

Appendice metodologica: il disegno della ricerca

Giovanni Di Franco

Publicato in G. Di Franco e S. Nobile, 1995, *L'Italia che si dispera e l'Italia che si innamora. Temi, valori e linguaggi in venticinque anni di canzone italiana*. Roma, Paper, pp. 167-195.

4.1 Premessa

Questo studio pionieristico sui testi delle canzoni italiane dell'ultimo ventennio (1971-1992) manca del sostegno che solitamente nella ricerca viene dal riferimento a lavori precedenti analoghi per obiettivi e per metodologia. Seppure negli ultimi anni siano emersi studi sulla produzione musicale - nelle sue varie articolazioni come, ad esempio, il rock, la musica leggera, il jazz, la musica popolare, la canzone politica, ecc. -, sulla musica come mezzo di comunicazione di massa, sulla fruizione e sul consumo musicale, mancano tuttavia ricerche empiriche sugli stessi temi, oppure, nei rari casi in cui sono state approntate, queste si configurano come casi isolati e parziali rispetto all'ambito spazio-temporale della ricerca.

Questo capitolo è rivolto a ripercorrere le fasi operative della ricerca condotta con la tecnica di analisi del contenuto sui testi delle canzoni italiane dell'ultimo ventennio. L'esplicitazione del percorso metodologico è invero il modo con cui vengono garantite la pubblicità, l'intersoggettività e la controllabilità dei risultati; ovvero quei pre-requisiti che soli rendono scientifico e oggettivo un lavoro di ricerca. Si intende, quindi, solo rendere conto della struttura concettuale e operativa elaborata nel concreto farsi di questa indagine.

4.2 La definizione dell'oggetto d'indagine

Come si è già detto, il nostro oggetto di studio è costituito dai contenuti testuali della produzione musicale italiana degli ultimi venti anni rivolta ad un pubblico di massa.

Questo studio ha come riferimento teorico-metodologico le discipline sociologiche e in particolare la sociologia delle comunicazioni di massa e la sociologia della musica. Uno dei problemi più propri delle discipline sociologiche risiede nella definizione del confine che separa lo studio scientifico dalla conoscenza del senso comune. La difficoltà che si ha nel tracciare questo confine è ben rappresentata in Italia dall'imbarazzo il cui si trova il commissario tecnico della nazionale di calcio al quale chiunque, all'interno dei milioni di appassionati, sportivi e tifosi e non solo tra i colleghi allenatori o, tutt'al più, tra i critici sportivi, si sente in grado di dare indicazioni, suggerimenti o magari, anche, prenderne il posto. Per chiudere brevemente questa digressione, che potrebbe benissimo esaurire le pagine di diversi libri, riteniamo di dovere ribadire che ciò che distingue la conoscenza sociologia dalla conoscenza del senso comune è il ricorso alla logica del metodo scientifico con le sue proprietà caratteristiche di pubblicità, intersoggettività e ripetibilità delle procedure.

Il processo logico-metodologico lungo il quale si è sviluppata la ricerca ha preso le mosse dalla definizione del problema oggetto d'indagine e dalla formulazione di una rete di ipotesi che ha

avuto una funzione di guida e di orientamento dell'intero *iter* della ricerca. Dall'individuazione della situazione problematica non può scaturire "immediatamente" la rilevazione empirica. E' necessario, infatti, superare una visione mistica e taumaturgica del potere euristico dei "dati" considerati come entità oggettive che "parlano da soli". Si può arrivare alla effettiva rilevazione dei dati solo dopo aver svolto una serie di operazioni che, nel loro insieme, costituiscono un processo di riduzione della complessità. Invero, la situazione problematica deve essere articolata in aree problematiche più specifiche, per ognuna si devono individuare concetti sottesi che poi sono scomposti in dimensioni concettuali costitutive. Per ogni dimensione concettuale si devono individuare gli indicatori migliori, sia ad un livello probabilistico che ad un livello semantico, e si arriva, infine, alle concrete operazioni di rilevazione empirica e di analisi dei dati.

4.3 Il disegno della ricerca

Dato il nostro oggetto d'indagine si è ritenuto importante partire dalla definizione del significato dei seguenti concetti-termini: 'testi', 'produzione musicale italiana' e 'pubblico di massa'.

Il primo concetto-termini è quello che crea meno problemi in quanto definisce la componente testuale delle produzioni musicali. La nostra ricerca ha preso in considerazione solo i testi composti prevalentemente in lingua italiana. Sappiamo infatti che la produzione musicale italiana si articola, dal suo nascere, anche in altre lingue, come possono essere considerati a pieno titolo i vari dialetti e su tutti il napoletano, che per tradizione storica sono precedenti all'uso della lingua italiana. C'è poi, in tempi più recenti, un frequente ricorso a lingue straniere, e tra queste prevale l'inglese, dovuto soprattutto all'influenza di alcuni generi musicali, ad esempio il rock, che ha spinto gli autori di testi all'uso di altre lingue ritenute più idonee per il tipo di musica proposta. Nella nostra ricerca, per esigenze di delimitazione dell'oggetto di studio, si è preferito analizzare solo i testi nei quali l'uso dell'italiano fosse prevalente, se non esclusivo.

Crea maggiori problemi il concetto 'produzione musicale italiana'. Ciò è dovuto ad una pluralità di fattori. Sin dalla nascita dell'industria discografica italiana e la conseguente produzione musicale di massa, sono stati conati moltissimi termini che, nel corso del ventennio preso in considerazione, hanno via via identificato stili, generi, autori, tendenze e mode. Ad accrescere la proliferazione di termini e di significati ha poi contribuito la critica e la stampa specializzata nell'intento di produrre classificazioni funzionali ad analisi critiche, valutazioni e ricostruzioni di tendenze evolutive della musica, nonché tutti i fenomeni, con relative nuove etichette semantiche, nati ai margini dell'industria musicale, che sono stati successivamente riassorbiti da quest'ultima o sono rimasti tendenze *underground* di ristrette *élite* di cultori. Accresce la confusione, infine, anche il fenomeno di trasformazione, estensione e stiramento di significato che alcuni termini, nel corso del tempo, hanno subito.

Effettuare una disamina della produzione delle molte etichette semantiche associate alla produzione musicale italiana non rientra negli scopi della nostra ricerca; abbiamo utilizzato, come criterio di selezione per l'individuazione dei materiali da sottoporre ad analisi, i dischi *long playng* (33 giri) che a partire dal 1971 (anno di avvio in Italia della compilazione delle classifiche di vendita per i dischi a 33 giri) sono entrati nelle classifiche di vendita in una posizione qualsiasi, dal primo al ventesimo posto, per almeno una settimana. In questo modo abbiamo potuto prescindere nella scelta dei dischi dai quali estrarre un campione rappresentativo, da differenze di genere

(musica leggera, cantautori, rock, e così via) e da presunte diversità, qualitative e implicitamente connotate da giudizi di valore, rispetto alla diversificazione delle produzioni musicali circa il loro potenziale più o meno commerciale.

Rimane, infine, da definire il concetto-termine 'pubblico di massa'. Con esso vogliamo definire i fruitori-consumatori, intesi nel senso più ampio possibile del termine, della produzione musicale che si esprime in lingua italiana. In questo primo studio, quindi, non sono state prese in considerazione tutte le produzioni musicali che hanno segmentato il pubblico in fasce e settori delimitati rispetto all'intero mercato discografico nazionale. Le ragioni che presiedono a tale scelta sono di ordine diverso, ma, in ultima analisi, sono da connettere alla considerazione di questa come una delle varianti in cui si manifesta la produzione simbolica della cultura di massa e in questo senso possiamo omologarla ad altri settori espressivi come il cinema, la televisione, la letteratura.

Le finalità della nostra indagine sono di tipo esplorativo, descrittivo e comparativo piuttosto che esplicativo, verificativo. Si intendono individuare, sulla base di riscontri empirici, le caratteristiche salienti della produzione musicale italiana di massa lungo un ventennio ovvero: descrivere l'offerta delle produzioni musicali nei termini delle caratteristiche morfologico-strutturali dei dischi e dei brani; dei contenuti tematici dei brani ossia degli argomenti-temi trattati e della loro frequenza; delle modalità qualitative cui si sviluppano i temi delle canzoni; delle forme linguistiche caratteristiche dei testi; delle dimensioni valoriali veicolate nelle canzoni.

Si è ipotizzato che tali obiettivi fossero dipendenti dalle seguenti variabili-caratteristiche:

1. genere degli interpreti delle canzoni;
2. genere degli interpreti filtrato dal periodo storico;
3. periodo storico.

Per la prima variabile indipendente, si è prodotta la seguente classificazione degli interpreti:

- **interpreti di genere melodico:** in questa classe rientrano le produzioni che hanno come caratteristica peculiare la forma canzone con una forte enfasi melodica;
- **interpreti di genere cantautore:** in questa classe rientrano le produzioni connotate dalla pluralità di ruoli (autore, compositore e interprete) degli interpreti e dall'attribuzione-riconoscimento, da parte degli altri musicisti, degli operatori del settore, dei critici, ecc., di questa pluralità di ruoli. In altri termini, lo status di cantautore non è definibile solo sulla base delle caratteristiche degli interpreti, ma necessita anche di un riconoscimento da parte dell'ambiente musicale nel suo complesso (industria musicale, pubblico, stampa specializzata) per questo, ad esempio, Umberto Tozzi, che è interprete e compositore dei propri brani non è ritenuto un cantautore, mentre Lucio Battisti, a sua volta interprete e compositore, lo è a pieno titolo;
- **interpreti di genere rock:** qui si collocano le produzioni che sono state influenzate dai vari sottogeneri in cui si è sviluppata la musica rock dalla seconda metà degli anni '50 in poi (rock and roll, beat, rock progressivo, hard rock, heavy rock, ecc.).

Per la seconda variabile indipendente si è prodotta più che una classificazione una tipologia in quanto essa risulta come prodotto della prima classificazione con il periodo storico, a sua volta distinto in due classi, anni settanta e anni ottanta. La tipologia è articolata in:

- **grandi interpreti:** in questa classe rientrano gli interpreti di genere melodico che hanno iniziato la loro carriera artistica a partire almeno dagli anni settanta e che sono tuttora in attività (rientrano in questa classe anche interpreti come Mina e Celentano, solo per fare qualche nome, che hanno esordito nel mercato discografico molti anni prima);
- **le meteore:** interpreti di genere melodico che hanno conosciuto un breve periodo di popolarità (ad esempio Sandro Giacobbe, Umberto Balsamo, Frà Cionfoli, ecc);

- cantautori anni settanta** (es. Fabrizio De Andrè, Francesco Guccini; Francesco De Gregori, Antonello Venditti, Lucio Dalla, ecc.);
- **cantautori anni ottanta** (es. Pino Daniele, Franco Battiato, Luca Barbarossa; Marco Masini, Paolo Vallesi, ecc.);
- **rock anni settanta**; (es. Area, Banco del Mutuo Soccorso, Premiata Forneria Marconi, Orme, Delirium, Formula tre, ecc.);
- **rock anni ottanta**: (es. Vasco Rossi, Zucchero, Litfiba, Ligabue, ecc.).

Il periodo storico è articolato in periodi quinquennali.

Per l'analisi si sono individuate le seguenti cinque aree problematiche:

1. morfologia dell'album;
2. morfologia del brano;
3. argomento del testo;
4. caratteristiche linguistiche del testo;
5. dimensioni valoriali del testo.

Queste aree sono state ristrutturare in dimensioni concettuali e indicatori ritenuti in ipotesi rilevanti per il nostro studio.

4.4 Lo strumento per la raccolta dei dati

L'indagine condotta attraverso la tecnica dell'analisi del contenuto è stata avviata dopo la definitiva messa a punto di una scheda di rilevazione che doveva essere in grado di rilevare gli aspetti che erano stati considerati rilevanti ai fini degli obiettivi. Per analisi del contenuto, seguendo il Krippendorff, si intende "una tecnica di ricerca atta a stabilire inferenze valide e replicabili a partire dai dati per arrivare al relativo contesto" (Krippendorff, 1983, p. 37). In questa definizione l'analisi del contenuto è uno strumento di ricerca empirica e in quanto tale esso comporta l'uso di procedure standardizzate applicate a dati empirici al fine di produrre "elementi di conoscenza, nuove capacità di comprensione, una rappresentazione di fatti e una guida per l'azione" (Krippendorff, 1983, p. 37). Come ogni strumento scientifico esso deve possedere gli attributi della validità, dell'attendibilità e della replicabilità.

Lo strumento cruciale per la nostra analisi del contenuto è quindi la scheda di rilevazione. Questa è un insieme logicamente organizzato di indicatori finalizzati a rilevare i contenuti, in ipotesi connessi al problema oggetto di indagine, che sono associati o corrispondenti alle dimensioni operativamente individuate. Ogni domanda contenuta nella scheda corrisponde ad una matrice di ipotesi posta in relazione ad un insieme formalizzato di concetti, a loro volta articolati in dimensioni, di cui le domande stesse sono i possibili indicatori.

Nel presente paragrafo si intendono presentare tutte le domande contenute nella scheda di rilevazione da noi utilizzata, che hanno condotto, dopo progressivi affinamenti e ritocchi, alla stesura definitiva, in termini di definizione operativa degli indicatori e quindi classificandone le modalità di rilevazione di questi, dello strumento di rilevazione. Le aree problematiche in cui si presenta suddivisa la scheda sono state individuate a partire dall'analisi del problema; la sua elaborazione è seguita, infatti ad una fase di *scouting*, nella quale sono state raccolte informazioni preliminari attraverso la ricognizione e l'analisi della letteratura sull'argomento (cfr. cap. 1). Le informazioni raccolte in questa fase esplorativa hanno permesso di stimolare intuizioni utili alla

formulazione di ipotesi e di individuare, in un secondo momento, a partire da concetti più generali, gli indicatori più adeguati allo specifico concetto di indagine. Gli indicatori sono definibili come concetti più specifici attraverso i quali il ricercatore instaura un rapporto di rappresentazione semantica con un concetto più generale (Marradi, 1980). In questo processo la scheda di rilevazione è uno strumento di rilevazione e di registrazione di stati su proprietà individuate attraverso traduzioni di indicatori in variabili: la procedura di definizione o riduzione dei concetti teorici in concetti osservativi (Pasquinelli, 1964) risulta indispensabile per conferme/disconferme di teorie e di ipotesi (Cannavò, 1984). Nel processo di operativizzazione, l'attenzione al contesto di indagine è messa in luce proprio da Marradi (1980) che sottolinea l'assenza di univocità concetto-indicatore e, in ordine a questo rapporto, la necessità di un riferimento al contesto culturale e sociale. La scelta degli indicatori più adeguati a rappresentare semanticamente un concetto in un contesto spazio-temporale preciso nasce dalla pluralità di significati che un indicatore, in quanto simbolo, può assumere in situazioni socio-culturali diverse. Quindi, continua Marradi, la scelta di indicatori dipende dalla conoscenza del contesto di analisi e dall'oggetto su cui si svolge la ricerca. Pertanto la ricerca di fondo ha consentito di realizzare un'appropriata connessione dei concetti teorici utilizzati nella letteratura del fenomeno da noi studiato, con gli indicatori, collegando, quindi, la rete teorica con la realtà empirica posta in osservazione.

Alla stesura della scheda si è arrivati al termine di un lungo percorso di elaborazione teorica iniziato con la tematizzazione dei concetti ed il reperimento di indicatori ad essi adatti. Per la messa a punto della scheda e per la sua socializzazione nel gruppo di analisti, è stato necessario un lungo periodo di *pretest* che ha permesso di tarare lo strumento e di controllarne al tempo stesso l'efficacia euristica. Per questa fase di collaudo si è lavorato in *équipe* con gli analisti per poter omogenizzare i criteri di rilevazione dei dati, che in ogni analisi del contenuto devono essere convenzionalmente fissati. Uno dei problemi metodologici affrontati nella fase di *pretest* ha riguardato la progressiva standardizzazione delle modalità di risposta alle varie domande presenti nella scheda di rilevazione. Si è progressivamente passati da una destrutturazione degli *items* ad una più elevata strutturazione. A questo scopo si è reso necessario analizzare un centinaio di canzoni diverse estrapolate da dischi di interpreti diversi per far emergere, fin dall'inizio, in maniera più chiara possibile le diversità stilistiche, di contenuto e di forma delle diverse produzioni musicali italiane.

Dal punto di vista tematico e strutturale la scheda di rilevazione utilizzata è stata suddivisa in cinque aree problematiche:

1. morfologia dell'album;
2. morfologia del brano;
3. argomento del testo;
4. caratteristiche linguistiche del testo;
5. dimensioni valoriali del testo.

Tale suddivisione si è ritenuta funzionale a fini descrittivi; ogni area problematica può essere suscettibile di "reagire" con tutte le altre e solo per comodità espositiva queste saranno illustrate separatamente.

1. Morfologia dell'album: l'articolazione di questa prima area problematica, dà luogo alle seguenti sotto aree espresse in termini di domande:

- D01 autore album,
- D02 titolo album,
- D03 anno pubblicazione album,
- D04 tipologia dell'album,

D05 composizione dell'album.

La scheda di rilevazione prevede nella sua prima sezione la raccolta di informazioni relative al disco e può essere assimilata alla raccolta di dati strutturali (anche detti "fissi") rispetto agli individui (unità d'analisi) su cui si vogliono raccogliere delle informazioni. La prima domanda definisce il nome dell'interprete del disco, la seconda il titolo dello stesso. In entrambi i casi, si sono lasciate aperte queste due prime domande considerando ogni autore e ogni titolo di disco come modalità diverse di due variabili lessicali. Per quanto riguarda l'anno (D03) di pubblicazione del disco gli analisti devono riportare nell'apposito spazio (vedi allegato 1) le ultime due cifre dell'anno di pubblicazione.

La quarta domanda classifica i dischi rispetto alla loro dimensione (singoli-doppi), che rappresentano le modalità più frequenti per quanto riguarda il mercato discografico italiano. Infatti, l'emissione di dischi di formato più ampio (come ad esempio i dischi tripli o quadrupli) è molto rara e, quando avviene, a causa del loro alto prezzo, non ne permette il loro accesso nelle classifiche di vendita. L'ultima caratteristica rilevata sull'intero disco (D05), riguarda la natura dei brani che compongono il disco. Rispetto questa caratteristica i dischi sono stati classificati in: a) raccolta di brani inediti, b) raccolta di brani editi, c) raccolta di brani eseguiti dal vivo e d) raccolta che comprende sia brani inediti che brani editi.

2. Morfologia del brano:

D06 titolo brano,

D07 posizione del brano all'interno del disco,

D08 tipologia del brano,

D09 modalità di esecuzione delle parti vocali del brano,

D10 presenza del ritornello,

D11 quantificazione delle parti cantate sul totale del brano.

Il primo *item* di quest'area (D06) prevede la registrazione del titolo del brano analizzato analogamente a quanto previsto per il titolo del disco, il secondo (D07) la sua collocazione (posizione del brano numerato progressivamente a partire a 1 fino a n) all'interno del disco. L'*item* D08 costituisce il filtro per l'analisi dei brani. Infatti, con questa domanda si individua se il testo del brano è in misura prevalente in lingua italiana o se è un brano strumentale o in dialetto o in altre lingue. Ai fini della nostra ricerca, come già detto, *si sono analizzati solo i brani che hanno testi, totalmente o in prevalenza, in lingua italiana* per cui tutti i brani che non possiedono questa caratteristica non sono stati analizzati. La domanda D09 riguarda le modalità di esecuzione delle parti testuali che sono state classificate in: a) una sola voce, b) due voci, c) tre o più voci e d) in forma corale. Le ultime due domande (D10 e D11) di questa sezione individuano se il testo ha una sua sezione o parte che ha funzioni di ritornello (vedi glossario) e il rapporto tra le parti cantate e le parti strumentali (quale delle due componenti prevale in termini temporali) nell'economia del brano.

3. Argomento del testo:

D12 individuazione del tema del testo,

D13 individuazione della forma attraverso la quale viene sviluppato il testo,

D14 rispetto ai verbi usati nel testo in quale persona prevalentemente sono coniugati,

D15 il tema del testo viene trattato da un punto di vista individuale o collettivo,

D16 livello di argomentazione del tema,

D17 individuazione di un destinatario,

D18 il testo viene presentato come una storia ecc,

D19 contesto denotativo,

D20 contesto connotativo,

D21 connotazioni del brano (differenziale semantico), D22 funzione della musica rispetto al testo.

La terza sezione della scheda di rilevazione definisce quale è l'argomento-tema del testo e in quali modalità questo si sviluppa. Attraverso la D12 si classificano gli argomenti-temi in sei classi: a) amore, b) altri rapporti affettivi, c) sfera personale, d) sfera sociale, f) altri temi-argomenti e g) assenza di un argomento-tema riconoscibile. L'interpretazione dell'argomento-tema costituisce il nodo cruciale della nostra analisi. L'interpretazione di un testo è un'operazione tutt'altro che semplice e univoca. Contributi da altre discipline hanno sottolineato il processo di relazioni tra testo, autore e fruitore-ascoltatore, ponendo il problema di come sia possibile per un analista-ricercatore decodificare il testo attribuendogli gli stessi significati dell'autore. Per i nostri scopi, non siamo interessati ad analisi sui significati "profondi" e "ultimi" dei testi, così come non interessano tutti i risvolti interpretativi che rimandano ai contenuti latenti dei testi, anche perché il tipo di materiali analizzato, quasi sempre, si caratterizza per uno stile comunicativo immediato, semplice e sintetico. Un testo di un brano musicale, per sue caratteristiche strutturali, ossia per essere inserito in una forma musicale definita che si sviluppa nell'arco di pochi minuti, deve arrivare direttamente e colpire, ovvero risvegliare l'interesse, l'ascoltatore. In questi termini abbiamo inteso l'interpretazione dell'argomento-tema del brano, ovvero la definizione dell'oggetto tematico sul quale viene sviluppato e costruito il testo. La prima classe tematica include tutti i testi che riguardano aspetti, situazioni, persone, comportamenti, sensazioni connesse a situazioni di rapporti d'amore. Sono state individuate, per questa prima classe, sei sotto-classi definite nel seguente modo:

1. amore in generale =riflessioni, interpretazioni, considerazioni sviluppate sull'amore senza citare un rapporto specifico d'amore;
2. fine di un rapporto = resoconti descrittivi di un amore finito;
3. malesseri d'amore = malesseri, disagi, sofferenze, provocate da un rapporto d'amore, questa classe include i temi del tradimento e della gelosia;
4. rapporto con il partner = resoconti descrittivi di un rapporto amoroso;
5. esaltazione del partner = il testo è rivolto al partner per enfatizzarne il ruolo, la presenza, le caratteristiche, ecc.;
6. erotismo/sensualità = il tema amore viene trattato, in modo esplicito o implicito, dal punto di vista del rapporto erotico- sensuale.

La classe degli altri rapporti affettivi è stata specificata nelle seguenti tre sotto-classi:

1. altri rapporti affettivi in generale = riflessioni, interpretazioni, considerazioni sui rapporti affettivi svincolate da un legame affettivo in particolare;
2. amicizia = il tema del testo è un legame di amicizia individuabile;
3. rapporti intergenerazionali = il tema riguarda a rapporti genitori-figli o figli-genitori.

La classe della sfera personale comprende i testi che sono incentrati sulla personalità dell'interprete e si articola in cinque sotto-classi:

1. autoreferenzialità = l'interprete parla di se stesso in maniera descrittiva-interpretativa;
2. malesseri disagi esistenziali = si rende conto di problematiche che sono connesse a difficoltà esistenziali ad esempio, solitudine, noia, insoddisfazione, ecc.;
3. riflessioni diacroniche = si sviluppa una riflessione personale in rapporto al tempo (passato-presente-futuro) assunto come dimensione prevalente;
4. relazioni interpersonali = le riflessioni personali sono sviluppare in relazione ai rapporti con gli altri;
5. rapporto con la religione = si analizza il rapporto sè- religione.

Nella classe della sfera sociale rientrano i temi che hanno un impatto e una rilevanza sociale:

1. sfera sociale in generale = analisi, commenti, interpretazioni in generale;
2. emergenze sociali = in questa classe ricadono i temi connessi con problemi di rilevanza sociale come la droga, l'aborto, la violenza, l'ecologia, ecc.;
3. umanità rispetto degli uomini = si tematizza il rispetto delle persone e si valorizza la vita;
4. religione = esame della religione a un livello sociale;
5. progresso evoluzione sociale = si considerano gli aspetti, sia positivi che negativi, legati al concetto di progresso- cambiamento sociale;
6. pacifismo antimilitarismo = si tematizza il problema della pace e l'inutilità della guerra;
7. emarginazione sociale = l'oggetto dei testi sono persone o gruppi caratterizzati da situazioni di discriminazione sociale;
8. politica = analisi, valutazioni di problemi o personaggi politici.

La quinta classe comprende gli argomenti-temi residuali, che non rientrano nelle precedenti classi, e si articola in:

1. luoghi = descrizione di luoghi fisici (città, nazioni, ecc.);
2. musica = si parla del mondo, del mestiere e di altri aspetti connessi con la musica;
3. biografie = descrizioni o resoconti su personaggi celebri o anonimi;
4. viaggi = si parla di viaggi in senso generale;
5. altri argomenti = altri argomenti che non rientrano nelle precedenti classi.

Per la definizione dell'argomento-tema del testo si è anche considerata la possibilità di rilevare una assenza, questo è il caso di testi costruiti su *nonsense* o su giochi linguistici onomatopeici.

Dopo aver individuato l'argomento del testo si passa all'analisi delle modalità di trattazione. La D13 permette la classificazione della forma (dialogica, personale, impersonale e corale) attraverso la quale si sviluppa il tema, mentre la D14 definisce le persone che sono usate prevalentemente per la coniugazione dei verbi. La D15 permette di stabilire se l'argomento-tema viene trattato a un livello individuale piuttosto che collettivo, mentre la domanda successiva classifica il tipo di trattazione più o meno analitica dell'argomento-tema (in termini di descrizione, interpretazione/analisi e individuazione di soluzione). La D17 definisce se nel testo è esplicitamente citato un destinatario e cioè una persona alla quale ci si rivolge, la D18 classifica la presentazione del testo nel modo seguente: a) come una storia-racconto, b) come una unica immagine, c) come una serie di riflessioni o una serie di ricordi, d) come un discorso diretto. Le domande 19 e 20 classificano i contesti denotativi e connotativi del testo; il primo, quando presente, si articola in a) svago-tempo libero, b) lavoro, c) scuola, d) politico, e) fantastico e f) altro; il secondo, quando presente, in a) famiglia, b) gruppo dei pari, c) relazioni con il partner e d) altro.

La D21 è una batteria di differenziale semantico che, in ipotesi, rende possibile la rilevazione di dimensioni soggiacenti che scaturiscono dall'interazione tra testo, musica e interpretazione. L'esigenza di non limitare l'analisi al solo contenuto testuale dei brani, è dovuta alla forma specifica di questo tipo di produzione culturale. Infatti, il testo di una canzone nasce proprio per essere accompagnato, o per accompagnare, una linea melodica e assume un significato in questo rapporto simbiotico. Altro significato si aggiunge poi con l'interpretazione che il cantante dà a questo testo-musicale e quindi, in ultima analisi, il senso di una canzone si acquisisce dalla simultanea percezione di una componente testuale, una musicale e una interpretativa. Per tentare di cogliere l'interazione dei tre elementi che costituiscono una canzone, si è predisposta la scala di differenziale semantico a cui gli analisti dovevano reagire dopo aver ascoltato il brano. Questa è

formata da 26 coppie di aggettivi opposti (vedi allegato 1) che sono separati da sette spazi, con la categoria centrale che rappresenta la neutralità.

L'ultima domanda di questa sezione (D22) insiste sulla componente musicale del brano in rapporto al testo, e classifica il rapporto musica testo nel seguente modo: a) descrizione = la musica segue fedelmente le parole si connota in termini di accompagnamento, b) enfaticizzazione = la musica produce un effetto di amplificazione e/o drammatizzazione del testo, c) attenuazione = la musica produce un effetto contrario rispetto al punto b, d) banalizzazione= la musica produce un effetto di perdita totale di credibilità del testo.

4. Caratteristiche linguistiche del testo:

D23 individuazione della/e frasi caratteristiche del testo,

D24 individuazione di espressioni o termini gergali nel testo,

D25 individuazione di espressioni o termini dialettali nel testo,

D26 individuazione di espressioni o termini in lingua straniera nel testo,

D27 individuazione di espressioni coprolaliche nel testo,

D28 altri riferimenti caratteristici del testo,

D29 individuazione di riferimenti a località geografiche,

D30 individuazione di espliciti riferimenti temporali,

D31 individuazioni a oggetti, simboli, attività modali tipiche.

In questa sezione della scheda si sono analizzati i testi rispetto al loro contenuto linguistico, il primo dato da rilevare (D23) è una domanda aperta nella quale gli analisti dovevano registrare, a loro giudizio, le frasi caratteristiche del testo. Anche in questo caso, si è progettata una variabile lessicale nella quale ogni risposta rappresenta una modalità di risposta diversa. Le domande successive permettono di cogliere la presenza nel testo di parole-frasi gergali (D24), parole-frasi dialettali (D25), parole-frasi in lingua straniera (D26) e parole-frasi coprolaliche (D27). La domanda 28 permette di specificare se nel testo sono presenti riferimenti all'arte, alla filosofia, alla letteratura, alla scienza, ecc. Le ultime tre domande di questa sezione indicano la presenza di riferimenti a località geografiche (D29), a chiare dimensioni temporali (passato, presente, futuro, D30) e a oggetti, simboli, attività modali tipiche di determinati periodi storici (D31).

5. Dimensioni valoriali del testo:

D32 individuazione dei riferimenti valoriali.

Nell'ultima sezione della scheda di rilevazione si dovevano individuare i valori che erano sollecitati nel testo rispetto all'argomento-tema trattato (vedi allegato 1) fino ad un massimo di cinque valori diversi.

4.5 La costruzione del campione

Lo scopo delle indagini campionarie è quello di permettere l'estendibilità dei risultati di una rilevazione condotta su una frazione di una determinata popolazione all'intera popolazione detta anche "universo" che in realtà è oggetto di interesse. Si ricorre a indagini di questo tipo ogni qual volta che le risorse economiche, di tempo e le caratteristiche della popolazione non permettono rilevazioni esaustive ed è questo il caso di gran parte delle ricerche delle discipline sociali ed in particolare della sociologia dove quasi sempre le popolazioni che si sottopongono a indagine sono di numerosità molto elevata.

4.6 Il piano di campionamento

Se per campionamento intendiamo il procedimento di individuazione del campione, il piano di campionamento è il processo di elaborazione del procedimento stesso. In questa attività è necessaria l'attenta considerazione dei vari problemi connessi al campionamento.

Perché un campione sia rappresentativo si pone subito un problema di definizione della popolazione da considerare, e dell'ampiezza del campione.

Non è facile valutare il comportamento degli uomini, e soprattutto nelle scienze sociali necessita da parte del ricercatore una profonda conoscenza della problematica dell'indagine e di tutte le questioni che possono esservi collegate: tale problematica va esaminata facendo riferimento alla propria esperienza, alla propria cultura, alla propria conoscenza del contesto socio-culturale entro il quale si colloca il fenomeno che è oggetto di interesse scientifico.

Il piano di campionamento viene, quindi, definito in relazione agli obiettivi dell'indagine e alla definizione della popolazione. E' necessario che gli scopi della ricerca siano ben precisati, per non incorre in valutazioni dispersive o errate, e per evitare che il campione risulti poco utile o addirittura estraneo ai temi di cui ci si sta occupando.

Oltre alla definizione della popolazione (cioè dell'insieme dei dati da cui viene tratto il campione) e all'ampiezza del campione stesso, si pongono vari problemi quali la scelta del periodo di realizzazione dell'indagine, la sua durata, la sua collocazione spaziale. Da valutazioni di questo genere può scaturire la scelta delle unità di campionamento. Tali unità possono rappresentare entità geografiche, oppure persone, o cose, o eventi, o tratti di risposte interpersonali, e così via; sarà proprio in relazione alla popolazione e agli scopi dell'indagine che esse verranno scelte e determinate in modo tale che non sussista dubbio se una certa entità debba farne parte oppure no.

L'individuazione delle unità di campionamento comporta l'elaborazione di metodi per la raccolta e la tabulazione dei dati; sarebbe impensabile programmare un piano di campionamento senza aver allestito un valido sistema di valutazione delle osservazioni.

Le considerazioni fin qui proposte evidenziano la necessità di una definizione concreta della popolazione (cioè delle persone, o cose o entità varie, che siano argomento dell'indagine) e che ne fissi gli elementi essenziali.

Nel caso della nostra indagine, in conformità degli obiettivi esposti in altra parte del presente lavoro, la popolazione di indagine è costituita dai dischi 33 giri aventi brani con testi in italiano, che sono entrati nelle classifiche di vendita, in un posto qualsiasi dal primo al ventesimo e per almeno una settimana, a partire dal 1-1-1971 fino al 31-12-1992.

Per l'operazione di campionamento si è dovuto procedere alla costituzione di un archivio che contenesse tutti i dischi che sono entrati nelle classifiche di vendita nel periodo considerato, dal primo al ventesimo posto. La scelta delle classifiche di vendita (Hit Parade) come criterio per l'individuazione del materiale empirico per l'analisi prescinde dalla validità di queste in merito alla loro composizione. E' noto infatti, che attraverso le classifica di vendita non si certificano in modo assoluto le vendite dei prodotti discografici (dischi, nastri, cd e altri supporti): a questo compito è istituzionalmente preposta la SIAE (Società Italiana Autori ed Editori), ma esse si configurano essenzialmente come mezzo di informazione e di promozione usato dalle case discografiche per la diffusione dei propri prodotti. Per gli scopi della nostra ricerca, le classifiche di vendita sono state usate come indicatori della popolarità dei dischi della musica italiana ed è per questo motivo che abbiamo usato, indipendentemente dalla posizione raggiunta e dal periodo di stazionamento, tutti i

dischi che sono entrati in classifica nel periodo considerato. Infatti rispetto all'offerta complessiva di musica italiana era necessario porre dei filtri per la definizione di quali materiali effettivamente analizzare ed il fatto che le classifiche di vendita sono uno dei principali canali di promozione dei prodotti discografici, ha fatto optare per questa soluzione, in quanto un disco entrato in classifica ha maggiori probabilità di essere noto al pubblico interessato rispetto ad un altro che non vi ha fatto ingresso.

Le fonti usate per la costituzione dell'archivio-popolazione dal quale campionare i dischi per l'analisi sono state una pubblicazione di Dario Salvatori (1989) che riporta le classifiche di vendita dei 33 giri dal 1971 al 1988, e, per gli anni dal 1989 al 1992, le classifiche riportate dal settimanale *Radio Corriere Tv*. Dall'archivio-popolazione, che in una prima fase conteneva tutti i dischi entrati in classifica, sono stati eliminati quelli che erano interamente strumentali e quelli, prevalentemente di interpreti stranieri, ma in qualche caso anche di interpreti italiani, i cui testi erano totalmente in lingua straniera. Dopo queste operazioni di filtro, l'archivio da cui estrarre il campione risulta composto da 692 dischi.

In accordo con gli obiettivi dell'indagine e con le procedure elaborate per la selezione e la raccolta dei dati, si deve decidere quale tipo di campionamento convenga realizzare; tra i metodi probabilistici più usati, nell'ambito delle discipline sociali ricordiamo il campionamento casuale semplice, il campionamento casuale sistematico, il campionamento stratificato semplice, il campionamento stratificato a più stadi, il campionamento per grappoli.

Nella nostra indagine si è optato per la scelta di un campionamento casuale semplice in quanto la disponibilità di un archivio esaustivo della popolazione forniva la massima garanzia di rappresentatività del campione estratto in quanto il procedimento di scelta garantisce a tutti i dischi (692) dell'archivio la stessa probabilità di essere estratti.

Per il problema della numerosità sufficiente del campione si è assunto un errore tollerato del $\pm 5\%$ ovvero con un livello di fiducia del 95%. La formula per stabilire la numerosità del campione è la seguente:

$$n = \frac{n^\circ}{1 + (n^\circ - 1)/N}$$

dove

$$n^\circ = \frac{Z^2 * P * Q}{e^2} = \frac{2,43^2 * (0,79 * 0,21)}{0,05^2} = \frac{5,9049 * 0,1659}{0,0025} = \frac{0,9796}{0,0025} = 391,85$$

con $e=0,05$ che indica l'errore tollerato del 5%; $Z = 2,43$ ossia l'ascissa della curva normale corrispondente alla probabilità di estrarre un campione non rappresentativo dell'5% e $P \times Q = 0,1659$ la stima della variabilità della popolazione rispetto alla distribuzione dischi di autori solisti-autori gruppi (complessi), dove abbiamo frequenze pari al 79% per i solisti e al 21% per i gruppi.

Sostituendo n° nella formula della numerosità campionaria otteniamo:

$$n = \frac{391,85}{1 + \frac{391,85 - 1}{692}} = \frac{391,85}{1 + 0,5648} = 250,4$$

La numerosità del nostro campione è quindi di 250 dischi che rappresentano il 36% della popolazione. La numerosità del campione è stata elevata a 258 unità (37% dell'universo) prevedendo un tasso di mortalità dell'1%, ovvero la possibilità di non riuscire a reperire qualche disco del campione cosa che peraltro, a posteriori, non si è verificata in quanto il reperimento dei dischi entrati a far parte del campione che non sono più distribuiti dalle case discografiche, sono stati reperiti presso la Discoteca di Stato in Roma.

4.7 Un controllo della rappresentatività del campione

Per stimare la rappresentatività del campione e quindi la possibilità di estendere i risultati ottenuti all'intera popolazione entro il livello di fiducia prefissato, è stato effettuato un confronto delle distribuzioni tra le unità della popolazione e quelle del campione relativamente a quelle variabili o caratteristiche della popolazione note.

Una buona rappresentatività del campione è valutabile, infatti, solo dopo aver studiato gli scostamenti tra le distribuzioni di frequenza nel campione rispetto alle stesse variabili note nella popolazione. E' chiaro che tanto più piccoli sono tali scostamenti e tanto più valido e/o rappresentativo potrà essere considerato il campione. In altre parole, possiamo ritenere un campione rappresentativo se tutte le possibilità e tutte le modalità in cui i caratteri considerati nella popolazione (ovvero le caratteristiche note della popolazione) sono presenti nel campione nelle giuste proporzioni con un errore ritenuto accidentale (dovuto al caso). Convenzionalmente gli scostamenti tollerati tra le due distribuzioni sono circa di cinque punti percentuali (sia in positivo che in negativo).

Le caratteristiche della popolazione di cui eravamo a conoscenza perché presenti nel nostro archivio erano le seguenti:

1. anno di ingresso nelle classifiche di vendita;
2. classificazione dell'interprete del disco;
3. classificazione del genere musicale;
4. piazzamento raggiunto in classifica di vendita;
5. settimane di stazionamento in classifica.

Con riguardo alla distribuzione per l'anno di ingresso dei dischi in classifica si registra tra la distribuzione campionaria e quella della popolazione un indice di scostamento medio pari allo 0,43% (cfr. tab. 1), quindi mediamente gli scarti sono ben al di sotto dell'1%. Analizzando con attenzione gli anni, si nota che nel campione il 1974 e il 1989 sono lievemente sovra rappresentati

(rispettivamente del 2,8% e del 2,1%), mentre il 1981 e il 1982 sono leggermente sotto rappresentati (2,4% e 2,1%, cfr. tab. 1).

Tab. 1 Distribuzione dei dischi per l'anno di ingresso in classifica nella popolazione e nel campione

Anno	Popolazione	Campione	Scarto
1971	3,5	3,5	0,0
1972	3,6	2,3	- 1,3
1973	4,0	4,6	+ 0,6
1974	4,9	7,7	+ 2,8
1975	4,3	4,2	- 0,1
1976	4,0	3,5	- 0,5
1977	3,2	2,7	- 0,5
1978	4,9	4,2	- 0,5
1979	3,8	4,2	+ 0,4
1980	3,8	5,4	+ 1,6
1981	5,5	3,1	- 2,4
1982	4,8	2,7	- 2,1
1983	4,3	3,9	- 0,4
1984	3,8	3,1	- 0,6
1985	5,6	6,2	+ 0,6
1986	5,8	6,2	+ 0,4
1987	4,2	4,6	+ 0,4
1988	4,5	5,0	+ 0,5
1989	4,5	6,6	+ 2,1
1990	4,9	3,9	- 1,0
1991	6,1	6,2	+ 0,1
1992	6,1	6,2	+ 0,1

indice di scostamento medio=0,43%

La seconda caratteristica controllata, è stata usata nell'operazione di definizione della numerosità del campione (vedi paragrafo precedente), rappresenta la classificazione degli interpreti dei dischi rispetto allo status di solista maschio, solista femmina e gruppi. In questo caso l'indice di scostamento medio tra la distribuzione della popolazione e quella nel campione è appena apprezzabile (0,2%).

Tab. 2 Distribuzione dei dischi per la classificazione dell'interprete nella popolazione e nel campione

Interprete	Popolazione	Campione	Scarto
Solista Uomo	60,9	60,6	- 0,3
Solista Donna	18,1	18,7	+ 0,6
Gruppi	21,0	20,7	- 0,3

indice di scostamento medio=0,2%

A proposito della classificazione del genere musicale degli interpreti, nella popolazione troviamo la seguente ripartizione; genere melodico 48,4%, genere cantautore 41,6% e genere rock 10%. Nel campione si registra una lieve sovra rappresentazione dei cantautori a scapito degli altri due generi; infatti la distribuzione per genere musicale nel campione è: genere melodico 47,4%, genere cantautore 42,9% e genere rock 9,7% (vedi tab. 3).

Tab. 3 Distribuzione dei dischi per la classificazione di genere musicale nella popolazione e nel campione

Genere	Popolazione	Campione	Scarto
Melodico	48,4	47,4	- 1,0
Cantautore	41,6	42,9	+ 1,3
Rock	10,0	9,7	- 0,3

indice di scostamento medio=0,43%

Le ultime due caratteristiche note nella popolazione sono le posizioni raggiunte in classifica dai dischi e il tempo di stazionamento (in termini di settimane) nelle stesse. La tab. 4 mostra le due distribuzioni rispetto alla posizione raggiunta nelle classifiche, in questo caso si registra un indice di scostamento medio pari allo 0,42%.

Tab. 4 Distribuzione dei dischi per la posizione raggiunta in classifica nella popolazione e nel campione

Posizione	Popolazione	Campione	Scarto
1	15,3	16,6	+ 1,3
2	9,0	9,3	+ 0,3
3	6,2	8,9	+ 2,7
4	6,5	6,6	+ 0,1
5	6,2	6,9	+ 0,7
6	4,9	5,0	+ 0,1
7	4,9	2,3	- 2,6
8	5,8	5,4	- 0,4
9	3,9	3,9	0,0
10	4,2	3,5	- 0,7
11	3,6	1,5	- 2,1
12	3,3	3,9	+ 0,6
13	3,8	2,3	- 1,5
14	3,2	3,9	+ 0,7
15	3,2	2,7	- 0,5
16	2,0	1,9	- 0,1
17	2,7	2,7	0,0
18	3,8	4,2	+ 0,4
19	4,0	3,5	- 0,5
20	3,5	5,0	+ 1,5

indice di scostamento medio=0,42%

Nell'analisi delle distribuzioni per posizione notiamo che nel campione i dischi che hanno raggiunto il terzo posto in classifica sono leggermente sovra rappresentati (+2,7%), mentre sono sotto rappresentati i dischi che hanno raggiunto il settimo e l'undicesimo posto (rispettivamente - 2,6% e -2,1%).

L'ultima caratteristica nota della popolazione, il tempo di stazionamento dei dischi in classifica espresso in settimane, è stato ripartito in cinque classi temporali: 1) fino a dieci settimane, 2) da 11 a 20 settimane, 3) da 21 a 30 settimane, 4) da 31 a 40 settimane e 5) oltre 41 settimane. L'indice di scostamento medio per questa caratteristica è pari allo 0,54%, mentre nella distribuzione campionaria si nota una lieve sovra rappresentazione della classe 31-40 settimane (+1,4%) e una lieve sotto rappresentazione della classe 1-10 settimane (-1,8%; vedi tab.5).

Tab.5 Distribuzione dei dischi per il tempo di stazionamento in classifica nella popolazione e nel campione

Settimane	Popolazione	Campione	Scarto
1-10	23,8	22,0	- 1,8
11-20	32,0	32,1	+ 0,1
21-30	28,7	27,8	- 0,9
31-40	10,9	12,3	+ 1,4
oltre 41	4,6	5,8	+ 1,2

indice di scostamento medio=0,54%

Possiamo concludere che la rappresentatività del campione rispetto alla popolazione è da considerarsi ben al di sotto del limite (+ o - 5%) di scostamento tollerato tra le modalità dei caratteri nella popolazione e nel campione.

4.8 le modalità di rilevazione

Stabilita la numerosità della popolazione di riferimento e individuato il campione di dischi da analizzare, dopo aver costruito e testato la scheda di rilevazione (queste prime due fasi si sono svolte dal mese di gennaio al marzo 1993) si è proceduto alla fase di rilevazione vera e propria che è stata effettuata in quattro mesi dall'aprile al luglio del 1993. In questa fase gli analisti hanno lavorato individualmente sui dischi che via via venivano loro distribuiti con criteri rigorosamente casuali. Il modello procedurale per la rilevazione era il seguente: l'unità di analisi erano i singoli brani in cui si articolavano i dischi. Gli analisti ascoltavano, nell'ordine in cui questi apparivano, i brani dei dischi loro assegnati, anche più volte, e poi, brano per brano, compilavano una scheda di rilevazione. Settimanalmente tutti gli analisti si riunivano per consegnare le schede compilate e per prendere dei nuovi dischi da analizzare. In queste occasioni si verificava il lavoro svolto e si risolvevano, collegialmente, gli eventuali casi dubbi che i singoli analisti avevano riscontrato. Nei casi in cui si registravano delle controversie interpretative, soprattutto in ordine all'individuazione del tema-argomento del testo, si procedeva ad un ascolto collegiale dei brani che avevano posto dei problemi e, successivamente, ad una votazione che permetteva di risolvere, per maggioranza, i problemi.

4.9 la meccanizzazione dei dati

Conclusa la fase della rilevazione dei dati, si è proceduto alle operazioni di codifica dei dati e al loro controllo per verificare la corretta somministrazione della scheda di analisi da parte degli analisti. Per ciascuna variabile, a seconda della sua natura classificatoria, misurativa, di conteggio o lessicale, sono stati usati i codici alfa-numeriche adeguati. I dati codificati sono stati poi trascritti in forma matriciale (casi per variabili) e sono stati archiviati su supporto magnetico, attraverso un apposito *software*, in un *file-dati*. Ogni riga della matrice corrisponde ad una unità di analisi che nel nostro caso è un brano musicale avente testo in italiano, le colonne rappresentano le variabili che sono state rilevate su tutte le unità di analisi. L'incrocio tra una riga x e una colonna y produce una cella della matrice che rappresenta il valore (o stato) del brano x sulla proprietà (trasformata in variabile) y.

Il file dati (matrice dei dati) si compone di 2.489 righe, vale a dire 2.489 brani diversi, e 77 colonne, ossia 77 variabili rilevate per ogni unità di analisi.

4.10 l'elaborazione dei dati

Dopo aver costruito la matrice dei dati, si sono effettuate le opportune elaborazioni che hanno permesso, in accordo con gli obiettivi dell'indagine, di descrivere-interpretare i risultati della ricerca. Il modello dell'analisi dei dati è parte integrante del disegno della ricerca, almeno nelle sue linee essenziali, e deve essere progettato insieme alle altre fasi in cui esso si articola. Una buona analisi dei dati non dipende dall'applicazione di una tecnica avanzata e sofisticata di elaborazione in astratto, ma dall'aver scelto, coerentemente con il disegno della ricerca adottato, la tecnica statistica più idonea rispetto agli scopi dell'indagine.

La fase dell'elaborazione si è articolata su tre diversi livelli:

1. analisi univariata;
2. analisi bivariata;
3. analisi multivariata.

Ogni livello è stato condotto in maniera opportuna rispetto alla natura delle variabili. Nella nostra ricerca si sono usate prevalentemente variabili categoriali (non ordinabili e ordinabili) e, in qualche caso, variabili cardinali (misurative e/o di conteggio: anno di pubblicazione del disco; posizione del brano nel disco; e la batteria di differenziale semantico; cfr. D03, D07 e D21 nell'allegato a). Si sono, inoltre, usate variabili di tipo lessicale (cfr. D02, D06 e D23 nell'allegato a) su cui sono state condotte analisi di tipo lessicale. Di seguito esporremo, per ogni tipo di variabili, le procedure di analisi usate e gli scopi per cui sono state usate, dedicando maggiore spazio a quelle tecniche che, si presume, siano meno note tra i lettori.

4.11 l'elaborazione delle variabili categoriali

A livello di analisi univariata, per questo tipo di variabili, si individua la frequenza per ogni modalità in valore assoluto e in percentuale. Tali percentuali sono riportate sulla copia del questionario riprodotta nell'allegato A alla fine del presente capitolo. Dall'analisi delle distribuzioni di frequenza, sono emerse le modalità tipiche (moda per le classificazioni semplici e mediana per quelle ordinabili) e, nel caso di modalità con frequenze trascurabili, si è considerata l'opportunità di effettuare delle riclassificazioni per evidenziare in maniera più netta le differenze più rilevanti all'interno delle singole variabili.

L'analisi bivariata più comune per le variabili categoriali consiste nella costruzione di tabelle a doppia entrata, anche dette di contingenza, che permettono di analizzare le relazioni tra due variabili simultaneamente considerate. Se si ipotizza una relazione asimmetrica, nel senso che una variabile detta indipendente produce un effetto sull'altra, detta dipendente, con l'uso delle tabelle di contingenza si può controllare, ma non confermare, la plausibilità di questa ipotesi di dipendenza. Se non si assume una ipotesi di dipendenza tra le due variabili incrociate, e quindi la natura della relazione è simmetrica, attraverso l'uso di un coefficiente di relazione statistica si può stabilire se la

relazione bivariata considerata è imputabile agli effetti del caso. Nel presente rapporto sono state pubblicate solo le tabelle più utili alla interpretazione dei dati rilevanti in connessione con le ipotesi formulate.

Le tecniche di analisi multivariata considerano simultaneamente più variabili e possono, fondamentalmente, essere finalizzate a fornire una lettura sintetica ed economica dei dati o per controllare l'adeguatezza di modelli di connessione causale tra le variabili elaborati dai ricercatori. Per le variabili categoriali le tecniche di analisi multivariata idonee per gli scopi sopra esposti sono l'analisi delle corrispondenze multiple (per la lettura sintetica) e i modelli log-lineari (per i modelli di relazione causale). Nel nostro caso, avendo come obiettivo prioritario la descrizione dei dati, si è fatto ricorso alla tecnica delle corrispondenze multiple che studia le relazioni tra le variabili categoriali (semplici e ordinali) e può essere considerata del tutto analoga all'analisi in componenti principali che analizza variabili cardinali. Essendo questa analisi di tipo fattoriale, essa ha l'obiettivo di sostituire alla matrice dei dati originale una nuova matrice di dimensioni ridotte che costituisca la migliore rappresentazione sintetica della stessa. La rappresentazione sintetica della matrice dei dati viene ottenuta attraverso l'individuazione di dimensioni soggiacenti alla matrice stessa che riassumono e comprendono le principali relazioni di interdipendenza tra le variabili analizzate in termini di un limitato, e di solito piccolo, numero di nuovi costrutti anche detti fattori.

Il punto di partenza dell'analisi delle corrispondenze multiple è costituito da una tabella di contingenza relativa a n soggetti su cui sono state rilevate p variabili categoriali. L'analisi delle corrispondenze multiple utilizza la metrica del Chi^2 per analizzare le relazioni, in termini di distanza, tra le righe (soggetti), tra le colonne (variabili) e tra esse congiuntamente, calcolata, di volta in volta, come somma delle differenze al quadrato tra le frequenze ponderate. Prima di poter effettuare l'analisi si deve trasformare la matrice dei dati originale (casi per variabili) in due nuove matrici dette: matrice logica-disgiuntiva completa e matrice di Burt. La prima è di dimensioni $n \times q$ dove n è il numero dei soggetti e q la somma di tutte le modalità delle p variabili. Per ogni unità d'analisi si assegna una colonna a ciascuna modalità di ogni variabile e si attribuisce il valore **0** se il soggetto non presenta quella modalità o il valore **1** se il soggetto presenta quella modalità.

Questo metodo di codifica si dice:

- disgiuntivo perchè le diverse modalità di ciascuna variabile si escludono a vicenda;
- completa perchè ad ogni soggetto viene attribuito sicuramente un valore (**0** o **1**) per ciascuna modalità di ciascuna variabile.

Per ogni variabile si crea così un insieme di nuove variabili, dette variabili indicatrici, che possono assumere valori 0/1 (assenza/presenza). La matrice disgiuntiva completa può essere considerata come una particolare tabella di contingenza doppia.

La matrice di Burt, anche detta matrice delle corrispondenze multiple, si ottiene incrociando tra loro tutte le modalità di una matrice disgiuntiva completa; questa è una tabella di frequenze di dimensione $q \times q$ ed è simmetrica rispetto alla diagonale principale. Attraverso queste elaborazioni della matrice dei dati originaria, con delle trasformazioni che tengono conto dei rapporti di ciascuna vettore riga o vettore colonna con i rispettivi totali marginali, vengono presi in considerazione i profili riga e i profili colonna e viene analizzata la dispersione dei profili stessi intorno al loro "centro di gravità" ("profilo marginale" di riga o di colonna) queste operazioni si effettuano usando la metrica del Chi^2 tramite la formula:

$$d^2 = \text{Sommatore}_{j \text{ da } 1 \text{ a } q} T/C_j \times (f_{ij}/F_i - f_{rj}/F_r)^2$$

dove ciascuna distanza tra due elementi j viene ponderata con il termine T/C_j (inverso del rapporto C_j/T , che è la frequenza relativa marginale della modalità j). Con questa metrica si ponderano le modalità in quanto, da una parte, si rivaluta il contributo dato dalle modalità con frequenza bassa e, dall'altra, si ridimensiona il contributo dato dalle modalità con frequenza alta. Per esemplificare i concetti ora illustrati, possiamo immaginare la matrice dei profili riga (costruita a partire da una matrice disgiuntiva completa), come un iperspazio a q dimensioni generato dalle q modalità e, entro questo iperspazio, si possono collocare le n unità di analisi sotto forma di una nuvola di punti. Per cui ogni unità d'analisi è rappresentata da un punto che ha come coordinate i q elementi del suo profilo- riga. Analogamente, invertendo il rapporto righe colonne, possiamo proiettare la matrice dei profili-colonna, in un iperspazio a n dimensioni generato dalle n unità d'analisi, in questo modo avremo una nuvola di q punti che rappresentano le q modalità. E', comunque, dimostrato che l'analisi condotta sui profili-riga è equivalente a quella condotta sui profili-colonna.

Obiettivo dell'analisi delle corrispondenze multiple, come di qualsiasi altra tecnica fattoriale, è trovare un *set* ridotto di nuove variabili in grado di "riprodurre" il massimo dell'inerzia desunta dal *set* di variabili rilevate (originarie) con un ridotto numero di fattori che siano espressione delle combinazioni delle q modalità originarie.

Una volta individuati i fattori, si possono rappresentare le modalità e/o i casi come punti proiettati all'interno di spazi fattoriali tra di loro ortogonali. Per questa rappresentazione, si usano le coordinate fattoriali, prodotte in funzione delle associazioni tra le variabili e i fattori prodotti. Questi fattori prodotti soddisfano le seguenti condizioni:

- sono ortogonali tra loro (indipendenti l'uno dall'altro);
- sono combinazioni delle modalità dalle quali parte l'analisi;
- riproducono ciascuno, in ordine decrescente, il massimo dell'inerzia (variabilità) della matrice originale dei dati.

L'ammontare dell'inerzia riprodotta da ciascun fattore è espressa dall'autovalore (o valore proprio) che permette di quantificare la percentuale di inerzia riprodotta di ogni fattore. Come per l'analisi in componenti principali, l'analisi delle corrispondenze multiple produce tanti fattori, quante sono le modalità delle variabili usate (meno uno). La somma di tutti gli autovalori dei fattori è pari al totale dell'inerzia: ne consegue che il rapporto tra l'autovalore del primo fattore e la somma di tutti gli autovalori è la proporzione di inerzia riprodotta dal primo fattore (o la %, se moltiplicata per 100).

Ciascun fattore riproduce, quindi, una quota dell'inerzia totale, ma occorre tener presente che questa va rivalutata in quanto la proporzione di inerzia viene calcolata sul numero delle modalità e non sul numero delle variabili usate. Come in analisi delle componenti principali, la loro importanza è decrescente e solo i primi k fattori vengono presi in considerazione per l'interpretazione dei risultati (d'altra parte se si considerassero tutti i q fattori estratti, verrebbe meno lo scopo principale dell'analisi, cioè quello della riduzione della dimensione della matrice originaria dei dati).

Definire il numero di fattori da usare per interpretare la massima quota di inerzia è una scelta che si effettua caso per caso sulla base di due criteri-obiettivi: la sintesi e l'economicità delle dimensioni (fattori) che devono essere interpretati. Solitamente, nell'analisi delle corrispondenze multiple si individua il fattore F oltre il quale il contributo di inerzia riprodotta diventa poco rilevante, ovvero, cumulando le inerzie dei fattori successivi non si registra un aumento significativo. Secondo Benzecri, a cui si deve la paternità della tecnica delle corrispondenze multiple, gli autovalori dei primi fattori subiscono una sottovalutazione dell'inerzia in quanto

subiscono un effetto di ridimensionamento a causa dell'elevato numero di modalità su cui di solito si opera. Per rivalutare l'inerzia riprodotta dai primi fattori estratti, Benzecri (1979) propone una formula, detta "ottimista", che tiene conto solo degli autovalori di entità non "trascurabile": considerando solo gli autovalori fino a quello che abbia valore superiore al rapporto $1/p$ (dove p =al numero delle variabili). In questo modo si ponderano gli autovalori "rilevanti" elevando al quadrato le differenze tra gli autovalori originali e la costante derivante dal rapporto $1/p$ e ricalcolando la percentuale di inerzia riprodotta sommando gli autovalori "rilevanti" e rapportando ciascun autovalore a questo nuovo totale.

Nell'analisi delle corrispondenze multiple, di possono definire le variabili in due categorie distinte; le variabili attive, ovvero quelle che concorrono direttamente alla formazione degli assi fattoriali, e le variabili supplementari o illustrative, che non contribuiscono alla formazione dei fattori, e sono utilizzate, attraverso la loro collocazione sugli assi fattoriali, per l'interpretazione dei fattori e/o per studiare eventuali legami di interdipendenza con questi.

Una volta che si è determinato il numero dei fattori da estrarre, rispetto alla loro importanza relativa alla quota di inerzia riprodotta, si passa ad analizzare i singoli fattori, per ottenere una interpretazione semantica degli stessi individuando le modalità che contribuiscono in maniera rilevante alla loro formazione.

Per interpretare semanticamente i fattori si utilizzano sia le coordinate fattoriali (che indicano la posizione dei q punti-modalità originali sui k assi nello spazio a k dimensioni, ed esprimono le relazioni tra modalità e i fattori), sia alcuni valori statistici che permettono di valutare l'importanza e il peso che ogni variabile-modalità attiva ha nella formazione dei fattori.

I valori statistici più rilevanti per l'interpretazione dei fattori sono:

- la **massa** (o **peso relativo**) di ciascuna variabile-modalità. E' il rapporto tra la frequenza della variabile-modalità e il numero totale di variabili-modalità attive;
- l'**indice di distorsione** (o **distanza dall'origine**) di ciascuna variabile-modalità, che dà indicazioni sul carattere periferico di certe variabili-modalità: di solito è inversamente proporzionale alla massa - quindi a valori alti di questo indice corrisponde una massa debole e viceversa - ed esprime, insieme alla massa l'importanza-irrelevanza della variabile- modalità;
- il **contributo assoluto** di ciascuna variabile-modalità, che rappresenta la parte di inerzia totale del fattore dovuta alla variabile-modalità cui si riferisce. Il contributo assoluto di una variabile-modalità è proporzionale sia alla sua massa che al quadrato della sua coordinata fattoriale. Se, per esempio, per la variabile-modalità j e per un fattore F , definiamo f = coordinata fattoriale, m = massa e λ = autovalore associato al fattore, il contributo assoluto (**ca**) della variabile-modalità j al fattore F sarà:

$$ca = mf^2/\lambda.$$

Per ogni asse fattoriale la somma di tutti i contributi assoluti è uguale a 100. Con il contributo assoluto si valuta dunque quanta parte una variabile-modalità ha avuto nel determinare un certo fattore, in rapporto all'insieme delle variabili-modalità;

- il **coseno quadrato** detto anche **contributo relativo** o **qualità della rappresentazione**, è un valore che permette di valutare il contributo che un certo fattore F fornisce alla riproduzione dell'inerzia della variabile-modalità. Se questo contributo è basso vuol dire che la variabile-modalità in questione non è ben rappresentata sull'asse fattoriale F , e la sua inerzia è riprodotta da altre dimensioni. Al contrario, se il contributo è elevato, è lecito analizzare il ruolo che essa gioca nella formazione dell'asse sul quale è ben rappresentata. Il coseno quadrato è espresso in proporzione di inerzia (varianza) della variabile-modalità riprodotta dall'asse fattoriale. I q fattori

riproducono congiuntamente il 100% dell'inerzia di ogni variabile-modalità, quindi in media ogni fattore ne può, teoricamente, riprodurre il $(100/q)\%$;

- le **coordinate fattoriali**, possono avere segno positivo o negativo, stabiliscono la posizione delle variabili-modalità sugli assi, sia in termini di distanza dal centro-origine, sia in termini di versante (positivo-negativo) dell'asse-fattore considerato. Le variabili-modalità che presentano valori alti nelle coordinate fattoriali (cioè quelle che sono più distanti dall'origine degli assi) sono quelle che, di solito, contribuiscono di più alla formazione dell'asse stesso. Tuttavia, il valore della coordinata e l'importanza della variabile-modalità non sono direttamente proporzionali in modo rigidamente determinato.

Una volta che si sono interpretati singolarmente i fattori, l'analisi delle corrispondenze multiple permette l'analisi grafica dei **piani fattoriali** risultanti dall'intersezione degli assi fattoriali considerati due a due. In questi piani fattoriali si possono proiettare le variabili-modalità o i casi, sotto forma di una nuvola di punti, utilizzando le coordinate fattoriali delle stesse. Valutando le posizioni dei punti rispetto agli assi e le distanze tra i punti stessi si possono desumere le strutture delle relazioni tra le variabili-modalità e tra queste e i fattori. In generale si può ritenere che più un punto è lontano dall'origine di un asse, maggiore è il suo contributo alla formazione dell'asse stesso, mentre più elevata è la prossimità tra due variabili-modalità, maggiore è l'interdipendenza tra queste.

Se si concentra l'analisi sull'individuazione e definizione dei singoli fattori (quindi valutando soprattutto i contributi assoluti più rilevanti) si adotta il cosiddetto approccio interpretativo **fattorialista**, mentre se si analizza la forma della nuvola di punti e le distanze tra questi (valutando soprattutto i coseni quadrati elevati) si adotta un approccio di tipo **geometrico-strutturale**. Questi due approcci interpretativi più che essere alternativi sono complementari in quanto ciascuno evidenzia delle specificità che sono assenti nell'altro.

4.12 l'elaborazione delle variabili cardinali

Per variabili cardinali si intendono quelle variabili i cui stati sono espressi in multipli e sottomultipli di una unità di misura o di conto. In altre parole, se la proprietà che si intende rilevare è continua (sono cioè possibili infinite cifre per ogni stato), abbiamo bisogno di una unità di misura che permetta, con delle necessarie approssimazioni, di rilevare gli stati dei casi su questo tipo di proprietà. Ci sono anche proprietà di tipo discreto (i cui stati si rilevano per incrementi o decrementi finiti) e, in questo caso, non abbiamo bisogno di una unità di misura, in quanto la rilevazione degli stati su proprietà di effettua tramite conteggio (cfr. Marradi, 1985). Nella nostra ricerca le variabili cardinali usate, in cui rientrano le variabili ottenute per conteggio, sono rispettivamente la D03 (anni di pubblicazione del disco) e la D07 (posizione dei brani analizzati all'interno del disco). Abbiamo poi inserito una batteria di differenziale semantico i cui *items*, anche se con una certa approssimazione, possono essere trattati come delle variabili variabili quasi-cardinali.

A livello di analisi univariata per questo tipo di variabili, si calcolano dei valori caratteristici di concentrazione (la media) e di dispersione (lo scarto quadratico medio) al fine di caratterizzare in maniera sintetica le loro distribuzioni.

L'analisi bivariata più comune per le variabili cardinali consiste nel calcolo del coefficiente di correlazione lineare, che permette di individuare una associazione di tipo lineare tra due variabili.

La tecnica di analisi multivariata usata per la batteria di differenziale semantico è stata l'analisi delle componenti principali che rientra nelle tecniche dette fattoriali, anche se, rispetto alle altre tecniche fattoriali, essa è una tecnica fondamentalmente descrittiva (come l'analisi delle corrispondenze multiple). Ci sono, infatti, due possibilità diverse nel processo di estrazione di fattori da una matrice: i fattori definiti e i fattori inferiti. Sono fattori definiti i fattori - sarebbe più opportuno chiamarli componenti - prodotti dall'analisi delle componenti principali e dall'analisi delle corrispondenze, in quanto non si fa alcuna assunzione inferenziale riguardo la struttura dei dati e le cause della loro variabilità. Con i fattori inferiti invece, si interpreta la variabilità dei dati osservati con delle variabili indipendenti latenti (i fattori) comuni a tutte le variabili. In questo secondo caso, ci sono diverse tecniche fattoriali che con tecniche diverse (massima verosimiglianza, minimi quadrati, ecc.) producono i fattori inferiti.

La tecnica delle componenti è, come già detto, del tutto analoga a quella delle corrispondenze multiple, anche se, per ragioni storico-cronologiche, sarebbe più corretto dire il contrario, ovvero che l'analisi delle corrispondenze multiple è la traduzione dell'analisi in componenti principali applicata a variabili categoriali. Obiettivo dell'analisi in componenti principali è esprimere una matrice di correlazioni tra variabili rilevate in termini sintetici, ovvero tramite un limitato numero di componenti che costituiscono delle combinazioni lineari delle variabili analizzate con una minima perdita di informazione. La prima componente estratta è quella a cui corrisponde la massima percentuale di varianza riprodotta, la seconda, ortogonale rispetto alla prima, esprimerà una quota di varianza riprodotta minore rispetto alla prima e così via fino ad estrarre tante componenti quante sono le variabili immesse nell'analisi. Naturalmente sarà necessario usare solo le prime componenti estratte per avere una efficace sintesi dei dati. Per l'interpretazione e la valutazione dell'analisi in componenti principali ci avvaliamo di una serie di valori statistici. L'autovalore è connesso ad ogni componente estratta permette di valutare il potere di sintesi di ogni fattore. Più è grande l'autovalore e maggiore sarà la percentuale di varianza riprodotta da quel fattore. La comunalità quantifica la percentuale di varianza riprodotta delle variabili analizzate dall'insieme delle componenti estratte. Ovviamente la comunalità delle variabili, sarà sempre uguale ad 1 se consideriamo tutte le possibili componenti (che sono uguali al numero delle variabili) e minore di 1 se lavoriamo solo con le prime componenti estratte (in questo si parla di analisi delle componenti principali troncata). Una volta stabilito il numero delle componenti principali da usare nell'analisi, che solitamente sono quelle che hanno un autovalore maggiore o uguale ad 1, si passa a considerare il rapporto che c'è tra queste e le variabili analizzate. A questo scopo, si utilizzano dei coefficienti chiamati *component loadings* (pesi componenziali) che indicano il peso che viene attribuito per ogni variabile a ciascuna componente principale; più questo è alto maggiore è la relazione lineare tra la variabile e la componente principale. Possiamo interpretare i *component loadings* come dei coefficienti di correlazione tra le variabili analizzate e le componenti estratte. Con la lettura dei *component loadings* e con una opportuna rotazione delle componenti estratte, riusciamo ad interpretare semanticamente le componenti principali individuando dei sottogruppi di variabili che sono connessi a singole componenti principali. Una volta interpretate le componenti principali, possiamo, infine, calcolare dei punteggi componenziali (*component scores*) per ogni caso rispetto ad ogni componente usata. In questo modo possiamo arricchire la nostra matrice dei dati con delle nuove colonne che rappresenteranno una sintesi di una parte o di tutta, o nel caso che avessimo rilevato tutte variabili cardinali, la matrice dei dati.

4.13 l'elaborazione delle variabili lessicali

Gli ultimi sviluppi della *scuola francese dell'analisi dei dati* hanno esteso le possibilità della tecnica dell'analisi delle corrispondenze a un nuovo tipo di variabili che si definiscono lessicali. Una variabile si dice lessicale quando le proprie modalità sono costituite da parole, frasi o, anche, testi abbastanza estesi (ad es. lettere, articoli e altri tipi di testi). Possiamo considerare una variabile lessicale una domanda aperta di un questionario, nella quale ogni risposta libera di ogni individuo intervistato è una modalità distinta. Normalmente, questo tipo di dati viene trattato con delle classificazioni, che aggregano le risposte semanticamente simili in gruppi omogenei, a cui vengono assegnati dei codici numerici che sono poi immessi nella matrice dei dati per le successive elaborazioni. Queste operazioni sono di solito molto faticose, sia in termini di energie che di tempo, e comunque comportano sempre una sensibile perdita di informazioni. Ricorrendo alle tecniche dell'analisi dei dati testuali, possiamo risolvere in maniera più efficace problemi di questo tipo considerando ed elaborando direttamente i testi senza doverli ricodificare e trasformare in valori numerici avendo, inoltre, la possibilità di mettere in relazione le variabili lessicali con gli altri tipi di variabili (categoriali e cardinali) raccolte.

Nella nostra ricerca le variabili lessicali usate sono il titolo del disco (D02), il titolo del brano (D06) e la D23 nella quale gli analisti dovevano selezionare le frasi caratteristiche dei pezzi analizzati. Quest'ultima variabile esige una operazione di estrapolazione dall'intero testo del brano analizzato di un suo segmento al tempo stesso significativo e rappresentativo dal punto di vista semantico-linguistico dell'intero testo. La necessità di operare subito in maniera sintetica è dovuta principalmente a ragioni di praticità e di economicità, in termini di tempo e di costi d'elaborazioni, in quanto un'analisi di tipo lessicale condotta in uno studio di analisi del contenuto non è priva di problemi. Se intendiamo usare l'analisi linguistica in una prospettiva di analisi del contenuto, e non, quindi, meramente lessicometrica, dobbiamo pre-trattare i lati lessicali in maniera opportuna. Un primo problema consiste nella differenziazione di quelle parole che pur essendo identiche hanno significati diversi (per es. amo=verbo e amo=oggetto usato per la pesca) e, all'opposto riunire parole che pur essendo diverse hanno lo stesso significato (es. amo, ami, ama, ecc.). L'importanza di una parola in un testo non si fonda tanto sulle sue ricorrenze quanto sull'univocità del suo significato, per cui è necessario operare con criteri di lemmatizzazione che permettono di ricondurre le diverse parole presenti nel testo alla loro comune radice (es. ha, hanno, abbia, avete, ecc. = verbo avere) o e parole con significato univoco allo stesso termine. L'operazione di riunione di due forme lessicali aventi lo stesso significante, ma un diverso significato (omografi), viene risolta attraverso un processo di disambiguazione. Osservando il contesto linguistico in cui è inserito il termine, potremmo, per esempio, distinguere tra il significante fine con significato equivalente a scopo e lo stesso fine con significato di termine o di conclusione. Il secondo caso prevede la fusione di alcuni termini sotto lo stesso "lemma" quando il ricercatore ritiene irrilevante ai fini della ricerca la differenza simbolica (Bolasco, 1992). Come visto, un primo tipo di fusione è dato dall'equivalenza grammaticale che consente di riunire le flessioni di uno stesso lemma (es. tutto, tutti, tutte, tutta), un altro consiste nella fusione di più sinonimi in un termine chiave. Per l'illustrazione delle operazioni di lemmatizzazione e di disambiguazione effettivamente svolte nella nostra ricerca si rimandano i lettori al Cap. 3.

Una volta che si sono predisposti i dati lessicali, si può procedere alle effettive operazioni di analisi usando un *software* adeguato per questo tipo di variabili (SPAD.T). A livello di analisi univariata, abbiamo una classificazione delle parole in ordine alfabetico e in ordine di frequenza.

Possiamo quindi calcolare degli indici di ricchezza linguistica (rapporto tra parole diverse sul totale delle parole) e le percentuali con le quali le parole e/o i lemmi sono presenti nell'intero *corpus* testuale. Questo primo livello d'analisi consente applicazioni che sono molto diffuse nella lessicologia e nella lessicografia.

Come già detto l'aspetto più innovativo del programma SPAD.T, è la possibilità di trattare le variabili lessicali in relazione con gli altri tipi di variabili. E' quindi possibile procedere ad un livello di analisi bivariata costruendo una tabella di contingenza forme linguistiche per individui: disponendo sulle righe le forme lessicali e sulle colonne gli individui aggregati secondo una variabile classificatoria. In questo modo è possibile individuare, con procedure statistiche basate sulla metrica del CHI^2 e sulle frequenze, quali forme linguistiche siano caratteristiche all'interno delle diverse categorie nelle quali abbiamo ripartito le nostre unità d'analisi (cfr. le procedure MOCAR e RECAR del programma SPAD.T, 1989).

Infine, a livello di analisi multivariata, possiamo applicare la tecnica delle corrispondenze lessicali. Essa è del tutto analoga alla tecnica delle corrispondenze multiple; l'obiettivo è sempre quello di sintetizzare l'informazione contenuta nei dati attraverso l'individuazione di nuove dimensioni sintetiche rispetto alla matrice dei dati. La novità dell'analisi delle corrispondenze lessicali consiste nell'immissione nella matrice dati anche di variabili lessicali. Le forme lessicali sono prese in considerazione come casi e costituiscono le righe della tabella di contingenza, mentre le colonne possono essere indifferentemente o gli individui, nel caso si analizzino diversi *corpus* testuali, o dalle modalità di alcune variabili associate alle forme lessicali, questo è in particolare il caso delle domande aperte di un questionario. L'elaborazione dei dati lessicali è essenzialmente di tipo descrittivo-esplorativo come sottolineano i due autori francesi (Lebart e Salem, 1988) ed è finalizzato ad ottenere una visione di sfondo delle relazioni esistenti tra i dati. Anche con l'analisi delle corrispondenze lessicali si producono dei grafici fattoriali che permettono di analizzare le associazioni significative tra le parole che emergono dai testi e le altre caratteristiche che sono state rilevate.

Allegato A:
***Scheda di analisi del contenuto dei testi delle canzoni italiane dal 1971 al 1992
e distribuzioni di frequenza dei relativi items***

Numero Scheda |__|__|__|__|

Analista _____

Sezione A: Morfologia Album

D01 Autore Album _____

Classificazione Autore A:

1. Melodico	1203	48,3%
2. Autore	1048	42,1%
3. Rock	238	9,6%
Tot.	2489	100,0%

Classificazione Autore B:

1. Grandi Interpreti	609	24,5%
2. Meteore	650	26,1%
3. Autori Anni '70	653	26,2%
4. Autori Anni '80	340	13,7%
5. Rock Anni '70	94	3,8%
6. Rock Anni '80	143	5,7%
Tot.	2489	100,0%

D02 Titolo Album _____

D03 Anno pubblicazione brano:

1970	10	0,4%
1971	93	3,7%
1972	73	2,9%
1973	93	3,7%
1974	133	5,3%
1975	100	4,0%
1976	30	1,2%
1977	85	3,4%
1978	73	2,9%
1979	103	4,1%
1980	70	2,8%
1981	76	3,1%
1982	79	3,2%
1983	83	3,3%
1984	75	3,0%
1985	140	5,6%
1986	139	5,6%
1987	174	7,0%
1988	168	6,7%
1989	152	6,1%
1990	122	4,9%
1991	182	7,3%
1992	236	9,5%
Tot.	2489	100,0%

D04 Tipo Album:

1. ___ Singolo	2278	91,5%
2. ___ Doppio	211	8,5%
Tot.	2489	100,0%

D05 L'Album è composto da:

1. ___ Brani inediti	1973	79,3%
2. ___ Raccolto brani editi	362	14,5%
3. ___ Brani dal vivo	154	6,2%
Tot.	2489	100,0%

Sezione B: Morfologia Pezzo

D06 Titolo Pezzo Analizzato _____

D07 Posizione Pezzo all'interno Dell'album: |___|___|

D08 Tipo Pezzo:

1. <input type="checkbox"/> Brano con testo in italiano	2206	88,6%
2. <input type="checkbox"/> Brano strumentale*	72	2,9%
3. <input type="checkbox"/> Brano totalmente in dialetto o in lingua straniera*	211	8,5%
Tot.	2489	100,0%

* = Se D08=2 O 3 la scheda finisce.

D09 Se d08=1 Modalità di esecuzione delle parti vocali del pezzo:

1. <input type="checkbox"/> Una sola voce	1972	89,3%
2. <input type="checkbox"/> Due voci o più voci	236	10,7%
Tot.	2206	100,0%

D10 Nel testo è individuabile chiaramente una sua sezione che ha funzioni di ritornello:

1. <input type="checkbox"/> Sì	1846	83,6%
2. <input type="checkbox"/> No	361	16,4%
Tot.	2206	100,0%

D11 Quantificazione delle parti cantate sul totale del pezzo:

1. <input type="checkbox"/> Prevalgono le parti cantate	1916	86,8%
2. <input type="checkbox"/> Si equivalgono	220	10,0%
3. <input type="checkbox"/> Prevalgono le parti strumentali	71	3,2%
Tot.	2206	100,0%

Sezione c: argomento del testo

D12 Qual è l'argomento principale del testo:

1. <input type="checkbox"/> Amore	1093	49,5%
2. <input type="checkbox"/> Rapporti affettivi	125	5,7%
3. <input type="checkbox"/> Sfera personale	432	19,6%
4. <input type="checkbox"/> Sfera sociale	267	12,1%
5. <input type="checkbox"/> Altro	214	9,7%
6. <input type="checkbox"/> Non presente	75	3,4%
Tot.	2206	100,0%

D12_1 Classificazione tema amore:

1. Amore in generale	123	11,3%
2. Fine di un rapporto	186	17,0%
3. Malesseri in amore	226	20,7%
4. Rapporti con il partner	336	30,8%
5. Esaltazione partner	157	14,4%
6. Erotismo	64	5,9%
Tot.	1093	100,0%

D12_2 Classificazione tema altri rapporti affettivi:

1. Rapporti affettivi in generale	23	17,7%
2. Amicizia	57	46,0%
3. Rapporti intergenerazionali	45	36,3%
Tot.	125	100,0%

D12_3 Classificazione tema sfera personale:

1. Autoreferenzialità	135	31,2%
2. Disagi esistenziali	130	30,0%
3. Riflessioni diacroniche	100	23,1%
4. Rapporto con gli altri	44	10,4%
5. Rapporto con la religione	23	5,3%
Tot.	432	100,0%

D12_4 Classificazione tema sfera sociale:

1. Sfera sociale in generale	73	27,9%
2. Emergenze sociali	36	13,4%
3. Umanità rispetto vita	34	12,6%
4. Religione	21	7,8%
5. Progresso sociale	24	8,9%
6. Pacifismo	20	7,4%
7. Emarginazione sociale	31	11,5%
8. Politica	28	10,4%
Tot.	267	100,0%

D12_5 Classificazione altri temi:

1. Luoghi geografici	32	15,0%
2. Canzoni sulla musica	67	31,5%
3. Biografie	81	38,0%
4. Canzoni sul viaggio	26	12,2%
5. Altro	8	3,3%
Tot.	214	100,0%

D13 Il testo viene sviluppato prevalentemente:

1. ___ In forma dialogica	75	3,4%
2. ___ In forma impersonale	146	6,6%
3. ___ In forma personale	1985	90,0%
Tot.	2206	100,0%

D14 I verbi presenti nel testo vengono coniugati prevalentemente:

1. ___ In prima persona singolare	818	37,1%
2. ___ In prima persona plurale	126	5,7%
3. ___ In seconda persona singolare	653	29,6%
4. ___ In seconda persona plurale	57	2,6%
5. ___ In terza persona singolare	358	16,2%
6. ___ In terza persona plurale	59	2,7%
7. ___ In forma impersonale	135	6,1%
Tot.	2206	100,0%

D15 L'argomento del testo viene trattato da un punto di vista:

1. ___ Individuale	2066	93,7%
2. ___ Collettivo	140	6,3%
Tot.	2206	100,0%

D16 Nel testo nel complesso si individua rispetto all'argomento:

1. ___ Una descrizione dell'argomento	1360	61,7%
2. ___ Una interpretazione - analisi dell'argomento	576	26,1%
3. ___ Una soluzione rispetto l'argomento	270	12,2%
Tot.	2206	100,0%

D17 Nel testo si può individuare - in modo esplicito - un destinatario:

1. <input type="checkbox"/> Partner	699	31,7%
2. <input type="checkbox"/> Presente non specificabile	192	8,7%
3. <input type="checkbox"/> Amico	58	2,6%
4. <input type="checkbox"/> Membro famiglia	29	1,3%
5. <input type="checkbox"/> Dio o Gesù	21	1,0%
6. <input type="checkbox"/> Non presente	1207	54,7%
Tot.	2206	100,0%

D18 Il testo viene presentato:

1. <input type="checkbox"/> Come una storia-racconto	232	10,5%
2. <input type="checkbox"/> Come una unica immagine	213	9,7%
3. <input type="checkbox"/> Come una serie di riflessioni o di ricordi	1173	53,1%
4. <input type="checkbox"/> Come un discorso diretto	588	26,7%
Tot.	2206	100,0%

D19 Il contesto denotativo del testo fa riferimento a:

1. <input type="checkbox"/> Svago tempo libero	249	11,3%
2. <input type="checkbox"/> Lavoro	74	3,4%
3. <input type="checkbox"/> Scuola	7	0,3%
4. <input type="checkbox"/> Politico	20	0,9%
5. <input type="checkbox"/> Fantastico	127	5,8%
6. <input type="checkbox"/> Non presente	1729	78,3%
Tot.	2206	100,0%

D20 Il contesto connotativo del testo fa riferimento a:

1. <input type="checkbox"/> Famiglia (rapporti extra-partnership)	56	2,5%
2. <input type="checkbox"/> Gruppo dei pari (relazioni amicali)	163	7,4%
3. <input type="checkbox"/> Relazioni di partnership	972	44,1%
4. <input type="checkbox"/> Non presente	1015	46,0%
Tot.	2206	100,0%

D21 Il brano nel suo complesso è connotabile in termini di:

Onestà	12,8%	□□□□	47,3%	□□□□	Disonestà	39,9%
Sicurezza	43,9%	□□□□	22,8%	□□□□	Insicurezza	33,3%
Divertimento	24,1%	□□□□	48,2%	□□□□	Noia	27,7%
Soddisfacente	43,8%	□□□□	15,6%	□□□□	Insoddisfacente	40,6%
Reale	22,7%	□□□□	15,9%	□□□□	Irreale	61,4%
Collaborativo	14,8%	□□□□	57,3%	□□□□	Competitivo	27,9%
Pacifico	25,1%	□□□□	43,7%	□□□□	Conflittuale	31,2%
Facile	44,8%	□□□□	26,1%	□□□□	Difficile	29,1%
Innovazione	31,9%	□□□□	36,9%	□□□□	Tradizione	31,2%
Prestigio	24,0%	□□□□	52,4%	□□□□	Senza Prestigio	23,6%
Piacevole	35,3%	□□□□	20,8%	□□□□	Spiacevole	43,9%
Democratico	12,9%	□□□□	68,4%	□□□□	Autoritario	18,7%
Dinamico	30,5%	□□□□	27,0%	□□□□	Statico	42,5%
Pulito	11,8%	□□□□	59,2%	□□□□	Sporco	29,0%
Utile	22,7%	□□□□	49,1%	□□□□	Inutile	28,2%
Desiderabile	29,3%	□□□□	19,7%	□□□□	Indesiderabile	51,0%
Leggero	33,8%	□□□□	22,9%	□□□□	Pesante	43,3%
Autonomia	38,8%	□□□□	31,8%	□□□□	Dipendenza	29,4%
Riuscito	31,6%	□□□□	43,5%	□□□□	Fallito	24,9%
Bello	18,3%	□□□□	42,6%	□□□□	Brutto	39,1%
Individuale	17,9%	□□□□	25,0%	□□□□	Collettiva	57,1%
Normale	29,6%	□□□□	47,1%	□□□□	Eccezionale	23,3%
Rabbia	29,2%	□□□□	47,7%	□□□□	Calma	23,1%
Tristezza	17,3%	□□□□	49,3%	□□□□	Felicità	33,4%
Norma	26,0%	□□□□	54,5%	□□□□	Trasgressione	19,5%
Benessere	38,0%	□□□□	33,0%	□□□□	Malessere	29,0%

D22 In rapporto al testo la musica ha una funzione:

1. □□□□	Descrizione - commento puntuale del testo	1994	90,4%
2. □□□□	Enfaticizzazione - drammatizzazione del testo	124	5,6%
3. □□□□	Attenuazione - ironizzazione del testo	88	4,0%
	Tot.	2206	100,0%

Sezione d: caratteristiche linguistiche del testo

d23 Quali sono le parole/frasi chiave del testo (ovvero le parole-frasi del testo più significative a giudizio dell'analista.):

1 _____

2 _____

3 _____

D24 Nel testo sono presenti espressioni o parole gergali riferibili a gruppi particolari:

1. <input type="checkbox"/> Sì	56	2,5%
2. <input type="checkbox"/> No	2150	97,5%
Tot.	2206	100,0%

D25 Nel testo sono presenti espressioni o parole o frasi dialettali:

1. <input type="checkbox"/> Sì	63	2,9%
2. <input type="checkbox"/> No	2143	97,1%
Tot.	2206	100,0%

D26 Nel testo sono presenti parole o frasi in lingua straniera:

1. <input type="checkbox"/> Sì	213	9,7%
2. <input type="checkbox"/> No	1993	90,3%
Tot.	2206	100,0%

D27 Nel testo sono rintracciabili espressioni coprolaliche:

1. <input type="checkbox"/> Sì	79	3,6%
2. <input type="checkbox"/> No	2127	96,4%
Tot.	2206	100,0%

Sezione e: altri riferimenti caratteristici del testo

D28 Nel testo sono contenuti riferimenti a:

	Sì	No
Arte	0,8%	99,2%
Filosofia	0,4%	99,6%
Letteratura – poesia	1,3%	98,7%
Scienza	0,3%	99,7%
Spettacolo	2,9%	97,1%
Altro	3,0%	97,0%

D29 Nel testo esiste un riferimento a località geografiche:

1. <input type="checkbox"/> Sì	273	12,4%
2. <input type="checkbox"/> No	1933	87,6%
Tot.	2206	100,0%

D30 Nel testo sono rintracciabili chiari riferimenti a particolari dimensioni temporali:

1. ___	Passato	128	5,8%
2. ___	Presente	122	5,5%
3. ___	Futuro	60	2,7%
4. ___	Non presente	1896	86,0%
	Tot.	2206	100,0%

D31 Nel testo sono rintracciabili chiari riferimenti a oggetti, simboli, attività modali tipiche di determinati periodi storici:

1. ___	No	2088	94,7%
2. ___	Sì	118	5,3%
	Tot.	2206	100,0%

Sezione f: dimensioni valoriali nel testo

D32 Nel testo si sollecitano o si fa riferimento a:

	v.a.	%	Valore
1.	13	0,4%	Abilità
2.	113	3,2%	Affettività
3.	18	0,5%	Aggressività
4.	926	26,1%	Amore
5.	82	2,3%	Amicizia
6.	53	1,5%	Auto Controllo
7.	32	0,9%	Bellezza
8.	258	7,3%	Benessere Psico Fisico
9.	9	0,3%	Bontà
10.	8	0,2%	Competenza
11.	9	0,3%	Competizione
12.	80	2,3%	Comprensione
13.	39	1,1%	Comunicazione
14.	78	2,2%	Coraggio
15.	72	2,0%	Creatività
16.	11	0,3%	Cultura
17.	76	2,1%	Determinazione
18.	26	0,7%	Disinvoltura
19.	78	2,2%	Divertimento
20.	36	1,0%	Ecologia
21.	28	0,8%	Egoismo-Individualismo
22.	72	2,0%	Erotismo
23.	91	2,6%	Evasione
24.	28	0,8%	Fede Religiosa
25.	31	0,9%	Fiducia
26.	33	0,9%	Gelosia Possessività
27.	36	1,0%	Giustizia
28.	26	0,7%	Impegno Politico
29.	64	1,8%	Impegno Sociale
30.	174	4,9%	Innovazione - Differenziazione
31.	12	0,3%	Integrazione - Conformismo
32.	41	1,2%	Intraprendenza
33.	145	4,1%	Libertà
34.	25	0,7%	Moralità
35.	33	0,9%	Onestà
36.	34	1,0%	Pacifismo
37.	5	0,1%	Partecipazione (Politica)
38.	11	0,3%	Pazienza
39.	8	0,2%	Prestigio
40.	10	0,3%	Prudenza
41.	125	3,5%	Razionalità
42.	35	1,0%	Rispetto della vita
43.	33	0,9%	Semplicità
44.	26	0,7%	Sicurezza
45.	79	2,2%	Sicurezza di sé
46.	42	1,2%	Sincerità
47.	95	2,7%	Sensualità
48.	79	2,2%	Solidarietà
49.	16	0,5%	Successo
50.	29	0,8%	Tenacia
51.	16	0,5%	Tolleranza
52.	49	1,4%	Trasgressione

Allegato b:
Elenco dei dischi estratti per la costruzione del campione

Dall'archivio dei 692 dischi che costituiscono la nostra popolazione di riferimento sono stati estratti i seguenti 258 che costituiscono il campione esaminato nella ricerca:

	Autore	Titolo	Anno
1.	Al Bano & Romina	"Vincerai"	1992
2.	Alice	"Falsi allarmi"	1983
3.	Alice	"Gioielli rubati"	1985
4.	Alice	"Il sole nella pioggia"	1989
5.	Alunni Del Sole	"E mi manchi tanto"	1973
6.	Alunni Del Sole	"Jenny"	1974
7.	Alunni Del Sole	"Liu"	1978
8.	Antonacci	"Liberatemi"	1992
9.	Arbore	"Quelli della notte"	1985
10.	Arbore	"Napoli punto e a capo"	1992
11.	Aznavour	"E fu subito Aznavour"	1971
12.	Baglioni	"Questo piccolo grande amore"	1972
13.	Baglioni	"Gira che ti rigira amore bello"	1973
14.	Baglioni	"E tu"	1974
15.	Baglioni	"Solo"	1977
16.	Baglioni	"E tu come stai"	1978
17.	Baglioni	"Alé oo"	1982
18.	Baglioni	"La vita è adesso"	1985
19.	Baglioni	"Oltre"	1992
20.	Balsamo	"Un pugno nella notte"	1992
21.	Barbarossa	"Non tutti gli uomini"	1988
22.	Battiato	"Orizzonti perduti"	1983
23.	Battiato	"Giubbe rosse"	1988
24.	Battiato	"Come un cammello in una grondaia"	1991
25.	Battisti	"Umanamente uomo il sogno"	1972
26.	Battisti	"Il nostro caro angelo"	1973
27.	Battisti	"Anima latina"	1974
28.	Battisti	"La batteria il contrabbasso ecc."	1976
29.	Battisti	"Don Giovanni"	1986
30.	Battisti	"La sposa occidentale"	1990
31.	Battisti	"C.S.A.R"	1992
32.	Bella Marcella	"Metamorfosi"	1974
33.	Bennato	"Io che non sono l'imperatore"	1975
34.	Bennato	"Uffa uffa"	1980
35.	Bennato	"E' goal"	1984
36.	Bennato	"Kaiwanna"	1984
37.	Bennato	"Ok Italia"	1987
38.	Bennato	"Abbi dubbi"	1989
39.	Bennato	"Edo rinnegato"	1990
40.	Bennato	"Capitan uncino"	1992
41.	Bennato & D'angiò	"Garofano d'ammore"	1977
42.	Berté	"E la luna busso"	1980
43.	Banco del Mutuo Soccorso	"Darwin"	1972
44.	Buongusto	"Fortunatamente ancora l'amore"	1981
45.	Buongusto	"Belle bugie"	1983

	Autore	Titolo	Anno
46.	Branduardi	"La pulce d'acqua"	1977
47.	Branduardi	"Cogli la prima mela"	1979
48.	Branduardi	"Angelo Branduardi"	1981
49.	Califano	"Ti perdo"	1979
50.	Camaleonti	"Amicizia e amore"	1973
51.	Canzian	"Io e Red"	1986
52.	Caputo	"No smoking"	1985
53.	Carboni	"Forever"	1985
54.	Carboni	"Luca Carboni"	1987
55.	Carrà	"Raffaella"	1978
56.	Celentano	"Yuppi du"	1975
57.	Celentano	"Soli"	1979
58.	Celentano	"Deus"	1981
59.	Cementano	"Uh uh"	1982
60.	Cementano	"I miei amici americani 2"	1986
61.	Cementano	"La pubblica ottusità"	1987
62.	Cementano	"Il tempo se ne va"	1987
63.	Cementano	"Stai lontana da me"	1988
64.	Cementano	"Il re degli ignoranti"	1991
65.	Cementano	"Super best"	1992
66.	Christian	"Un'altra vita un altro amore"	1982
67.	Christian	"Sere"	1985
68.	Christian	"Guardando il cielo"	1988
69.	Christian	"Se non è amore deve essere amore"	1989
70.	Cinguetti	"Cantando con gli amici"	1971
71.	Ciondoli	"Un altro modo di vivere"	1982
72.	Cocciante	"Anima"	1974
73.	Cocciante	"L'alba"	1975
74.	Cocciante	"Concerto per margherita"	1976
75.	Cocciante	"E io canto"	1979
76.	Cocciante	"Cocciante"	1982
77.	Cocciante	"La grande avventura"	1987
78.	Cocciante	"Viva"	1988
79.	Cocciante	"Riccardo Cocciante"	1991
80.	Concato	"Punto e virgola"	1991
81.	Conte	"Aguaplano"	1987
82.	Conte	"Parole d'amore scritte a macchina"	1990
83.	Cugini Di Campagna	"Le canzoni dei Cugini di Campagna"	1988
84.	D'angelo	"Inseparabili"	1989
85.	D'avena	"I tuoi amici in tv n.4"	1990
86.	D'avena	"L'Europa siamo noi"	1991
87.	Dalla	"Automobili"	1976
88.	Dalla	"Dalla"	1979
89.	Dalla	"Lucio Dalla"	1980
90.	Dalla	"1983"	1983
91.	Dalla	"Cambio"	1990
92.	Dalla	"Amen"	1992
93.	Dalla & Morandi	"Dalla & Morandi"	1988
94.	Daniele	"Nero a metà"	1980
95.	Daniele	"Scio"	1984
96.	Daniele	"Ferryboat"	1985
97.	Daniele	"Un uomo in blues"	1991

	Autore	Titolo	Anno
98.	De Andre'	"La buona novella"	1970
99.	De Andre'	"Le nuvole"	1990
100.	De Andre'	"Il viaggio"	1991
101.	De Gregori	"De Gregori"	1974
102.	De Gregori	"De Gregori"	1978
103.	De Gregori	"Viva l'Italia"	1979
104.	De Gregori	"La nostra storia"	1987
105.	De Gregori	"Mira mare 19.4.1989"	1989
106.	De Gregori	"Niente da capire"	1990
107.	De Piscopo	"Bello carico"	1988
108.	De Sio	"Tre"	1983
109.	Delirium	"Dolce acqua"	1971
110.	Delirium	"Lo scemo e il villaggio"	1972
111.	Di Michele	"Raccolta"	1990
112.	Drupi	"Drupi"	1986
113.	Drupi	"Amica mia"	1992
114.	Esposito	"Il grande esploratore"	1984
115.	Esposito	"As tu as"	1985
116.	Facchinetti	"Roby Facchinetti"	1984
117.	Ferradini	"Una catastrofe bionda"	1983
118.	Ferri	"Sempre"	1973
119.	Finardi	"Dal blu"	1983
120.	Finardi	"Dolce Italia"	1987
121.	Finardi	"Millennio"	1991
122.	Fiorello	"Veramente falso"	1991
123.	Fogli	"Collezione"	1982
124.	Formula 3	"Formula 3"	1971
125.	Formula 3	"La grande casa"	1973
126.	Fortis	"Fortis"	1979
127.	Fortis	"Tra demonio e santità"	1980
128.	Fossati	"La mia banda suona il rock"	1980
129.	Fossati	"La pianta del tè"	1988
130.	Giacobbe	"Io vorrei"	1990
131.	Goggi	"C'è poesia"	1986
132.	Goggi	"Punti di vista"	1989
133.	Guccini	"Stanze di vita quotidiana"	1974
134.	Guccini	"Via Paolo Fabbri 43"	1976
135.	Guccini	"Amerigo"	1978
136.	Guccini	"Signora Bovary"	1987
137.	Jovanotti	"La mia moto"	1989
138.	Jovanotti	"Una tribù che balla"	1991
139.	Juli & Julie	"Una storia d'amore"	1975
140.	Ladri di Biciclette	"Ladri di biciclette"	1989
141.	Litfiba	"Sogno ribelle"	1992
142.	Mango	"Odissea"	1986
143.	Mannoia	"Fiorella Mannoia"	1986
144.	Mannoia	"Di terra e di vento"	1989
145.	Martini	"Mia Martini"	1972
146.	Martini	"Ritratti"	1972
147.	Martini	"Nel mondo una cosa"	1973
148.	Masini	"Malinconioia"	1991
149.	Matia Bazar	"Melancholia"	1985

	Autore	Titolo	Anno
150.	Matia Bazar	"Melo"	1987
151.	Matia Bazar	"Tutto il mondo dei Matia Bazar"	1991
152.	Matia Bazar	"Il tempo del sole"	1991
153.	Matia Bazar	"Tournée"	1991
154.	Mietta	"Canzoni"	1990
155.	Mina	"Baby Gate"	1974
156.	Mina	"Del mio meglio"	1975
157.	Mina	"Del mio meglio 3"	1975
158.	Mina	"Mina con bigne"	1977
159.	Mina	"Mina live '78"	1978
160.	Mina	"Finalmente ho conosciuto il conte"	1985
161.	Mina	"Uiallalla"	1987
162.	Mina	"Rane supreme"	1987
163.	Mina	"Oggi ti amo di più"	1988
164.	Mina	"Ridi pagliaccio"	1988
165.	Morandi	"Le italiane sono belle"	1987
166.	Morelli	"Leano Morelli"	1978
167.	Nannini	"Puzzle"	1984
168.	Nannini	"Malafemmina"	1988
169.	New Trolls	"Ut"	1972
170.	Nomadi	"Ma noi no"	1992
171.	Orme	"Contrappunti"	1974
172.	Orme	"Storia o leggenda"	1977
173.	Orme	"Verità nascoste"	1977
174.	Osanna	"Palepoli"	1973
175.	Oxa	"E' tutto un attimo"	1986
176.	Oxa	"Tutti i brividi del mondo"	1989
177.	Oxa	"Di questa vita"	1992
178.	Paoli	"Cosa farò da grande"	1987
179.	Paoli	"L'ufficio delle cose perdute"	1988
180.	Pfm	"Jet lag"	1977
181.	Pitura Freska	"Na brutta banda"	1992
182.	Pooh	"Opera prima"	1971
183.	Pooh	"Alessandra"	1973
184.	Pooh	"Fforse ancora poesia"	1975
185.	Pooh	"Un po' del nostro tempo migliore"	1975
186.	Pooh	"Rotolando respirando"	1977
187.	Pooh	"Buona fortuna"	1981
188.	Pooh	"Aloha"	1984
189.	Pooh	"Un altro pensiero"	1989
190.	Pooh	"Uomini soli"	1990
191.	Pooh	"Il cielo è blu sopra le nuvole"	1992
192.	Pravo	"Di vero in fondo"	1971
193.	Pravo	"Per aver visto un uomo piangere"	1971
194.	Pravo	"Pazza idea"	1973
195.	Pravo	"Mai una signora"	1974
196.	Pravo	"Incontro"	1975
197.	Pupo	"Lo devo solo a te"	1979
198.	Ramazzotti	"Cuori agitati"	1985
199.	Ramazzotti	"Musica è"	1988
200.	Rettore	"Rettoressa"	1988
201.	Ricchi e Poveri	"E penso a te"	1981

	Autore	Titolo	Anno
202.	Ricchi e Poveri	"Mamma Maria"	1982
203.	Rocchi	"Volo magico n.1"	1971
204.	Ron	"Una città per cantare"	1980
205.	Ron	"Calypso"	1983
206.	Ron	"Tutti cuori viaggianti"	1983
207.	Ron	"Apri le braccia e poi vola"	1990
208.	Ron	"Le foglie e il vento"	1992
209.	Rossi Vasco	"Va bene va bene così"	1984
210.	Rossi Vasco	"Liberi liberi"	1989
211.	Rossi Vasco	"Vasco live"	1990
212.	Rossi Vasco	"Vaggiando"	1991
213.	Rossi Luciano	"Gazie"	1981
214.	Ruggeri	"Enrico VIII"	1986
215.	Ruggeri	"Difesa francese"	1986
216.	Ruggeri	"La parola ai testimoni"	1988
217.	Ruggeri	"Contatti"	1989
218.	Ruggeri	"Peter Pan"	1991
219.	Sorrenti	"Aria"	1972
220.	Sorrenti	"Di notte"	1980
221.	Sorrenti	"Radici"	1992
222.	Squallor	"Arraphao"	1983
223.	Stadio	"Siamo tutti elefanti inventati"	1991
224.	Tazenda	"Limba"	1992
225.	Togni	"Segui il tuo cuore"	1986
226.	Tozzi	"E' nell'aria ti amo"	1977
227.	Tozzi	"Gloria"	1979
228.	Tozzi	"Notte rosa"	1981
229.	Vallesi	"Paolo Vallesi"	1991
230.	Vandelli	"29 settembre '89"	1989
231.	Vanoni	"Ah l'amore l'amore quante cose..."	1971
232.	Vanoni	"L'amore"	1972
233.	Vanoni	"Dettagli"	1973
234.	Vanoni	"La voglia di sognare"	1974
235.	Vanoni	"A un certo punto"	1974
236.	Vanoni	"2301 parole"	1981
237.	Vanoni	"Ornella"	1986
238.	Vanoni	"Stella nascente"	1992
239.	Vanoni & Paoli	"Insieme"	1985
240.	Vecchioni	"Robinson"	1979
241.	Vecchioni	"Hollywood Hollywood"	1982
242.	Vecchioni	"Ippopotami"	1986
243.	Vecchioni	"Milady"	1989
244.	Venditti	"Quando verrà natale"	1974
245.	Venditti	"Lilly"	1975
246.	Venditti	"Sotto il segno dei pesci"	1978
247.	Venditti	"Buona domenica"	1979
248.	Venditti	"Cuore"	1984
249.	Venditti	"Venditti e segreti"	1986
250.	Venditti	"Centocitta"	1987
251.	Venditti	"Da San Siro a Samarcanda"	1992
252.	Wess & Dory Ghezzi	"Wess & Dory Ghezzi"	1974
253.	Zanicchi	"Caro Theodorakis"	1971

	Autore	Titolo	Anno
254.	Zero	"Soggetti smarriti"	1986
255.	Zucchero	"Rispetto"	1986
256.	Zucchero	"Blues"	1987
257.	Zucchero	"Live at the Kremlin"	1991
258.	Zucchero	"Miserere"	1992