polirematiche a tutti gli effetti. Degli aggettivi individuati in realtà solo tre, pur essendo polirematiche peraltro comuni per tutti gli strumenti, ossia *chitarra solista* (343 occorrenze), *prima chitarra* (122 occorrenze) e *seconda chitarra* (147 occorrenze) non si riferisce allo strumento bensì al musicista per metonimia. Gli altri aggettivi individuati sono tutte polirematiche che definiscono luoghi di costruzione (*chitarra brasiliana*, *chitarra hawaiana*, *chitarra spagnola*), stili costruttivi (*chitarra acustica*, *chitarra elettrica*, *chitarra classica*, *chitarra basso*), stili strumentali (*chitarra ritmica*, *chitarra moderna*, *folk*, *chitarra flamenco*), stili relativi a diverse epoche storiche (*chitarra barocca*, *chitarra romantica*).

Anche solamente osservando il numero delle polirematiche stabilizzate nel lessico musicale (anche in quello che è penetrato nella lingua comune) si registra una regolarità proporzionale tra il numero delle polirematiche e la frequenza d'uso della forma base monorematica. Il lemma *liuto* ad esempio presenta 9 forme polirematiche quattro delle quali di tipo diatopico (*arabo, greco, siriano, cinese*), tre di tipo cronologico (*rinascimentale, barocco, medievale*) e due di carattere costruttivo (*tiorbato* che si riferisce all'aggiunta di corde, *membrafono* che si riferisce al meccanismo di produzione del suono).

<b>Lemma: Liuto – Aggettivi frequenti</b>	<b>Occorrenza in ItWac</b>
arabo	52
rinascimentale	16
barocco	8
medievale	4
greco	3
siriano	2
cinese	2
tiorbato	1
membrafono	1

**Tabella 4 –Determinazioni aggettivali del lemma ‘liuto’ in ItWac**

In altri casi come quello delle *launeddas*, strumento a fiato a tre canne, che registra 225 occorrenze in ItWac, l'unica determinazione è quella geografica, ossia relativa alla origine *sarda* dello strumento (5 occorrenze); la *chiarina*, tipo di tromba di epoca romana (106), ha una sola determinazione di tipo cronologico *chiarina medievale*; l'*harmonium* (o *armonium*), strumento a tastiera usato soprattutto per la musica sacra (173), presenta nel corpus tre polirematiche *indiano, enarmonico, maya*.

Osservando globalmente la relazione tra frequenza e numero di polirematiche generate dalla forma base possiamo individuare una tendenza analoga alla nota legge di Zipf sul numero dei significati (Zipf 1949) per cui la relazione tra frequenza di un lemma e il numero delle sue accezioni è direttamente proporzionale. Esaminando la

Tabella 5 che contiene la sezione alta (ranghi bassi) e quella bassa delle frequenze degli strumenti musicali, con accanto l'indicazione sul numero di polirematiche Nome/Aggettivo.

<i>Lemma</i>	<i>ItWac</i>	<i>N° polirematiche N/Agg</i>
chitarra	41209	17
<u>pianoforte</u>	17703	11
arpa	13634	19
violino	12596	21
flauto	7913	27
fisarmonica	3719	17
<u>violoncello</u>	3451	10
contrabbasso	3382	9
clarinetto	3226	11
<u>gong</u>	2311	6
mandolino	2008	13
sintetizzatore	1971	10
sassofono	1522	6
cetra	1460	8
oboe	1356	13
liuto	1323	15
-----		
maracas	275	3
<u>oud</u>	256	7
saxofono	225	2
launeddas	225	1
mandola	225	1
marimba	203	1
ottavino	177	1
harmonium	173	3
koto	150	2
didgeridoo	136	2
chiarina	106	1

**Tabella 5 - Legge di proporzionalità tra frequenza e numero delle polirematiche**

Complessivamente la relazione è piuttosto stabile, tuttavia si registrano alcune anomalie, ad esempio, *pianoforte* che pur essendo a rango 2 ha solo 11 polirematiche, oppure *oud*, che nella fascia bassa della lista ha 7 polirematiche. Occorre qui aprire una parentesi metodologica: la valutazione della sussistenza di una relazione di proporzionalità diretta tra frequenza e numero delle polirematiche (chiamiamola *legge/tendenza sul numero delle polirematiche*), per essere correttamente valutata, necessiterebbe la considerazione di tutte le polirematiche sviluppate da una testa lessicale, non solamente il sottoinsieme delle polirematiche N/Agg che si sono prese in esame nel presente lavoro. Ad esempio, per *pianoforte* risulterebbero oltre alle undici forme (*verticale* 63, *classico* 59, *acustico* 57, *digitale* 51, *moderno* 26, *elettrico* 25, *solista* 22, *elettronico* 9, *contemporaneo* 4, *automatico* 4, *meccanico* 3) da aggiungere anche le forme 5 forme *pianoforte a coda* (235), *pianoforte a 4/6/8 mani* (116), *pianoforte a mezza coda/mezzacoda* (17), *pianoforte giocattolo* (15), *pianoforte a grancoda* (4), arrivando a 16 polirematiche che hanno come testa il lemma *pianoforte*.

Uno degli obiettivi di questa ricerca sarà proprio di verificare l'ipotesi generale relativa al rapporto frequenza / numero delle polirematiche attestate tenendo conto delle diverse tipologie grammaticali attive per ciascun nome di strumento. Osservando poi da vicino le tipologie di polirematiche dei nomi di strumento del tipo N/Agg si possono iniziare a ipotizzare gamme di qualificazioni raggruppabili in poli non esclusivi, del tipo:

- a) qualificazione tecnica e costruttiva (*acustico, elettronico, digitale, verticale, orizzontale*);
- b) qualificazione geografica o di provenienza (*spagnolo, siriano, hawaiano, balcanico, napoletano*);
- c) qualificazione cronologica (*antico, rinascimentale, barocco, romantico, contemporaneo*);
- d) qualificazione stilistica ed esecutiva (*folk, moderno, ritmico, bluegrass*).

La non esclusività si riferisce non solo al fatto che alcune qualificazioni sono trasversali ossia sono attribuite a molti strumenti, mentre altre sono peculiari di uno solamente, ma si riferisce, in altro senso, anche al fatto che un'etichetta che sia in origine di tipo geografico o tecnico, può poi assumere un senso di tipo stilistico o cronologica, per



cui l'etichetta di fatto è polisemica, se non all'origine, nell'osservazione attuale. Ad esempio, laddove una qualificazione nasce come tecnica, come nel caso di *chitarra acustica* (che è la polirematica più frequente con 2.137 occorrenze) che si oppone alla chitarra senza qualificazioni (*chitarra spagnola*, la quale assume poi la denominazione *classica* appunto per contrapposizione alle altre tipologie tecniche ed esecutive), e soprattutto a *chitarra elettrica*, nel tempo finisce per significare, non solamente una chitarra senza amplificazione e con tutte le corde in metallo (a differenza della classica che ne presenta tre in nylon), ma anche uno stile esecutivo di impronta popolare e spesso caratterizzato da ritmo ed esecuzioni con accordi fissi o quasi. Sarebbe dunque utile e fruttuoso osservare lo sviluppo di tali polirematiche di qualificazione e in quali direzioni si sviluppano le estensioni semantiche nel tempo.

#### **4. Qualificazioni soggettive di stato d'animo degli strumenti musicali**

Una seconda linea di questa ricerca mira a mettere in luce espressioni non polirematiche che esprimano le caratteristiche di stato d'animo di ciascuno strumento o famiglia di strumenti o meglio lo stato d'animo che inducono nell'ascoltatore poi attribuito, quasi come tratto di personalità, allo strumento nel senso comune. Esempi di questo tipo di qualificazione sono *arpa magica*, *violino stridulo*, *chitarra tagliente*, *tromba angelica*, *corno possente*, *violoncello caldo*, *violino straziante*, *flauto lamentoso*. Per questa analisi ci si è sempre avvalsi dello stesso corpus, ItWac, descritto nel § 3. Anche in questo caso si è circoscritta l'analisi alle costruzioni di tipo N/Agg (con aggettivo pre- o postposto).

Complessivamente emergono una serie di caratteristiche interessanti che riguardano la *varietà orizzontale di aggettivazioni* (come gli stessi aggettivi si attribuiscono a diversi strumenti) e la *varietà verticale di aggettivazione* (come uno stesso strumento ha il potere di produrre una grande varietà di qualificazioni).

Quelle che possiamo chiamare qualificazioni di tipo soggettivo (ossia quella che non riguarda le peculiarità costruttive, geografiche, cronologiche ed esecutivo-stilistiche) nel senso comune rappresentato dal corpus generalista ItWac sono molto numerose (a differenza di

corpora specialistici di ambito musicale)<sup>7</sup>. Alcune qualificazioni sono molto ricorrenti e, per così dire, orizzontali (v.

Tabella 6), ossia si applicano a un gran numero di strumenti pressoché indistintamente. È il caso di qualificazioni come *magico*, *stonato*, *dolce*, *lontano*, *allegro*, *angelico*, *bello*, *elegante*, *mitico*, *prezioso*, *struggente*, *divino*, *languido*, *soave* che hanno un ampio spettro di attribuzioni a strumenti, come si può vedere per *magico* (v.

Tabella 7) che spazia su 20 strumenti dal flauto (per l'evidente riferimento al titolo della celeberrima opera mozartiana) passando per la viola, l'arpa, la chitarra, ma anche sitar, cornamusa, liuto e mandolino.

Altri esempi sono *acuto* che si applica a tromba, 6 occorrenze, clarinetto 2, ocarina 1; *angelico* a tromba 12, flauto 4, arpa 2, trombone 2, tuba 2; *struggente* a violino 7, pianoforte 4, clarinetto 1, fisarmonica 1, sassofono 1; *dolce* a chitarra 13, fagotto 6, tuba 4, mandolino 3, violino 3, cetra 1, contrabbasso 1, liuto 1.

Se invece proviamo ad assumere il punto di vista opposto, ossia l'osservazione della varietà di qualificazione attribuiti a ciascuno strumento riusciamo ad avere una rappresentazione della ricchezza delle associazioni e, anche indirettamente delle potenzialità espressive riconosciute a ciascuno strumento. Ci sono strumenti più versatili anche timbricamente, ed altri confinati a ruoli molto circoscritti.

Se esaminiamo la

Tabella 8 scopriamo che nella fascia alta (fino a nove tratti) della lista di varietà verticale delle qualificazioni troviamo al rango 1 il *violoncello* come lo strumento che insieme alla *chitarra* e, sorprendentemente, alla *fisarmonica*, risulta il più ricco di attribuzioni qualificative di tipo soggettivo. Facendo qualche esempio:

- il *violoncello* ha ben 26 qualificazioni (*caldo* 8, *pulito* 4, *prezioso* 4, *romantico* 3, *magico* 3, *malinconico* 2, *elegante* 2, *implorante* 1, *spumeggiante* 1, *mesto* 1, *drammaturgico* 1, *soave* 1, *pacato* 1, *possente* 1, *accattivante* 1, *delizioso* 1, *funebre* 1, *incisivo* 1, *lirico* 1, *solitario* 1, *tenero* 1, *allegro* 1, *mitico* 1, *rotondo* 1, *severo* 1, *lontano* 1);

---

<sup>7</sup> Si è iniziato il lavoro anche sul confronto con un corpus specialistico per osservare proprio la differenza di qualificazione e d'uso delle costruzioni polirematiche e delle qualificazioni soggettive, ma per motivi di spazio non è stato possibile presentarlo e discuterlo qui.

- l'*arpa* ne ha 12 (*magico* 18, *divino* 7, *lunare* 6, *etereo* 5, *soave* 5, *aggraziato* 3, *soffice* 3, *armonioso* 3, *celeste* 3, *vibrante* 2, *angelico* 2, *lamentevole* 1);
- la *grancassa* 12 (*fragoroso* 9, *assordante* 3, *sussiegoso* 2, *stonato* 2, *battuto* 2, *infernale* 2, *incessante* 1, *marziale* 1, *ritmico* 1, *spettacolare* 1, *eccessivo* 1, *potente* 1).
- i *sonagli* 7 (*rumoroso* 2, *sonante* 1, *allucinato* 1, *strisciante* 1, *mistico* 1, *divertente* 1, *potente* 1);
- il *corno* ne ha 5 (*magico* 16, *possente* 8, *robusto* 8, *caratteristico* 8, *antico* 7);

Abbiamo così alcune indicazioni su come si tenta di dare un contesto espressivo, spesso umanizzato, allo specifico timbro di ciascuno strumento e su come si tenta di identificare la qualità del suono mediante l'uso metaforico, evocativo e filtrato dal soggetto (ad esempio *caldo/freddo*). Le qualificazioni di questo tipo non sono puramente soggettive, ma hanno una serie di costrizioni di tipo esterno e culturali, dipendenti ad esempio dalla tecnica costruttiva e le possibilità sonore e timbriche dello strumento (difficilmente si potrà dire che la grancassa è acuta, per ragioni tecniche) e dal contesto abituale d'uso di ciascuno strumento nella musica occidentale e nella specifica tipologie musicali di forma e di stile.

<i>lemma</i>	<i>n° strumenti diversi</i>
magico	20
stonato	10
dolce	8
lontano	6
allegro	5
angelico	5
bello	5
elegante	5
mitico	5
prezioso	5
struggente	5
divino	4
languido	4

soave	4
sonoro	4
splendido	4
acuto	3

**Tabella 6 - Aggettivi orizzontali in ItWac**

<i>strumento</i>	<i>ItWac</i>
flauto	638
viola	38
arpa	18
chitarra	16
corno	16
piffero	15
violino	14
sitar	4
cetra	3
violoncello	3
balalaika	2
cornamusa	2
fiarmonica	2
tamburello	2
trombone	2
launeddas	1
liuto	1
mandolino	1
ocarina	1
sassofono	1

**Tabella 7- Strumenti 'magici' in ItWac**

<i>Strumento</i>	<i>N° Agg diversi</i>
violoncello	26
chitarra	25
fiarmonica	22
cetra	21

flauto	18
sintetizzatore	18
sax	16
violino	16
pianoforte	14
sassofono	14
mandolino	13
arpa	12
clarinetto	12
contrabbasso	12
grancassa	12
gong	11
tromba	11
liuto	9

**Tabella 8 - Ricchezza qualificativa in ItWac**

### **Alcune conclusioni**

Questa prima analisi dei nomi di strumenti musicali fa emergere alcune questioni sia dal punto di vista descrittivo, sia metodologico. Da quest'ultima prospettiva emergono regolarità nel rapporto tra frequenza del nome di strumento e sua capacità di produrre polirematiche stabili nella lingua che si cristallizzano nei testi non solo specialistici. Tale regolarità ci sembra poter essere estesa a qualunque altro dominio linguistico, ma questo va verificato ampliando la base di dati ed estendendola a diverse strutture sintattiche tipiche.

Si sottolinea inoltre la crucialità dell'integrazione di risorse lessicografiche integrate da analisi e dati estratti da grandi corpora per l'osservazione del comportamento testuale e dell'uso delle polirematiche nei testi. In particolare nel dominio musicale ci permette di riflettere sulle tipologie semantiche prototipiche dei nomi di strumento e su come queste poi vengano inserite in un circolo di estensione semantica che finisce per attraversare le tipologie orizzontalmente (ma ovviamente lo stesso metodo può essere utilmente adoperato per osservare ad esempio la formazione delle polirematiche avverbiali, come *moderato cantabile*, *adagio con brio*, *con agitazione*, *fuori tempo* o verbali, ad esempio).

Per quanto riguarda invece l'analisi delle qualificazioni che non riguardino le peculiarità costruttive, geografiche, cronologiche ed esecutivo-stilistiche è interessante osservare dal punto di vista semantico e testuale come si costruiscono una sorta di profili di personalità di ciascuno strumento, analizzabili anche in famiglie di strumenti (cordofoni, membranofoni, ecc.) dei quali è possibile non solo individuare qualitativamente il carattere, ed eventualmente confrontarlo con analoghi usi in contesti specialistici, ma anche stimare quantitativamente da una parte la varietà e ricchezza, dall'altra individuare quali sono le qualificazioni che orizzontalmente risultano più usate per descrivere il carattere e il suono di uno strumento. Questa analisi permette inoltre di mettere in luce quali siano le strategie più comuni per cogliere o attribuire senso, un senso verbalmente esprimibile, ai caratteri del suono mediante processi metaforici e di trasferimento di caratteristiche di stato d'animo dall'ascoltatore allo strumento produttore del suono.

## **BIBLIOGRAFIA:**

- ALLORTO, R. e A. FERRARI (1959). *Dizionario di musica*. Milano, Ceschina.
- ARISTOMENIS, T., F. NIKOS, et al. (2002). Comparative evaluation of collocation extraction metrics.
- BARONI, M. e A. KILGARRIF (2006). *Large linguistically-processed Web corpora for multiple languages*. Proceedings of the Eleventh Conference of the European Chapter of the Association for Computational Linguistics:, Association for Computational Linguistics 87-90.
- BASSO, A. (1983). *Dizionario enciclopedico universale della musica e dei musicisti. Il Lessico*. Torino, UTET.
- BENSON, M. (1989). *The Structure of the Collocational Dictionary*. "Int J Lexicography" 2(1): 1-14.
- BERTOLINI, C. e D. ZANNONI (1997). *Prontuario dei termini musicali*. Monza, Casa musicale Eco.
- BISETTO, A. (2004). Composti e polirematiche. *La formazione delle parole in italiano*. M. Grossmann and F. Rainer. Tübingen, Niemeyer: 56-69.

- BOCCAGNA, D. L. (1999). *Musical terminology : a practical compendium in four languages*. Stuyvesant, NY, Pendragon Press.
- BRACCINI, R. (2007). *Musica. Dizionario terminologico italiano, francese, inglese, tedesco*. Bologna, Zanichelli.
- CACCIARI, C. (1989). *La comprensione delle espressioni idiomatiche, il rapporto tra significato letterale e significato idiomatico*. "Giornale italiano di psicologia" 16: 413 - 37.
- CACCIARI, C., F. Vespignani, et al. (2008). *Aspettative semantiche ed espressioni idiomatiche: aspetti psicolinguistici ed evidenze elettrofisiologiche. Neuropsicologia della comunicazione: 139-61*.
- CALBI, O. (1992). *A-Z lessico della musica : teoria della musica in ordine alfabetico*. Napoli, Simeoli.
- CALZOLARI, N., C. Fillmore, et al. (2002). *Towards best practice for multiword expressions in computational lexicons. Proceedings of the 3rd International Conference on Language Resources and Evaluation (LREC 2002)*. Las Palmas, Canary Islands: 1934-40.
- CANDÈ, R. d. (1968). *Dizionario di musica*. Milano, V. Bompiani.
- CASADEI, F. (1994). *La semantica delle espressioni idiomatiche*. "Studi italiani di linguistica teorica ed applicata" 1: 61-81.
- CASADEI, F. (1995). *Flessibilità lessico-sintattica e produttività semantica delle espressioni idiomatiche: un'indagine sull'italiano parlato. L'italiano che parliamo*. F. Casadei, G. Fiorentino and V. Samek-Lodovici. Santarcangelo, Fara Editore: 11-33.
- CASADEI, F. (1995). *Per una definizione di "espressione idiomatica" e una tipologia dell'idiomatico in italiano*. "Lingua e Stile" 2: 335 - 58.
- CASADEI, F. (1996). *Metafore ed espressioni idiomatiche: uno studio semantico sull'italiano*. Roma, Bulzoni.
- COOVER, J. (1958). *Music lexicography; including a study of lacunae in music lexicography and a bibliography of music dictionaries*. Denver, Bibliographical Center for Research, Rocky Mountain Region, Denver Public Library.
- COWIE, A. P. (1998). *Phraseology : theory, analysis, and applications*. Oxford, Clarendon Press.
- D'AGOSTINO, E. e A. ELIA (1998). *Il significato delle frasi: un continuum dalle frasi semplici alle forme polirematiche. Ai limiti del linguaggio*. F. e. a. Albano Leoni. Roma- Bari, Laterza: 287-310.

- De MAURO, T. (1999). *Grande Dizionario Italiano dell'uso*. Torino, UTET.
- De MAURO, T. (1999). Introduzione. *Grande Dizionario Italiano dell'uso*. T. De Mauro. Torino, UTET: VII-XLII.
- De MAURO, T. (1999). Postfazione. *Grande Dizionario Italiano dell'uso*. T. De Mauro. Torino, UTET: 1163-83.
- De MAURO, T. (2005). *La fabbrica delle parole : il lessico e problemi di lessicologia*. Torino, UTET libreria.
- De MAURO, T. e M. Voghera (1996). Scala mobile. Un punto di vista sui lessemi complessi. *Italiano e dialetti nel tempo. Saggi di grammatica per G. Lepschy*. B. Paola, G. Cinque, T. De Mauro and N. Vincent. Roma, Bulzoni: 99-131.
- DELLA CORTE A. e G. M. Gatti (1930). *Dizionario di musica*. Torino [etc.], G. B. Paravia & c.
- EGGBRECHT, H. H. (1955). *Studien zur musikalischen Terminologie*. Mainz,, Verlag der Akademie der Wissenschaften und der Literatur; in Kommission bei F. Steiner.
- EGGBRECHT, H. H. (1972). *Handwörterbuch der musikalischen Terminologie*. Wiesbaden,, F. Steiner.
- EGGBRECHT, H. H. a. (1981). Zur Wissenschaft der auf Musik gerichteten Wörter. *Cultural and historical aspects of musical terminology*: 776-8.
- ELIA, A., E. D'AGOSTINO, et al. (1985). Tre componenti della sintassi italiana: frasi semplici, frasi a verbo supporto e frasi idiomatiche. *Atti del XVII congresso internazionale della Società di linguistica italiana*, Bulzoni: 311-26.
- EVERT, S. (2004). *The Statistics of Word Cooccurrences: Word Pairs and Collocations*. Stuttgart, University of Stuttgart.
- EVERT, S. e B. KRENN (2001). Methods for the qualitative evaluation of lexical association measures. *Proceedings of the 39th Annual Meeting of the Association for Computational Linguistics*. Toulouse, France: 188 - 95
- FELBAUM, C. (1998). *WordNet : an electronic lexical database*. Cambridge, Mass, MIT Press.
- FELBAUM, C. (2007). *Idioms and collocations : corpus-based linguistic and lexicographic studies*. London ; New York, Continuum.
- FIRTH J. (1957). Modes of Meaning. *Papers in Linguistics 1934-51*. Oxford, Oxford University Press: 190-215.



- HALLIDAY, M. (1966). Lexis as a Linguistic Level. *In memory of J.R. Firth*. C. E. Bazell and J. R. Firth. London., Longmans: 148-63.
- HOANG, H. H., S. N. KIM et al. (2009). A re-examination of lexical association measures. *ACL/IJCNLP Multiword Expressions Workshop*: 31-9
- JACKENDOFF, R. (1997). *The architecture of the language faculty*. Cambridge, Mass., MIT Press.
- KILGARRIF, A. (2006). Collocationality (and how to measure it). *Proceedings of the 12th EURALEX International Congress*. E. Corino, M. C. and C. Onesti. Torino, Edizioni dell'Orso, Alessandria: 997-1004.
- KILGARRIF, A., P. Rychly, et al. (2004). The Sketch Engine *Proc. Euralex*. Lorient, France: 105-16.
- KILGARRIF, A. e D. TUGWELL (2001). WORD SKETCH: Extraction and Display of Significant Collocations for Lexicography. *Proc. ACL workshop on COLLOCATION: Computational Extraction, Analysis and Exploitation*. Toulouse: 32-8.
- LIMENTA, F. (1940). *Dizionario lessicografico musicale italiano-tedesco-italiano*. Milano., U. Hoepli.
- MURARO, M. T., Ed. (1995). *Le parole della musica. Studi sul lessico della letteratura critica del teatro musicale in onore di Gianfranco Folena, vol. II*. Firenze, Olschki.
- NAPOLI, E., Ed. (2006). *Dizionario dei termini musicali* Milano, B. Mondadori.
- NICOLODI F. e P. TROVATO, Eds. (1994). *Le Parole della musica. Studi sulla lingua della letteratura musicale in onore di Gianfranco Folena, vol. I*. Firenze, Olschki.
- NICOLODI, F. e P. TROVATO (1996). *Tra le note : studi di lessicologia musicale*. Fiesole [Italy], Cadmo.
- NICOLODI, F. e P. TROVATO, Eds. (2000). *Le parole della musica. Studi di lessicologia musicale, vol. III*. Firenze, Olschki.
- NICOLODI, F. e P. TROVATO (2000). *Studi di lessicologia musicale*. Firenze, Olschki.
- O'DELL, F. e M. McCARTHY (2008). *English collocations in use*. Cambridge ; New York, Cambridge University Press.
- PAPAGNO, C. (2008). Comprensione di espressioni idioma-tiche: evidenze neuropsicologiche. *Neuropsicologia della comunicazione*: 121-37.

- PEARCE, D. (2002). A Comparative Evaluation of Collocation Extraction Techniques. *Third International Conference on Language Resources and Evaluation*. Las Palmas, Spain: 1530-6.
- PETROVIC, S., J. SNAJDER, et al. (2006). *Comparison of collocation extraction measures for document indexing*. Information Technology Interfaces, 2006. 28th International Conference on, 451-6.
- PFISTER, M. (1993). *A Concise Dictionary of Musical Terminology*. "Zeitschrift für Romanische Philologie" 109: 1-2.
- POZZI, R. (1995). *Tendenze e metodi nella ricerca musicologica*. Atti del Convegno internazionale Latina, 27-29 settembre 1990. Firenze, Olschki.
- SINCLAIR, J. (1991). *Corpus, concordance, collocation*. Oxford, Oxford University Press.
- SINCLAIR, J. M. e R. CARTER (2004). *Trust the text : language, corpus and discourse*. London ; New York, N.Y., Routledge.
- SINCLAIR, J. M., S. JONES, et al. (2004). *English collocation studies : the OSTI report*. London ; New York, Continuum.
- STUBBS, M. (2002). *Two Quantitative Methods of Studying Phraseology in English*. "International Journal of Corpus Linguistics" 7(2): 215-44.
- TONANI, E. (2005). *Storia della lingua italiana e della musica : italiano e musica nel melodramma e nella canzone : atti del IV Convegno ASLI Associazione per la storia della lingua italiana, Sanremo, 29-30 aprile 2004*. Firenze, F. Cesati.
- VIETRI, S. (1986). *Lessico e sintassi delle espressioni idiomatiche*. Napoli, Liguori.
- VIETRI, S. (2004). *Lessico-grammatica dell'italiano*. Torino, UTET.
- VOGHERA, M. (1994). *Lessemi complessi: percorsi di lessicalizzazione a confronto*. "Lingua e stile" 28: 185-214.
- VOGHERA, M. (2004). *Le polirematiche. La formazione delle parole*. M. Grossmann and F. Rainer. Tübingen, Max Niemeyer Verlag: 56-68.
- ZIPF, G. K. (1949). *Human Behavior and the Principle of Least Effort*, Addison-Wesley Press.