



Luca Giuliano
Gevisa La Rocca

L'ANALISI AUTOMATICA E SEMI-AUTOMATICA DEI DATI TESTUALI

II
STRATEGIE DI RICERCA E APPLICAZIONI



ISBN 88-7916-472-6

Copyright 2010

LED Edizioni Universitarie di Lettere Economia Diritto

Via Cervignano 4 – 20137 Milano

Catalogo: www.lededizioni.com – E-mail: led@lededizioni.com

I diritti di traduzione, di memorizzazione elettronica e pubblicazione
con qualsiasi mezzo analogico o digitale
(comprese le copie fotostatiche e l'inserimento in banche dati)
sono riservati per tutti i paesi.

Le fotocopie per uso personale del lettore possono essere effettuate nei limiti del 15% di ciascun volume o fascicolo di periodico dietro pagamento alla SIAE del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941 n. 633.

Le riproduzioni effettuate per finalità di carattere professionale, economico o commerciale o comunque per uso diverso da quello personale possono essere effettuate a seguito di specifica autorizzazione rilasciata da:

AIDRO, Corso di Porta Romana n. 108 – 20122 Milano

E-mail segreteria@aidro.org – sito web www.aidro.org

In copertina:

R. Lull, *Arbre de ciencia*, Barcelona, Pere Posa, 1505

Stampa: Digital Print Service



INDICE

PREFAZIONE	7
1. Informatica e statistica (p. 8). – 2. Chi ama il confronto non teme la misura (p. 9). – 3. Un sussidio artificiale al servizio dell'intelligenza (p. 11). – 4. Il lento, il veloce e il digitale (p. 13). – 5. L'equilibrio tra incertezza e ambiguità (p. 14). – 6. La sintesi possibile tra spiegazione e interpretazione (p.15).	
1. I DISCORSI DEL GIURAMENTO: OBAMA E BUSH A CONFRONTO	19
1.1. Introduzione (p. 19). – 1.2. Ipotesi di lavoro (p. 21). – 1.3. Il corpus Obama-Bush: prime osservazioni (p. 21). – 1.4. Analisi delle forme pronominali e delle parole specifiche (p. 23). – 1.5. I discorsi di Obama e di Bush osservati attraverso l'indice di negatività (p. 27). – 1.6. Conclusione (p. 29). – Strategia di analisi (p. 29) – Riferimenti bibliografici (p. 34).	
2. CONSUMI E IMMAGINI DAI <i>TWEENS</i> . I FANS DI HANNA MONTANA	35
2.1. Introduzione (p. 35). – 2.2. I temi delle discussioni online (p. 36). – 2.3. Miley/Hanna e Hanna/Miley: elementi di identificazione (p. 38). – 2.4. I <i>tweens</i> come target di vendita (p. 40). – 2.5. Lo slang (p. 41). – 2.6. Delineando gli scenari (p. 43). – Strategia di analisi (p. 44). – Riferimenti bibliografici (p. 51).	
3. IL LESSICO DELLA GUERRA NELLA COMUNICAZIONE ONLINE	53
3.1. Introduzione (p. 53). – 3.2. Ipotesi di lavoro (p. 55). – 3.3. Il corpus I-raq28 (p. 56). – 3.4. Il sub-corpus Guerra-Iraq (p. 58). – 3.5. Analisi delle corrispondenza dei segmenti ripetuti (p. 58). – 3.6. Conclusione (p. 64). – Strategia di analisi (p. 65). – Riferimenti bibliografici (p. 69).	
4. FRAMMENTI DI FANTASCIENZA ONLINE E L'IMMAGINAZIONE DEL FUTURO	71
4.1. Introduzione (p. 71). – 4.2. La tecnica di analisi (p. 74). – 4.3. Le caratteristiche del futuro (p. 77). – 4.4. La costruzione di una realtà (p. 80). – 4.5. Chi(osare) (p. 81). – Strategia di analisi (p. 83). – Riferimenti bibliografici (p. 90).	

5. L'IMMAGINE DELLA CINA NELLA STAMPA QUOTIDIANA DAL 2000 AL 2008	93
5.1. Introduzione (p. 93). – 5.2. Ipotesi di lavoro (p. 95). 5.3. Il corpus Cina2008 (p. 96). – 5.4. Analisi delle corrispondenze sulle forme peculiari e specifiche del corpus Cina2008 (p. 99). – 5.5. Il tono positivo e negativo degli articoli nella stampa (p. 107). – 5.6. Conclusione (p. 109). – Strategia di analisi (p. 110). – Riferimenti bibliografici (p. 113).	
6. ISTANTANEE DELLA FAMIGLIA IN PUBBLICITÀ: TRADIZIONALE E RICOMPOSTA	115
6.1. Introduzione (p. 115). – 6.2. Immagini della famiglia tradizionale (p. 117). – 6.3. Immagine della famiglia ricomposta (p. 119). – 6.4. Alcune considerazioni (p. 123). – Strategia di analisi (p. 124). – Riferimenti bibliografici (p. 133).	
7. LE CANZONI DEL FESTIVAL DI SANREMO: LESSICO E ANALISI DEL CONTENUTO	135
7.1. Introduzione (p. 135). – 7.2. Ipotesi di lavoro (p. 136). – 7.3. Il corpus Sanremo2010 (p. 136). – 7.4. Le fasi di sviluppo del lessico delle canzoni di Sanremo (p. 139). – 7.5. Parole vecchie e nuove (p. 146). – 7.6. Conclusione (p. 150). – Strategia di analisi (p. 152). – Riferimenti bibliografici (p. 153).	
8. SELF REMEDIATION: L'IMPORTANZA DI PARTECIPARE AL GRANDE FRATELLO	155
8.1. Introduzione (p. 155). – 8.2. La ricerca (p. 156). – 8.3. I significati della partecipazione al Grande Fratello (p. 158) – 8.4. Delineare i tratti distintivi delle dimensioni (p. 162). – Strategia di analisi (p. 163). – Riferimenti bibliografici (p. 166).	

Il progetto del volume è equamente condiviso dai due autori. La prefazione e i capitoli 1, 3, 5, 7 sono stati scritti da Luca Giuliano; i capitoli 2, 4, 6 e 8 sono stati scritti da Gevisa La Rocca.



PREFAZIONE

Questo volume rappresenta la continuazione ideale di un lavoro pubblicato nel 2008: *L'analisi automatica e semi-automatica dei dati testuali. Software e istruzioni per l'uso* (L. Giuliano e G. La Rocca, LED).

Le ricerche qui presentate fanno uso delle tecniche e dei software illustrati là con dettaglio e sulla base degli approcci metodologici argomentati e discussi soprattutto nei capitoli 1, 2, 3 e 7 dedicati all'integrazione tra qualità e quantità, alla *Grounded Theory* e all'analisi lessicometrica. La conoscenza approfondita degli uni e degli altri non è strettamente necessaria per acquisire una conoscenza critica dei risultati delle ricerche e dei temi affrontati. Tuttavia, le applicazioni dei diversi software utilizzati nell'analisi dei dati testuali hanno lo scopo di fornire all'utente una spiegazione dei percorsi e delle strategie che sono state adottate per compiere il consueto cammino che dalla formulazione delle ipotesi porta al loro controllo empirico. Al termine di ciascun capitolo sono sviluppati e descritti, anche con riferimenti al software, i passaggi principali che hanno guidato l'analisi fase per fase.

Le riflessioni che seguono ¹ si propongono di aiutare il lettore a inquadrare queste scelte strategiche e pertanto non riguardano i contenuti specifici di ciò che verrà trattato nei diversi capitoli, quelli che di solito vanno a formare la “parte sostantiva” di una ricerca. Si tratta piuttosto di alcune considerazioni generali (e personali) che riguardano la metodologia delle scienze sociali intesa come gestione intelligente e consapevole delle informazioni e dei dati empirici. Oggi più che mai la critica dei metodi e delle tecniche di ricerca, per essere socialmente accreditata, deve sollecitare la formazione di competenze

¹ Il testo è una rielaborazione con integrazioni dell'intervista rilasciata alla redazione della rivista *Informatica Umanistica* in occasione della pubblicazione del vol. 1 (2009): <http://www.ledonline.it/informatica-umanistica/>

che rispondano ai problemi posti dalla convergenza di complessità, incertezza e variabilità dei fenomeni sociali. Il fatto che in questo volume i fenomeni oggetto di studio siano dei comportamenti sociali che si manifestano come “fenomeni linguistici”, “visivi” o “comunicativi” non deve apparire come un vincolo limitativo: la lingua non descrive il mondo che c’è là fuori; la lingua è parte costitutiva del mondo che è frutto delle nostre interazioni sociali e dei processi che ci fanno essere quello che siamo.

1. INFORMATICA E STATISTICA

Nel 1975 ero un giovane borsista presso l’istituto di Statistica e ricerca sociale della Facoltà di Scienze statistiche, alla Sapienza, e mi capitò di assistere a un seminario di Jean-Paul Benzécri sulle applicazioni dell’analisi delle corrispondenze a dati linguistici. Proprio su questo tema dalla metà degli anni Sessanta Benzécri stava sviluppando le sue prime sperimentazioni di *analyse des données* mettendo in evidenza come la matematica del finito e della combinatoria che si rifà a strutture algebriche astratte permettesse di ricostruire modelli di senso più semplici soggiacenti alla complessità presente in un corpus di testi. Si trattava di mettere a frutto le nascenti disponibilità di calcolo automatico per realizzare il sogno degli studiosi di statistica del passato, a partire da Charles Spearman, il pioniere dell’analisi fattoriale (1904), fino a Louis L. Thurstone (1931) e, soprattutto, Harold Hotelling con l’analisi in componenti principali (1933) e l’analisi della correlazione canonica (1936). L’idea era di trovare delle variabili “di sintesi”, come combinazione lineare delle variabili originarie, in grado di massimizzare la riproduzione della variabilità complessiva di una matrice di dati. L’obiettivo era prevalentemente descrittivo, anziché esplicativo, e pertanto non era necessario stabilire a priori alcuna distinzione tra variabili indipendenti e variabili dipendenti, come accadeva invece nella classica analisi multivariata inaugurata da Sewell Wright con la *path analysis* (1918-1934) e da Ronald Fisher con l’analisi della varianza (1926-1935). L’analisi in componenti principali apriva la strada all’analisi multidimensionale su matrici di centinaia di variabili, cardinali e categoriali, con il contributo decisivo di Louis Guttman (1941), Cyril Burt (1950), Maurice G. Kendall (1955) e John W. Tukey (1960), che anticipavano tecniche che saranno poi pienamente sviluppate grazie alle scienze dell’informazione. Non è un caso che due dei termini più noti dell’informatica, *bit* (*binary digit*) e *software*, furono conati proprio da Tukey, il primo negli anni Quaranta, mentre lavorava nei laboratori della AT&T al fianco di Claude E. Shannon che ne fece uso in una pubblicazione del 1948; il secondo nel 1958 in un articolo pubblicato dallo stesso Tukey sull’*American Mathe-*

matical Monthly (la paternità fu poi rivendicata successivamente da Paul Niquette che riteneva di averlo utilizzato già dal 1953).

L'analisi delle corrispondenze di Benzécri possedeva i tratti di una sintesi efficace e affascinante, soprattutto se applicata alle proprietà squisitamente qualitative delle parole e del lessico. L'informatica ha reso possibile il dialogo tra contesto della scoperta e contesto della giustificazione, tra comprensione e spiegazione, tra intuizione e onere della prova, tra momento induttivo e momento deduttivo, tra parole e numeri. Già nella seconda metà del Seicento, Leibniz – da matematico e logico geniale ma anche da umanista memore della lezione rinascimentale – abbozzò un progetto di linguaggio (*calculus*) universale in cui tutti gli elementi essenziali del pensiero avrebbero dovuto trovare una corrispondenza simbolica tale da permettere lo sviluppo di un'arte per la soluzione di tutti i problemi. Oggi le scienze umane hanno raccolto questa sfida e, dopo oltre un secolo di malinteso positivismo, sono sempre più numerosi coloro i quali concepiscono il sapere come unitario, ricordando che Galileo Galilei descriveva se stesso come un filosofo e un umanista. In questa prospettiva sono auspicabili percorsi diagonali di ogni tipo e incontri disciplinari che permettano lo scambio di procedure e metodi tra campi di applicazione diversi, sfruttando analogie e punti di convergenza in cui l'informatica – per la sua logica combinatoria – è particolarmente adeguata e creativamente produttiva.

2. CHI AMA IL CONFRONTO NON TEME LA MISURA

Louis De Jaucourt nella *Encyclopédie* di Diderot e D'Alambert definisce la scienza come “conoscenza chiara e certa di qualche cosa, fondata o su principi evidenti per se stessi o su dimostrazioni”. La voce “Scienza” compilata da Jaucourt traccia i confini di un sapere che è fortemente ancorato all'empirismo e al naturalismo di Bacone, ma non esita a richiamare l'attenzione del lettore sulla necessità di essere uomini di lettere per accedere a questa conoscenza. Siamo ancora lontani da quella dissennata separazione tra pensiero umanistico e pensiero scientifico che produrrà i suoi frutti più perversi nella vulgata scienziata tra la fine dell'Ottocento e gli inizi del Novecento. Eppure, prima dei cultori di discipline umanistiche, sono stati proprio i “maniaci della misura” a porre problemi inconsueti e nuovi che derivavano soprattutto dal successo della statistica e dalla sua iniziale frenesia deterministica. Lo sguardo “quanto-frenico” di questi precursori oggi li potrebbe esporre facilmente al ridicolo, ma non dobbiamo dimenticare che essi operavano in una cultura dominante, specialmente nei paesi della Riforma, che aveva sostituito il determinismo teologi-

co di Lutero e Calvino con il determinismo scientifico di Newton.

Così il grande matematico Johann Bernoulli, contemporaneo di Leibniz, nei suoi diari di viaggio ci fornisce più informazioni sulle dimensioni dei quadri antichi che aveva avuto l'occasione di osservare che non sui loro autori e sul loro valore artistico. John Arbuthnot, medico e matematico scozzese, amico di Swift e di Newton, in una memoria del 1710 per la Royal Society pretendeva di aver dimostrato l'esistenza della provvidenza divina sulla base del fatto che i maschi alla nascita erano più delle femmine per compensare le morti più frequenti di quest'ultimi in guerra e nei naufragi. Johann P. Süssmilch, cappellano di Federico II di Prussia e membro dell'Accademia delle scienze di Berlino, pretendeva di aver stabilito una misura dell'ordine divino attraverso un'analisi dettagliata delle statistiche demografiche.

Al di là di queste annotazioni curiose, l'approccio quantitativo ha dato un contributo importante alla linguistica grazie alle osservazioni induttive di studiosi come il matematico russo Viktor Bunjakovskij, che fu tra i fondatori della statistica in Russia con Pafnuty Chebyshev, e che nei suoi lavori tracciò un abbozzo di aritmetica del linguaggio che tenesse conto della frequenza delle parole e della loro lunghezza. Altri contributi analoghi sono venuti dagli inventori della stenografia, come Jean-Baptiste Estoup (1907), e dalla psicolinguistica di Adolf Buseman, con il suo studio sul linguaggio dei bambini (1925). George K. Zipf, positivista convinto, con il suo "principio di frequenza relativa" delle forme grafiche, negli anni Trenta ha aperto la strada alla lessicometria e all'analisi quantitativa del linguaggio.

Certamente oggi noi sappiamo, grazie anche a Charles S. Peirce, che l'induzione non può essere sorretta dalla mera osservazione dei casi ripetuti, pur se numerosi. Osservare e misurare la ripetibilità dei fenomeni e trovare risultati che si conformano alle legge di Gauss non significa poter concludere che la curva normale sia la rappresentazione di una legge naturale. Il caso non è, come credevano i deterministi, l'espressione della nostra ignoranza sulle vere cause dei fenomeni osservati. Semmai è logicamente più plausibile affermare che le cause verosimili emergono da processi casuali. L'interesse degli statistici non è più rivolto all'accertamento della media ma alla stima incerta delle variazioni probabili rispetto alla media. Seguendo un suggerimento di Nassim N. Taleb, se non si sa nuotare è sconsigliato guardare un fiume del quale è nota la profondità media di un metro e mezzo. È molto più importante conoscere il campo di variazione degli esiti possibili che non il risultato finale delle osservazioni in un determinato periodo di tempo.

Le scienze sociali, oggi, hanno tutto da guadagnare dalle immense risorse offerte dall'informatica e dalle metodologie di analisi sviluppate nell'ambito delle scienze statistiche e della probabilità. L'analisi dei processi di comunica-

zione non può prescindere ormai dalla sempre più diffusa digitalizzazione delle informazioni e delle grandi basi di dati testuali che ne sono la conseguenza. Bibliografia, classificazione dei testi e dei documenti, editoria elettronica, gestione della conoscenza basata su dati non strutturati, come ad esempio nella gestione della documentazione giuridica, pongono continuamente problemi nuovi le cui soluzioni non possono che scaturire dalla collaborazione intensa tra informatici, statistici ed esperti delle discipline sostantive.

3. UN SUSSIDIO ARTIFICIALE AL SERVIZIO DELL'INTELLIGENZA

Ci troviamo sui confini di un territorio ancora in gran parte ignoto che dobbiamo esplorare con flessibilità critica e apertura mentale, disponibili a tollerare l'inadeguatezza dei nostri modelli senza percorrere precipitosamente strade già note perché esse sarebbero inevitabilmente condannate a prospettare soluzioni euristicamente regressive, come direbbe Imre Lakatos. Tuttavia il lavoro scientifico nella sua routine quotidiana solo raramente è innovativo. C'è molta ipocrisia accademica nella sopravvalutazione della originalità della ricerca. Lo studioso serio si lascia guidare dal rigore e, quando è necessario, esercita la sua creatività cercando di cogliere i segnali di errore che incontra lungo il cammino. La ricerca scientifica è tale proprio perché nel metodo è disponibile ad ammettere il suo fallimento e a ricominciare da capo, passando da una domanda all'altra ed escludendo di volta in volta le risposte che si sono dimostrate sbagliate. Dobbiamo essere preparati ad esaltare l'asimmetria tra verifica e falsificazione delle congetture teoriche come ci propose Karl Raimund Popper quando, di fronte alla imprevedibilità del mondo, ci invitò ad accettare l'idea che era più ragionevole l'inferenza negativa piuttosto che l'inferenza positiva: migliaia di prove a sostegno di un'ipotesi non dimostrano nulla, mentre una sola prova contraria è sufficiente a demolirla per sempre.

Rispetto alle modalità di insegnamento delle scienze sociali e delle discipline umanistiche, la cultura scientifica dei fisici, dei chimici, dei biologi è predisposta a favorire l'innovazione. Gli studenti universitari di scienze matematiche, fisiche e naturali si preparano alla professione su manuali scritti appositamente per loro nei quali gli elementi fondamentali di conoscenza sono sintetizzati e predisposti per delineare lo "stato dell'arte" e metterli nelle condizioni di utilizzare le competenze acquisite nelle applicazioni richieste. Al termine del percorso formativo, nella tesi finale si chiede loro di sviluppare un progetto, più o meno articolato e sperimentale, ma quasi sempre orientato alla riformulazione di un problema che trova classicamente espressione in una ipo-

tesi di ricerca. Raramente questo tipo di studente viene invitato a leggere i classici e a percorrere storicamente e criticamente lo sviluppo delle discipline che fanno parte del proprio orizzonte culturale. La prova finale di laurea lo porterà a leggere i prodotti di ricerca più recenti e innovativi, ad assumere modalità di comunicazione e linguaggio caratteristici della propria specializzazione e – quando la sua formazione è completa ed efficace - a saper aggiornare continuamente le conoscenze acquisite se non a innovarle diventando egli stesso un ricercatore. Gli studenti di storia, di lingue, filosofia, archeologia, letteratura sono invece costantemente incoraggiati a documentarsi sull’immenso patrimonio culturale che li ha preceduti, sulla storia e sullo sviluppo delle loro discipline, sulla critica delle fonti, sulla struttura semantica e sintattica degli schemi interpretativi che si sono succeduti nel tempo. La formazione di uno studente di fisica avviene prevalentemente in una cornice conoscitiva che azzerla la tradizione, rendendo le discipline indifferenti rispetto alla loro esistenza storico-sociale nella prospettiva, che si è rivelata effettivamente efficace, di porre in essere innovazioni teoriche e, soprattutto, applicative e tecnologiche. La formazione di uno studente di lettere avviene invece sotto il peso di una tradizione soverchiante, in cui lo sviluppo della conoscenza è soprattutto approfondimento e risistemazione dell’esistente, focalizzazione di dettagli, recupero o aggiornamento di nuove interpretazioni di ciò che era già stato interpretato.

In questo quadro, l’informatizzazione dei documenti passati al setaccio della tradizione alleggerisce il peso della memoria. In un recente convegno promosso dalla Fondazione Telecom Italia al Politecnico di Torino (*2060: con quali fonti si farà la storia del nostro presente?* 8-9 aprile 2010) storici, sociologi, archivisti, filosofi e informatici si sono interrogati sui problemi connessi con quanto del nostro presente è già consegnato al futuro in formato digitale per il solo fatto di essere nato nella “società in rete” (comunicazioni interpersonali, transazioni economiche e finanziarie, pubblicazioni e attività di ricerca scientifica, cronaca quotidiana e dibattito politico, solo per ricordarne una piccola parte). Ma a questo si aggiunge il crescente impegno con cui archivi e biblioteche stanno digitalizzando in modo retrospettivo testi e documenti analogici che fanno parte del loro patrimonio e che per la prima volta verranno messi a disposizione di una consultazione di massa (per esempio la Bibbia “a 42 linee” di Gutenberg resa disponibile online sul sito della British Library).

Il software rende possibile la formalizzazione e la disposizione ordinata delle cose. In un certo senso le procedure informatiche e la digitalizzazione delle conoscenze rappresentano la “nuova mnemotecnica” del nostro tempo: un sussidio artificiale al servizio dell’intelligenza. Esse permettono al “pensiero divergente”, che sperimenta nuovi approcci e nuove soluzioni giocando con i

mondi possibili, di muoversi con flessibilità tra le diverse opzioni senza smarrire il fulcro del proprio lavoro ancorato alla operatività del “pensiero convergente”, mettendo nel contempo a frutto quella originalità creativa che è pur sempre auspicabile per favorire lo sviluppo della conoscenza.

4. IL LENTO, IL VELOCE E IL DIGITALE

Fin qui credo di aver prospettato i vantaggi per le scienze e la cultura umanistica dell’approccio sistematico e procedurale offerto dalla digitalizzazione dei documenti. Il prodotto di sintesi più noto che l’informatica stessa ha sviluppato sotto la diretta influenza di esigenze tipiche delle discipline letterarie, storiche e sociali è l’ipertesto. Da quando Vannevar Bush nel 1945 diffuse l’idea del “memex” (*memory extension*) come strumento di consultazione e indicizzazione associativa di un archivio, secondo il modello di funzionamento della mente che non è lineare ma connettivo, fino alla sua prima realizzazione in *Xanadu* (1960), prototipo vero e proprio della *Literary Machine* di Ted Nelson, e poi alle sue applicazioni tecnologiche concrete del *World Wide Web* di Tim Berners Lee (1991) e della piattaforma *Wiki* di Ward Cunningham (1995), l’ipertesto e la nozione di link sono entrati a far parte nel nostro orizzonte culturale a tal punto da averne dimenticato la costituzione recente e da averlo assimilato come modalità naturale della produzione intellettuale. Oggi per noi il link è come un nuovo segno di interpunzione nel testo scritto: una convenzione normativa che aiuta a rappresentare nel testo l’organizzazione del pensiero. La punteggiatura si è presentata nella scrittura occidentale con il prevalere della lettura silenziosa, mentale e veloce, sulla lettura ad alta voce, lenta e scandita dalle esigenze della respirazione. Il link (e l’ipertesto che lo giustifica) rappresenta oggi lo sviluppo della lettura digitale, guidata dalla mente e resa operativa dalla tastiera del computer o dal *touch screen* dell’iPad e dei pc-tablet che si avvia a sostituirla.

Gran parte delle innovazioni che si sono succedute in questi anni nella *Information Technology* sono state il prodotto di processi che sono iniziati nelle comunità virtuali e nei sistemi organizzativi in rete. Il cosiddetto Web 2.0 è fondato sul concetto di interattività, performance, opera aperta, interpretazione soggettiva e spazio cognitivo dell’utente, costruzione di senso, pragmatica esperienziale, scambio tra sistemi informativi, tutti elementi che fanno capo all’interazione sociale, alle forme di comunicazione e alla “connettività totale” più che alle pure esigenze di supercalcolo che erano all’origine dello sviluppo dell’informatica. Ciò che era lento nella trasmissione della cultura orale è di-

ventato veloce nell'affermarsi della cultura scritta. Ora la cultura digitale si affranca dallo stesso fluire del tempo e della sua irreversibilità. L'ipertesto non ha né un prima né un dopo, è chi lo usa che ne stabilisce la sequenza. Nella società dell'ipertesto l'annuncio di un evento, come ha dimostrato il reverendo Jones con la sua promessa di bruciare copie del Corano nel nono anniversario dell'11 settembre, si fa evento prima ancora del suo accadere e produce i suoi effetti devastanti di morte e distruzione prima ancora che accada e, come ora sappiamo, senza che accada. Il testo digitalizzato (in questo caso il rimbalzare sui social network della dissennata minaccia di uno pseudoprofeta malato di protagonismo) non è né lento né veloce, è semplicemente attuale.

5. L'EQUILIBRIO TRA INCERTEZZA E AMBIGUITÀ

L'informatica si presenta all'esperienza dell'utente come applicazione tecnica. È stato sempre così, già da quando nel 1834 Charles Babbage progettò (senza portarla a termine) la "macchina analitica" (*analytical engine*) che è considerata l'antenata del computer. La macchina analitica era un dispositivo che avrebbe dovuto eseguire delle sequenze di operazioni (come un odierno programma di software) sulla base di due congegni costituiti di complessi cilindri dentati (la memoria e l'unità centrale) che ricevevano istruzioni da schede perforate.

L'applicazione pratica, la semplificazione operativa dell'algoritmo, spesso occultano le scelte teoriche che danno una forma al mondo. Tuttavia non vi è un esercizio del fare senza presupporre un orizzonte di distinzioni e differenze, più o meno consapevoli. La scienza, rappresentando il punto più alto di un sapere che non si limita alla trasmissione di quanto si è acquisito ma si propone come un modello di accrescimento del sapere, opera consapevolmente in un mondo di differenze che, come ha ben argomentato Gregory Bateson², debbono a loro volta esser differenziate e classificate. Le implicazioni teoriche sono dunque inevitabili, oltre che auspicabili, e ignorarle può soltanto indurre in equivoci ed errori.

Il sapere critico e, più in generale, una riflessione su come dobbiamo procedere e sui modi migliori di procedere nella conoscenza del mondo, non solo ha aperto la strada all'informatica, alle scienze decisionali e alle scienze cognitive, ma rappresenta oggi il terreno ideale di incontro tra studiosi di discipline diverse nell'oggetto ma unite da un problema comune: il difficile equilibrio tra incertezza e ambiguità. Da una parte, l'incertezza che deriva dalla dif-

² G. Bateson, "Forma, sostanza e differenza", *General Semantics Bulletin*, 37, 1970; ripr. in *Verso un'ecologia della mente*, Milano, Adelphi, 1976, pp. 488-508.

ficoltà di leggere e spiegare le immense variazioni della realtà osservata; dall'altra l'ambiguità che rappresenta la difficoltà di scegliere tra diverse interpretazioni della realtà, sempre molteplici e spesso contraddittorie.

L'informatica ha aperto la strada ad applicazioni che intersecano scienze umane e scienze naturali (una distinzione a mio avviso insostenibile sul piano epistemologico e che trova qualche giustificazione solo nella didattica) dimostrando che lo sguardo del soggetto conoscente è sempre umano mentre l'oggetto del conoscere è indeterminato e sfuggente. Tuttavia, proprio grazie alla loro indeterminatezza gli eventi sono sottoposti a controllo tramite le leggi del caso. Nello spazio digitale valanghe di numeri, dominati dalla probabilità e dalla statistica, acquistano senso e ci restituiscono un'immagine del mondo "usabile" per le nostre esigenze. Ma senza la mediazione delle idee e del sapere critico, cioè di "differenze che producono la differenza" (e qui è ancora Bateson che ci offre il suo aiuto), non potremmo arginare in nessun modo l'incertezza dei nostri modelli né dominarne l'ambiguità.

6. LA SINTESI POSSIBILE TRA INTERPRETAZIONE E SPIEGAZIONE

L'immenso contributo delle nuove tecnologie digitali nella gestione di grandi masse di informazioni e dati ha interessato tutte le discipline empirico-formali che puntano a un certo grado di generalizzazione dei risultati. Sul versante strettamente metodologico esse non hanno avuto conseguenze specifiche per le scienze sociali tali da dover essere considerate come un "mondo a parte". Le conseguenze della digitalizzazione, invece, sono state decisive per il modo di concepire i "documenti" di cui le scienze sociali fanno largo uso e che costituiscono le basi empiriche di controllo delle ipotesi di ricerca. I documenti di ricerca, se non sono preparati nel momento della loro rilevazione in una forma strutturata e classificabile a priori, come accade in un test o in un questionario i cui risultati sono trasferiti immediatamente in codici alfanumerici, sono irriducibilmente ancorati alla loro forma fisica originaria: la lettera, l'articolo di giornale, il manifesto, il diario, la fotografia ecc. I documenti analogici, prodotti per qualsiasi motivo nella vita sociale e solo in un secondo momento portati all'attenzione del sociologo o dell'antropologo sulla base di specifici interessi e scelte di ricerca, non possono essere sottoposti alla generalizzazione empirica se non con una immensa fatica e con un enorme impiego di tempo. Thomas e Znaniecki nel 1920 portarono a termine il loro lavoro *The Polish Peasant in Europe and America* analizzando in sei anni almeno mille lettere e ottomila documenti vari tratti da giornali dell'epoca, ma nessuno sa esattamente quanti fos-

sero, quali furono scartati e perché. Questi problemi furono messi in luce molto tempo dopo la pubblicazione, in un convegno del 1938 durante il quale si seppe che gran parte della documentazione originale era andata distrutta. Una base di dati fondata sull'analisi dei documenti naturali, tipicamente d'archivio, ha un carattere difficilmente ispezionabile, se non riproducendo il percorso, spesso singolare e intuitivo, dello studioso che lo ha compiuto.

La digitalizzazione dei documenti, in particolare dei testi, può cambiare radicalmente questa situazione. I testi possono essere condivisi da diversi ricercatori. L'analisi del contenuto può essere condotta con software CAQDAS (*Computer Assisted Qualitative Data Analysis Softwares*) appositamente sviluppati per seguire il percorso di concettualizzazione, di operativizzazione e di classificazione. Testi non strutturati, come le interviste, i focus group, i forum in Internet, le email o i messaggi in Facebook, possono essere classificati automaticamente ed essere restituiti all'esame critico della comunità scientifica dotati di tutte le meta-informazioni che permettono la ricostruzione e la trasparenza delle scelte metodologiche compiute dai ricercatori.

L'associazione e l'incontro tra l'analisi qualitativa assistita dal computer e l'analisi statistica dei dati testuali rendono operativo metodologicamente il "circolo dell'interpretazione" teorizzato da Paul Ricoeur³, la dialettica tra spiegare e comprendere. Parlare di "ermeneutica quantitativa" significa esprimersi con un ossimoro un po' provocatorio che, tuttavia, trova una sua giustificazione nel tentativo di conciliare il lavoro di interpretazione che si focalizza sul testo, sulla intenzionalità dell'autore, sull'universo che in esso si esprime e sulla sua singolare costruzione di senso con il lavoro di spiegazione che opera un distacco del testo dal suo autore, che decontestualizza le parole per portarle a sostegno di ipotesi di ricerca, sottoporle a confronti, misure, formalizzazioni che ne permettono la generalizzazione.

L'ermeneutica quantitativa è una prospettiva – non ancora realizzata ma potenzialmente presente nella digitalizzazione del testo – che permette l'incontro tra scienze che hanno interesse a mettere in evidenza le proprietà singolari e uniche degli oggetti e scienze che hanno interesse a selezionare le proprietà che costituiscono classi di oggetti. Il ricercatore si muove così dal testo preso nella sua completezza per risalire alle parole che costituiscono il testo. Questo "doppio movimento" dalla qualità alla quantità e dalla quantità alla qualità diventa essenziale e irrinunciabile quando il ricercatore ha a disposizione basi testuali e corpora costituiti da centinaia di migliaia o da milioni di oc-

³ Il tema è stato trattato da P. Ricoeur a partire da un articolo del 1970, "Che cos'è un testo?", ripr. in *La sfida semiologica*, Roma, Armando Editore, 2006, pp. 212-230.

correnze. Allora, nella impossibilità di leggere il testo per interpretarne il senso egli mette in atto strategie lessicometriche e quantitative per individuare parole-chiave, segmenti ripetuti, unità semantiche che presentano qualche interesse rilevante, qualche peculiarità di presenza all'interno del corpus. Poi, successivamente, assume un punto di vista ermeneutico quando seleziona le concordanze o individua nel corpus parti di testo che presentano un interesse particolare in base alla presenza delle parole chiave individuate.

Negli esempi che seguono questa potenziale “ermeneutica quantitativa” si offre al confronto critico utilizzando le tecniche tipiche della statistica testuale e alcuni dei software più accessibili e accreditati tra i ricercatori: Atlas.ti, Nvivo, TaLTaC e SPAD. Per ogni applicazione è indicata, al termine del capitolo, la strategia di analisi utilizzata e alcune note esplicative sui diversi passaggi richiesti nella utilizzazione del software. L'auspicio, come sempre, è di aver fatto qualcosa di utile per diffondere l'uso consapevole di strumenti sempre più indispensabili per decifrare il mondo in cui viviamo.

Roma, 19 settembre 2010

LUCA GIULIANO

1.

I DISCORSI DEL GIURAMENTO: OBAMA E BUSH A CONFRONTO

1.1. INTRODUZIONE

Il 20 gennaio 2009 Barak Obama ha parlato come presidente agli Stati Uniti d'America e al mondo. Il discorso del giuramento è stato ripreso dai mezzi di comunicazione vecchi e nuovi, dai giornali alla rete, e tutti hanno sottolineato l'importanza di una cerimonia politica nazionale che mai aveva avuto un seguito così vasto. Qualcuno ha ricordato il giubileo della regina Vittoria oppure il giuramento di John F. Kennedy. Tutti i confronti sono leciti, tenendo conto della enorme diffusione che hanno oggi i media ad ogni livello, ma in ogni caso mai, in assoluto, vi è stato nella storia un momento di attesa così forte rispetto al discorso di un nuovo capo di Stato. Le parole della politica, enfatizzate dalla presenza reale e virtuale di una comunità mondiale in ascolto, sono diventate *instrumentum regni* e, come si conviene in questi casi, hanno assunto anche una funzione sacrale di rappresentazione collettiva di una società intera. È qualcosa che va al di là della consueta "politica spettacolo" per diventare una "religione della politica". In un mondo disincantato e scettico, che si dice privo di ideologie e poi si trova a condividere globalmente solo il linguaggio asettico dell'economia e della finanza, il giuramento di Barak Obama esprime il bisogno, se non la necessità, di una forza contraria, di un movimento che risale la corrente e riporta l'idea di democrazia e di pluralismo al centro della vita sociale.

L'elezione di Barak Obama alla presidenza degli Stati Uniti ha avuto un impatto molto significativo su quello che gli americani chiamano la "gamma razziale" (*racial spectrum*), in particolare tra gli americani di origine africana. Tuttavia le interpretazioni di questo importante momento storico dal punto di

vista dell'affermazione di una eguaglianza razziale nella società americana non sono state sempre concordi (Reed e Louis, 2009; Winant, 2009). Un presidente afro-americano alla Casa Bianca offre il fianco a commenti del tipo: “non ci sono più scusanti, ora, per i neri” (*now blacks have no more excuses*). Inoltre americani bianchi e neri adottano diversi criteri di giudizio per misurare le distanze razziali, in una tendenza generale che percepisce questo cambiamento sociale come un gioco a somma zero in cui a un guadagno dei Neri corrisponde una perdita dei Bianchi (Eibach e Purdie-Vaughns, 2009). Alcuni non esitano ad affermare che, pur essendo questo un passaggio storico molto importante, il significato della sua elezione in termini di eguaglianza delle opportunità nella rappresentanza politica è solo un fatto simbolico e non un fatto reale. La retorica della “razza” sarebbe servita solo a estendere la base di consenso di Obama durante la sua campagna elettorale ma nulla induce a credere che un nero nell'Ufficio Ovale possa mutare la struttura di fondo di un sistema di potere largamente fondato sulle ineguaglianze di razza, di genere e di reddito (Harlow, 2009).

L'analisi con metodi statistici del discorso tenuto da Barak Obama nella cerimonia di giuramento come presidente degli Stati Uniti si inquadra all'interno degli studi compiuti in questi ultimi anni sulla retorica del discorso politico in particolare da Norman Fairclough (*Language and Power*, 1989) e da Teun van Dijk (*Ideology: a multidimensional approach*, 1998). Il mondo della politica è uno degli ambiti in cui le pratiche sociali sono quasi completamente discorsive. Lo spazio cognitivo della politica è inevitabilmente fondato su prospettive d'azione ideologica e le ideologie politiche si riproducono tramite i discorsi.

L'analisi del discorso, negli studi sulla comunicazione linguistica, si presenta classicamente come un metodo di ricerca che studia l'enunciazione, e cioè una situazione in cui un parlante si propone intenzionalmente di influenzare un ascoltatore. Gli “eventi discorsivi” già in Michel Foucault si articolano in un contesto di rapporti tra persone (io, tu, noi, voi), nei luoghi in cui si manifesta la relazione e nello spazio-tempo che essi invocano con le argomentazioni (ieri, oggi, domani, passato, presente, futuro).

In particolare, l'analisi del *president inaugural speech* ha una tradizione piuttosto recente. I principali lavori riguardano i discorsi inaugurali di Franklin D. Roosevelt (Houck e Nocasian 2002), Bill Clinton (Winter e Weintraub, 2005), George W. Bush (Dillon *et alii*, 1990; Northcott, 2004; Gvosdev e Saunders, 2005) e Barak Obama (Berson e Berson, 2009).

1.2. IPOTESI DI DI LAVORO

In questo lavoro leggeremo, attraverso la lente dell'analisi quantitativa, le 2.365 parole (occorrenze) in cui Obama ha raccontato al mondo la sua visione del potere non per trarne delle indicazioni di tipo statistico (che sarebbero improprie per una estensione lessicale così ridotta), ma per individuarne, tramite l'analisi quantitativa, le unità semantiche più forti e le caratteristiche distintive del discorso nel suo complesso.

Le analisi linguistiche traggono maggiore forza empirica e descrittiva soprattutto quando si possono mettere a confronto i "parlanti". In questo caso, tenendo conto delle dovute differenze di contesto, il confronto può essere effettuato con il discorso del secondo giuramento di George W. Bush, nel 2005. Il discorso di Bush è complessivamente equivalente a quello di Obama per il numero di parole utilizzate (2.061 occorrenze) ma è ben differenziato per quanto riguarda i contenuti e il tono complessivo. L'analisi è stata effettuata con TaL-TaC2 utilizzando, come vedremo, risorse linguistiche esterne.

Il corpus pertanto è stato costruito (Step 1) allo scopo di rintracciare le caratteristiche distintive dei due presidenti con una esplicita ipotesi di corrispondenza tra contenuti/forme linguistiche e orientamento politico (democratico per Barak Obama e conservatore per George W. Bush). Il pragmatismo politico di Obama, nel passaggio storico che la sua elezione rappresenta, induce a ritenere che nel suo discorso si riflettano con maggiore trasparenza le difficoltà in cui versa la politica economica e internazionale americana rispetto a un presidente come Bush che, al suo secondo mandato, doveva soprattutto rassicurare i cittadini statunitensi sul problema della sicurezza e della difesa della democrazia minacciata dalle guerre che si sono succedute dopo il drammatico 11 settembre del suo primo mandato.

1.3. IL CORPUS OBAMA-BUSH: PRIME OSSERVAZIONI

Le misure lessicometriche del corpus (Step 2) ne mettono in evidenza la fragilità dal punto di vista strettamente "statistico" (tab. 1.1). Si tratta di un corpus molto piccolo, con una percentuale di hapax (parole che compaiono una sola volta nel corpus) che supera ampiamente la soglia consigliata per un'analisi quantitativa, pari al 50%; ma ancora più indicativo è il rapporto tra parole distinte (V) e occorrenze (N) del 29,37% che non permette di assumere il corpus come sufficientemente esteso dal punto di vista lessicale (soglia consigliata: inferiore al 20%). Sulla base di questi parametri possiamo condurre un'analisi quantitativa incentrata in gran parte sugli aspetti descrittivi, facendo affidamento – per l'interpretazione – sul contribu-

to che possono dare alcuni strumenti qualitativi come le “concordanze” e la classificazione delle parole.

Tab. 1.1 – Misure lessicometriche del corpus Obama-Bush

Occorrenze	N	4.426
Forme grafiche	V	1.395
Type/Token ratio	(V/N)*100	31,52
Percentuale di hapax	(V1/V)*100	66,74
Frequenza media generale	N/V	3,17

La prima osservazione che possiamo condurre riguarda la frequenza delle parole all'interno dei due testi che compongono il corpus (tab. 1.2); dalla tabella sono state escluse le parole grammaticali di minore interesse (parole “vuote”).

Tab. 1.2 – Parole piene più frequenti nei discorsi del giuramento di Obama (2009) e Bush (2005)

Forme grafiche	B. Obama	Forme grafiche	G. W. Bush
our	67	our	50
we	62	we	37
us	23	freedom	27
but	20	will	22
will	19	America	20
they	17	liberty	15
not	16	not	15
who	14	you	12
can	13	your	12
you	12	their	10
nation	12	nation	10

Nel discorso di Obama, se escludiamo le consuete parole grammaticali, le prime occorrenze significative in termini di contenuto si esprimono con *our* (67), *we* (62) e *us* (23): i “nostri antenati”, i “nostri documenti fondativi”, la “nostra economia”, prima di tutto, e poi la “nostra salute”, la “nostra scuola”, il “nostro pianeta” e così di seguito. *We* appare per la prima volta per ribadire la fedeltà agli ideali dei fondatori (*We the People have remained faithful to the ideals of our forebears*) per poi essere scandito nei primi periodi con precisi riferimenti alle future azioni di governo: *we*

will act. we will build. we will restore. we will transform. we will do (agiremo, costruiremo, restituiremo, imbrighieremo, trasformeremo, faremo); fino a ribadire quel *we can* (4) che è stato lo slogan di avvio della campagna elettorale. *Us* (23) è richiamato per tre volte facendo riferimento al legame che unisce gli americani di oggi all'oscuro lavoro di uomini e donne del passato, e per tre volte agli obiettivi e alle domande in attesa di risposta. Il primo, e unico, sostantivo ad apparire in ordine di frequenza fino alla soglia di 10 di occorrenze è *nation* (12)

Le parole che risaltano con maggiore enfasi nel discorso di Bush, oltre a *our* (50) e *we* (37), utilizzate comunque con meno intensità rispetto a Obama, sono *freedom* (27), *America* (20) e *liberty* (15). “Libertà” è in assoluto il sostantivo più frequente utilizzato da Bush, sia nella sua accezione più sociale e concreta (*history of freedom; triumph of freedom; cause of freedom; messages of freedom; america's ideal of freedom; fire of freedom; call of freedom; advance of freedom; expansion of freedom*) che nella sua versione più astratta e formale (*proclaims liberty; author of liberty; definition of liberty; promise of liberty; appeal of liberty; success of liberty; survival of liberty*).

1.4. ANALISI DELLE FORME PRONOMINALI E DELLE PAROLE SPECIFICHE

Entrando più in dettaglio nel campo semantico definito dalla forme pronominali e dagli aggettivi possiamo osservare dalle concordanze (Step 3) come in Obama *our* è riferito a un ampio spettro di temi sociali, storici e valoriali:

our ability	our commerce
our adversaries	our common dangers
our all	our common defense
our ambitions	our common good
our ancestors	our common humanity
our better history	our confidence
our borders	our creed
our business	our darkest hours
our capacity	our economy
our cars	our enduring spirit
our cause	our example
our challenges	our factories
our character	our fate
our children 's children	our forebearers
our collective failure	our founding documents

our goods and services	our politics
our government	our planet
our gross domestic product	our power
our hardship	our prosperity
our health	our revolution
our history	our safety
our ideals	our schools
our individual ambitions	our security
our journey	our spirit
our land	our success
our liberty	our system
our minds	our time
our nation	ourway of life
our patchwork heritage	ourworkers

Nel discorso di Bush l'aggettivo *our* è riferito a un campo semantico più circoscritto e generico (*country, goal, land, world*), oppure ancorato ai temi della sicurezza nazionale (*enemies, national security, soldiers, vulnerability*).

our best	our nation' security
our character	our national life
our constitution	our own
our country	our own freedom
our deepest beliefs	our own lives
our duties	our own style
our efforts	of government
our enemies	our people
our enemies's defeat	our policies
our fathers	our relations
our fellow americans	our response
our founders	our schools
our founding	our society
our friends	our soldiers
our generation	our strength
our goal	our time
our ideals	our vulnerability
our land	ourworld
our nation	our youngest citizens

Lo strumento tecnico per mettere in rilievo il diverso uso delle parole secondo i parlanti è l'analisi del linguaggio specifico (Step 4). Il corpus è di dimensioni ridotte, pertanto l'analisi è stata condotta sulle forme che hanno una frequenza maggiore o uguale a 5 (tab. 1.3).

Tab. 1.3 – Parole specifiche nei discorsi di Obama e di Bush

Parole	Occ. corpus	Occ. Obama	Parole	Occ. corpus	Occ. Bush
us	26	23	freedom	28	25
but	25	20	liberty	17	15
they	20	17	your	15	12
new	12	11	own	9	9
what	8	8	human	6	6
been	8	8	seen	6	6
let	7	7	justice	6	6
less	7	7	tyranny	5	5

Come possiamo osservare la concretezza del discorso di Obama si presenta molto sfumata nelle parole specifiche utilizzate con soglia 5. Le uniche parole con un significato distinto e non con una funzione esclusivamente grammaticale sono le forme *us* e *new*. In Bush si presentano invece in modo molto evidente le parole che caratterizzano pienamente il discorso inaugurale del secondo mandato: *freedom*, *liberty*, *human* (*human freedom*, *human being*, *human dignity*, *human rights*, *human liberty*, *human choice*), *justice*, *tyranny*; un presidente che si propone alla nazione come il difensore dei principi e dei valori minacciati dalla catastrofe dell'11 settembre 2001. Nella retorica del presidente Bush non c'è distinzione tra "interessi americani" e "valori americani". Si tratta di una semplificazione del problema che si esprime anche nella scelta di articolare il discorso inaugurale intorno al tema della libertà (*freedom*, *liberty*). Valori comuni non esprimono necessariamente interessi comuni come dimostrano bene le scelte politiche divergenti compiute da Francia e Germania rispetto alla guerra in Iraq (Gvosdev e Saunders, 2005).

In Obama, invece, risaltano due aspetti: il senso del "noi" (*us*), che presenta 23 occorrenze su 24 in totale, e il valore storico del "nuovo" (11 su 12 nell'insieme del corpus) che la sua elezione rappresenta sui diversi fronti della politica americana. Dell'utilizzo del pronome personale *us* si è già detto: le concordanze mettono in evidenza il forte richiamo ideale che unisce il presente al passato, il senso di responsabilità collettiva, la consapevolezza dell'importanza storica dell'evento.

Tab. 1.4 – Concordeanze della forma pronominale **us** nel discorso di Obama

I stand here today humbled by the task before **us**,
men and women obscure in their labor -- who have carried **us** up the long
for **us**, they packed up their few worldly possessions
for **us**, they toiled in sweat shops and settled the west;
for **us**, they fought and died,
that feed our commerce and bind **us** together.
that have consumed **us** for so long no longer apply.
and those of **us** who manage the public's dollars
nor is the question before **us** whether the market is a force for good or ill.
but this crisis has reminded **us** that without a watchful eye,
they understood that our power alone cannot protect **us**,
nor does it entitle **us** to do as we please.
you cannot outlast **us**,
as we consider the road that unfolds before **us**,
they have something to tell **us** today,
it is precisely this spirit that must inhabit **us** all.
a friend lose their job which sees **us** through our darkest hours.
what is required of **us** now is a new era of responsibility
the knowledge that god calls on **us** to shape an uncertain destiny.
so let **us** mark this day with remembrance,
let us remember these timeless words.
let **us** brave once more the icy currents,
with eyes fixed on the horizon and god's grace upon **us**,

Anche nell'evocare il senso del "nuovo", Obama imprime al discorso un valore epocale: *new age* (2); *new era* (2); *new life*, *new jobs*, *new foundation*, *new threats*, *new way forward*. In questo Obama coniuga sapientemente le radici del suo passato con le speranze del futuro, come era già nel suo programma politico e sociale illustrato in un libro di grande successo come *Dreams from My Father: A Story of Race and Inheritance* del 1995 (Shima, 2009):

Tab. 1.5 – Concordeanze della forma **new** nel discorso di Obama

prepare the nation for a **new** age.
traveled across oceans in search of a **new** life
we will act -- not only to create **new** jobs,
but to lay a **new** foundation for growth.
colleges and universities to meet the demands of a **new** age.

we can meet those **new** threats that demand even greater effort
that America must play its role in ushering in a **new** era of peace.
to the muslim world, we seek a **newway** forward
our challenges may be **new**.
the instruments with which we meet them may be **new**.
what is required of us now is a **new** era of responsibility

Nel discorso di Barak Obama i valori fondanti della democrazia americana sono diffusi e disseminati senza una evidente concentrazione in poche parole d'ordine, ma con un'attenzione convinta verso ideali concreti sui quali orientare la propria azione politica: *prosperity, freedom, cooperation, market, job, wealth, school*. La rilevanza di questi valori nell'azione politica di Barak Obama era già stata segnalata dagli studiosi che avevano seguito i primi passi del senatore Obama ancor prima della sua candidatura a presidente (Atwater, 2007; Kantor, 2007). Concretezza e pragmatismo che derivano dai suoi studi in legge all'Università di Chicago (Schultz, 2009; Winant, 2009).

Barak Obama non elude il problema della crisi economica (4) in atto e la stessa crisi della politica. Il richiamo alla forza dello spirito e ai valori religiosi si inserisce pienamente nella tradizione della retorica politica americana: *God* (3), *Christians, scripture, destiny*. Nel discorso di Obama vi sono due avverbi illocutivi - *now* (6) e *today* (7) - che accanto alla parola *new* (11) restituiscono appieno il senso di un evento che è già Storia nel momento in cui avviene: *now taken the presidential oath; now stand before you to take a most sacred oath; now, there are some who question the scale of our ambitions; now is a new era of responsibility*. In queste parole si esprime tutto l'impegno del giuramento "qui e ora": *today humbled by the task before us; today I say to you that the challenge we face are real; today we must pick ourselves up; today from the grandest capitals to the small village; today just as the fallen heroes who lie in Arlington whisper*.

1.5. I DISCORSI DI OBAMA E DI BUSH OSSERVATI ATTRAVERSO L'INDICE DI NEGATIVITÀ

L'oratoria di Barak Obama, fortemente caratterizzata dalla consueta retorica nazionale americana, è stata sicuramente emozionante per il pubblico che era presente a Washington e anche per il pubblico in ascolto, sebbene non sia particolarmente segnata da parole di carattere "emotivo". Le più significative sono *courage, faith, hatred, willingness, interest* e *trust*.

Per il confronto dei due discorsi e per l'impatto che essi hanno avuto sul pubblico in ascolto è interessante poter osservare in forma sintetica e compa-

rabile quanto essi si differenziano rispetto alla positività/negatività delle parole utilizzate. Ovviamente questo approccio è troppo semplicistico per poter essere assunto come misura dell'ottimismo/pessimismo che si esprime nei due discorsi ma può essere un buon modo per valutarne il tono complessivo.

Per questo scopo si può fare ricorso a una risorsa già sviluppata, nel corso del tempo, dal *General Inquirer system* con il suo *Inquirer dictionary* (Id) che utilizza le categorie ricavate da quattro fonti diverse: il dizionario di Harvard IV-4; il dizionario di Laswell sviluppato su alcune categorie valoriali; altre categorie di recente costruzione e poi dalle categorie grammaticali.

Nell'Id è disponibile una lista di termini positivi e negativi costituita da 1.915 lemmi contrassegnati come positivi e 2.291 lemmi come negativi. L'applicazione di questa lista, tuttavia, è condizionata alla disponibilità di un corpus lemmatizzato. Pertanto, utilizzando la categorizzazione grammaticale effettuata con TreeTagger (Step 5) è stato necessario ricostruire il corpus in modo che esso fosse costituito dai lemmi anziché dalle forme grafiche originali.

Nel suo primo discorso alla nazione Obama non nasconde le sue preoccupazioni per le sfide che lo attendono. Si è già detto dei riferimenti espliciti alla crisi economica: *That we are in the midst of the crisis is now well understood*. L'analisi del bilanciamento tra le parole che esprimono positività e negatività ne costituisce una conferma (tab. 1.6). L'indice di negatività (il rapporto tra le parole negative e positive) nel discorso di Obama è di 0,52 a fronte di uno 0,33 di Bush nel 2005.

Tab. 1.6 – Indice di negatività dei discorsi inaugurali di B. Obama (2009) e G. W. Bush (2005)

Parlante	Lemmi negativi	Lemmi positivi	Negativi/Positivi
Barak Obama	106	203	0,52
Gorge W. Bush	75	246	0,30

Alcuni studi condotti dagli anni '50 in poi del secolo scorso da Osgood, Souci e Tannenbaum (1957), poi confermati successivamente da altri autori (Bolasco e Della Ratta-Rinaldi, 2004; Giuliano, La Rocca, 2010), hanno messo in evidenza la fondatezza della cosiddetta *Pollyanna hypothesis*, secondo la quale sussiste una tendenza nella comunicazione che vuole che si utilizzino più frequentemente parole a valenza positiva piuttosto che negativa. Secondo questa ipotesi, in seguito ad alcune prove empiriche su diversi tipi di testi, si ritiene che un valore del rapporto tra termini negativi e positivi maggiore di 0,40 esprima un propensione del testo verso

il tono negativo. Su questa base, il discorso di Bush del 2005 si presentava con una netta tendenza alla positività mentre il discorso di Obama del 2009 si presenta con il segno opposto.

1.6. CONCLUSIONE

L'analisi automatica dei discorsi inaugurali dei due presidenti degli Stati Uniti ci ha permesso di mettere in evidenza, in modo quantitativo e sintetico, alcune delle osservazioni già evidenziate dagli studiosi della "retorica presidenziale". Pur consapevoli, sul piano metodologico, della "fragilità" dal punto di vista quantitativo di un corpus che – per la sua dimensione – ben si presta ad altre metodologie qualitative che si basano sulla lettura critica dei testi, si è voluto mettere in evidenza come, attraverso tecniche di statistica linguistica emergano caratteristiche fondate sulla frequenza delle parole, purché le osservazioni siano rigorosamente limitate alle evidenze empiriche ed eventualmente arricchite da valutazioni qualitative "di contesto" come quelle offerte dall'analisi delle concordanze.

I due discorsi si differenziano – ovviamente – per i due momenti storici in cui si collocano, ma anche per i temi che in essi si evidenziano. Tra i temi che erano nelle attese degli analisti (e degli elettori) spiccano i riferimenti all'unità della nazione e al "nuovo" clima che l'elezione di Obama ha rappresentato rispetto all'amministrazione che lo ha preceduto.

L'analisi del "tono" valutativo del discorso di Obama rispetto a quello di Bush ha confermato il senso di responsabilità del nuovo presidente nel mettere in primo piano le difficoltà economiche e politiche del momento, senza lasciarsi andare a retoriche trionfalistiche e a false promesse. Con un'espressione un po' forte, possiamo dire che in questo "pessimismo della ragione" Barak Obama esprime uno *Smart Power*, come lo ha chiamato in un commento a caldo il politologo Joseph Nye, un potere intelligente che sa mediare tra persuasione e coercizione, mostrando il nuovo volto dell'America di oggi.

STRATEGIA DI ANALISI

STEP 1 – ORGANIZZAZIONE E NORMALIZZAZIONE DEL CORPUS

Il corpus in analisi è costituito dai discorsi inaugurali dei due presidenti. I discorsi del giuramento di G.W. Bush e B. Obama sono stati scaricati dal sito The Avalon Project della Yale Law School nel quale sono raccolti tutti gli I-

naugural Adresses dei Presidenti degli Stati Uniti (consultato il 15 gennaio 2010: http://avalon.law.yale.edu/subject_menus/inaug.asp). I testi sono stati acquisiti in TaLTaC2 con identificazione della variabile “parlanti” con due modalità “Obama” e “Bush”.

Il corpus è di piccole dimensioni (4.426 occorrenze) pertanto è stato possibile procedere manualmente alla riduzione delle maiuscole tenendo conto delle forme grafiche che necessariamente devono rimanere con l'iniziale maiuscola (nomi propri, sigle ecc.). In ogni caso questa è un'operazione che non può essere eseguita automaticamente da un programma come TaLTaC2 perché richiede l'utilizzo di risorse linguistiche in inglese. Tuttavia durante l'acquisizione del corpus in TaLTaC il programma esegue, dopo il parsing, un controllo formale sulla punteggiatura e sui caratteri che rappresenta comunque un passaggio utile e fondamentale ai fini dell'analisi.

STEP 2 – GENERAZIONE DEL VOCABOLARIO E MISURE LESSICOMETRICHE

Le operazioni di parsing e generazione del vocabolario del corpus non sono condizionate dalla lingua. Molte operazioni importanti per la statistica linguistica possono essere condotte con tutti i software di analisi automatica dei testi indipendentemente dalla lingua in cui è scritto il testo stesso. Il parsing identifica le forme grafiche, le codifica e procede al loro conteggio. Inoltre possiamo chiedere a TaLTaC2 di eseguire il calcolo delle occorrenze per ciascun parlante dal menu **Analisi** – *Pre-trattamento* – *Calcolo delle sub-occorrenze*.

Questa operazione, oltre a permetterci di apprezzare immediatamente le prime differenze di linguaggio tra i due parlanti, rappresenta un passaggio fondamentale per eseguire tutti i confronti successivi.

STEP 3 – ANALISI DELLE CONCORDANZE

Dal menu **Analisi** – *Analisi testuale* – *Recupero di informazione: concordanze* – *Concordanze semplici e complesse* possiamo osservare le forme grafiche di maggiore interesse all'interno del contesto rappresentato dalle parole adiacenti (che precedono o seguono) una forma grafica assunta come punto di riferimento. L'output delle concordanze è sempre effettuato nel contesto delle parole che precedono e seguono la forma grafica prescelta. Nel primo caso, avendo come riferimento l'aggettivo *our*, l'analisi delle concordanze è stata effettuata solo rispetto alle pa-

role che seguono. Tuttavia, negli approfondimenti successivi riferiti al pronome personale *us* e all'aggettivo *new*, le concordanze sono state effettuate in entrambe le "direzioni".

STEP 4 – ANALISI DELLE FORME SPECIFICHE

Dal menu **Analisi** – *Analisi Lessicale* – *Analisi delle specificità* mettiamo a confronto direttamente le parole utilizzate dai due parlanti con l'obiettivo di individuare quali parole sono presenti, per esempio, nel discorso di Obama in misura maggiore (o minore) di quanto sarebbe nelle nostre attese rispetto al discorso di Bush se i due presidenti utilizzassero il loro linguaggio attingendo allo stesso vocabolario e agli stessi argomenti.

Le differenze nella frequenza delle parole utilizzate nel discorso di un parlante rispetto al complesso delle parole utilizzate nei due discorsi sono evidenziate applicando un test statistico di significatività fondato sulla legge ipergeometrica. Il test richiede di fissare una soglia di occorrenze delle parole da prendere in esame che, di solito, è maggiore o uguale a 10 nel corpus. In questo caso – con un corpus di piccole dimensioni – il test ha un valore puramente indicativo e di sussidio alla descrizione delle differenze tra i parlanti nei limiti dei due discorsi analizzati ma non può essere assunto come rappresentativo di una vera e propria specificità lessicale.

STEP 5 – LEMMATIZZAZIONE CON TREE'TAGGER

TreeTagger è un software sviluppato nell'Institute for Computational Linguistics dell'Università di Stoccarda che permette di classificare le parole di un testo scritto in diverse lingue (francese, inglese, tedesco, italiano, spagnolo bulgaro, russo, greco e portoghese) attribuendo ciascuna parola a una categoria grammaticale e a un lemma ritenuto adeguato secondo determinati modelli di linguistica computazionale dotati di specifici parametri. Sul sito di riferimento è possibile ottenere la documentazione scientifica, le istruzioni e i file di installazione: <http://www.ims.uni-stuttgart.de/projekte/corplex/TreeTagger/>.

TreeTagger è in grado di elaborare, per esempio, un corpus in input avente il seguente formato per TaLTaC2:

```
****ObamaJan20-2009 *P=Obama  
Endrowvar
```

my fellow citizens: I stand here today humbled by the task
before us, grateful for the trust you have bestowed, mindful
of the sacrifices borne by our ancestors.

L'output di TreeTagger ha il seguente formato:

```
****ObamaJan20-2009  JJ  ****ObamaJan20-2009
*P=Obama             NN  *P=Obama
endrowvar            NN  endrowvar
my                    PP$ my
fellow                JJ  fellow
citizens              NNS citizen
:                     :   :
I                     NP  I
Stand                 VVP stand
here                  RB  here
today                 NN  today
humbled               VVN humble
by                    IN  by
the                   DT  the
task                  NN  task
before                IN  before
us                    PP  us

'                     '   '
grateful              JJ  grateful
for                   IN  for
the                   DT  the
trust                 NN  trust
you                   PP  you
have                  VHP have
bestowed              VVN bestow

'                     '   '
mindful              JJ  mindful
of                   IN  of
the                   DT  the
sacrifices           NNS sacrifice
borne                VVN bear
by                    IN  by
our                   PP$ our
ancestors             NNS ancestor
.                     SENT .
```

Questo file di output può essere acquisito in TaLTaC2 con integrazione nel vocabolario del corpus delle informazioni grammaticali corrispondenti.

STEP 6 – RICOSTRUZIONE DEL CORPUS LEMMATIZZATO E APPLICAZIONE DEL TAGGING SEMANTICO CON LISTA DA FILE ESTERNO

Il corpus, già acquisito in TaLTaC a seguito del passaggio in TreeTagger, dal menu **File** – *Esporta* – *Ricostruzione corpus* – *Testo annotato con ...* dalla finestra di comando “Tipo di ricostruzione” si spuntano le opzioni “Lemma” e formato “Testo libero”. L’esecuzione del comando produce un file di testo (da salvare nella cartella di lavoro) che in sostituzione delle forme grafiche originali (forma flesse; ad esempio: *citizens*) contiene il testo lemmatizzato (ad esempio: <citizen>):

```
***ObamaJan20-2009 *P=Obama
my fellow citizen: I stand here today humble by the task before us, grateful for the trust you have bestow, mindful of the sacrifice bear by our ancestor. I thank president Bush for his service to our nation, aswell as the generosity and cooperation he have show throughout this transition. forty-four American have now take the presidential oath. theword have be speak during rise tide of prosperity and the stillwater of peace. yet, every so often, the oath be take amidst gather cloud and rage storm. (...)
```

Il corpus così lemmatizzato è acquisito nuovamente in TaLTaC2 con parsing e generazione del vocabolario (N = 4.391; V = 1.163). Dal menu **Analisi** – *Analisi lessicale* selezioniamo il comando *Tagging* – *Semantico* – *Vocabolario*. Nella finestra di dialogo selezioniamo “Lista da file esterno” e carichiamo la lista da utilizzare che avremo preparato in precedenza: una lista di termini positivi con un solo elemento su ogni riga. Nel campo “Categoria da attribuire a CAT_SEM” scriveremo “POS”. Mandando in esecuzione il comando ciascun lemma del corpus corrispondente alla lista dei termini positivi sarà annotato nella colonna CAT_SEM della tabella Vocabolario della sessione con l’etichetta POS. La stessa operazione sarà eseguita con la lista dei lemmi negativi e con l’annotazione dell’etichetta NEG nella tabella Vocabolario.

Il passo successivo consiste nel conteggiare le occorrenze dei lemmi positivi e negativi. Nella tabella selezionare la colonna “CAT_SEM” e, dal menu **Calcola**, il comando *Fusioni*. Verrà visualizzata una tabella riassuntiva in cui le categorie semantiche presenti nella colonna selezionata sono associate con le occorrenze dei lemmi positivi e negativi per ciascun parlante e sulla quale si potrà calcolare il rapporto lemmi negativi/lemmi positivi.

RIFERIMENTI BIBLIOGRAFICI

- ATWATER D. F. (2007) "Senator Barack Obama", in *Journal of Black Studies*, vol. 38, n. 2, pp. 121-129.
- BERSON I. R, BERSON, M. J. (2009) "Making Sense of Social Studies with Visualization Tools", in *Social Education*, vol. 73, n. 3, pp.124-126.
- BOLASCO S., DELLA RATTA-RINALDI F. (2004) "Experiments on Semantic Categorisation of Texts: Analysis of Positive and Negative Dimension", in G. Purnelle, C. Fairon, A. Dister (eds) *Les poids des mots. Actes des 7^e JADT*, Louvain, Presse Universitaire de Louvain, pp. 202-210.
- DILLON G. L., DOYLE A., EASTMAN C. M., KLINE S., SILBERSTEIN S., TOOLAN M. (1990) "The Rhetorical Construction of a President", in *Discourse & Society*, vol. 1, n. 2, pp.189-200.
- EIBACH R.P., PURDIE-VAUGHNS V. (2009) "Change We Can Believe in?: Barack Obama's Framing Strategies for Bridging Racial Divisions", in *Du Bois Review: Social Science Research on Race*, vol. 6, n. 1, pp. 37-151.
- GIULIANO L., LA ROCCA G. (2010) "Validity and reliability of the automatic classification of texts according to the negative-positive criterion", in S. Bolasco, I. Chiriari, L. Giuliano, *Statistical Analysis of Textual Data*. Sapienza University of Rome, *Actes des 9^{èmes} JADT*, Rome 9-11 June, Milano: LED, vol. 1, p. 61-72.
- GVOSDEV N, K., SAUNDERS P. J. (2005) "On Liberty", in *The National Interest*, n. 79, pp. 5-7.
- HARLOW R. (2009) "Barack Obama and the (In)Significance of His Presidential Campaign", in *Journal of African American Studies*, vol. 13, n. 2, pp. 164-175.
- HOUCK D. W., NOCASIAN M. (2002) "FDR's First Inaugural Address: Text, Context, and Reception", in *Rhetoric & Public Affairs*, vol. 5, n. 4, pp. 649-678.
- KANTOR J. (2007) "In law school, Obama found political voice", *New York Times*, January 28, pp. 1, 21.
- NORTHCOTT M. (2004) "An Angel Directs the Storm': The Religious Politics of American Neoconservatism", in *Political Theology*, vol. 5, n. 2, pp. 137-158.
- REED W. L, LOUIS B. M. JR (2009) "'No More Excuses': Problematic Responses to Barack Obama's Election", in *Journal of African American Studies*, vol. 13, n. 2, pp. 97-109.
- SCHULTZ B. (2009) "Obama's Political Philosophy: Pragmatism, Politics, and the University of Chicago", in *Philosophy of the Social Sciences*, vol. 39, n. 2, pp. 127-173.
- SHIMA A. (2009) "The Fictive Force of Barack Obama's Dreams from My Father", in *American Studies in Scandinavia*, vol. 41, no. 1, pp. 3-22, fall.
- WINANT H. (2009) "Just Do It: Notes on Politics and Race at the Dawn of the Obama Presidency", in *Du Bois Review: Social Science Research on Race*, vol. 6, n. 1, pp. 49-70.
- WINTER D.G. WEINTRAUB W. (2005) "William Jefferson Clinton: Personality Traits and Motivational Biases: Motivations and Mediation of Self-Other Relationships", in Jerrold M. Post, *The psychological assessment of political leaders: With profiles of Saddam Hussein and Bill Clinton*, Ann Arbor, MI: The University of Michigan Press, pp. 303-312.

2.

CONSUMI E IMMAGINI DAI *TWEENS*. I FANS DI HANNAH MONTANA

2.1. INTRODUZIONE

Fino agli anni Sessanta il termine *tween* non esisteva nemmeno. In uno dei primi testi dedicati a questi piccoli consumatori Eugene Gilbert (1957) utilizzava la categoria e il termine di *subteen*, quindi preadolescenti, compresi in una fascia d'età tra i dieci e i tredici anni. Per incontrare il termine *tween* bisogna aspettare il 1987: esso appare in un articolo di Carol Hall sulla rivista *Marketing and Media Decision*. Infatti, sebbene l'infanzia come settore di vendita sia apparsa negli anni immediatamente successivi alla seconda Guerra Mondiale, è solo intorno agli anni Novanta che i ragazzi tra i nove e i quindici anni diventano un target di consumo. Nello specifico con il termine *tweens* si identificano i ragazzi tra i nove e i dodici anni. Thomas Cook e Susan Kaiser (2004) in *Between and be Tween* dicono che *they're younger than teens, older than tots and know as "tweens"*; occorre però precisare che quella di *tweens* è una definizione inventata dagli esperti del settore, non è pertanto una categoria o etichetta in cui i ragazzini si riconoscono.

La figura dei *tweens* non può essere compresa se non si considera il mercato dell'infanzia e l'indotto che vi gira intorno. A quell'età l'appartenenza è segnata dal consumo – inteso secondo la lezione di Thorstein Veblen (1971; ed. or. 1899) – che omologa o differenzia: tra gli obiettivi di questi ragazzi c'è il raggiungimento della popolarità, la voglia di essere alla moda, sono quindi facile presa di pubblicità accattivanti, di generi di consumo creati su misura, del desiderio di emulare i personaggi televisivi. “Il giovanissimo consumatore, facilmente manipolabile, è a sua volta maggiormente in grado di manipolare un genitore pieno di sensi di colpa (Quart, 2003, p. 83)”. Un esempio di come si possa gene-

rare un fenomeno di consumo per questi ragazzi è il caso Hannah Montana, che da serie tv è fuoriuscita dai teleschermi e invade librerie, negozi di musica, cinema. Per capire come i *tweens* si aggregino intorno a un personaggio televisivo e creino una community intorno a esso si rivela utile l'analisi dei messaggi, delle credenze, dei pensieri frutto della loro interazione e lasciati su di un forum.

I messaggi selezionati a tale scopo sono stati scaricati dal forum ufficiale di Hannah Montana; fra i vari topic presenti si è scelto di lavorare con quelli dedicati specificatamente ad Hannah e a Miley Cyrus. Come tutti gli appassionati sapranno le due protagoniste sono la stessa persona: Miley, una ragazzina “normale”, quando canta diventa Hannah Montana; per farlo indossa una parrucca per proteggere la sua identità e soprattutto la sua comune e tranquilla vita da adolescente.

Miley Cyrus (il vero nome dell'interprete) è nata il 23 Novembre 1992 a Nashville, Franklin, nello stato del Tennessee. Proviene da una famiglia di artisti: il padre, Billy Ray Cyrus, è un cantante country mentre la madre è un'attrice. Ha ottenuto all'età di undici anni la parte di Miley Stewart/Hannah Montana per i tipi della Disney, che durante il provino sono rimasti molto colpiti dall'energia di questa ragazzina. Così, grazie alla serie tv Hannah Montana, Miley, a soli 15 anni, è diventata una delle *teenagers* attrici-cantanti più famose al mondo. L'impatto che questa dirompente ragazzina ha sulla vita dei pre-adolescenti è notevole: per quantificare l'ingerenza nella vita dei *tweens* basta guardare le vendite dei suoi cd, dei libri e gli incassi dei film a lei dedicati.

Il forum si presta bene ad analizzare l'immagine che di Hannah e Miley si formano i loro fans. Infatti mediante la lettura, la codifica e poi l'interpretazione di quanto emerge è possibile sondare la costruzione dell'immagine della/delle protagoniste della serie tv identificando i tratti che maggiormente rimangono impressi nei piccoli telespettatori.

Obiettivo del presente lavoro è quindi:

- ricostruire l'idea che di Hannah e di Miley hanno questi ragazzi, verificandone le differenze;
- far emergere i tratti dei *tweens* come target di vendita;
- tratteggiare il linguaggio utilizzato dai *tweens*.

2.2. I TEMI DELLE DISCUSSIONI ONLINE

Mediante la lettura, il corpus è stato frammentato in 244 segmenti ai quali sono stati applicati 256 codici. È il topic dedicato ad Hannah Montana ad aver fatto registrare la necessità di creare più codici; da questo corpus sono infatti emersi 18 codici in più rispetto a quello dedicato a Miley (fig. 2.1).

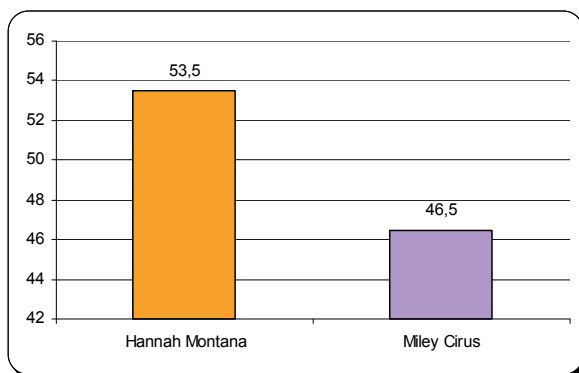


Fig. 2.1 – Distribuzioni percentuali dei codici nei corpora

L'immagine e l'immaginario che accompagna un personaggio televisivo nelle fantasie degli adolescenti è frutto di diversi aspetti. La serie oltre a ritrarre scenari possibili è essa stessa uno strumento di socializzazione: degli episodi e di quanto vi gira intorno si parla con gli amici, addirittura se ne conversa on line. Come fa notare Dafna Lemish (2007) "i bambini non sono telespettatori passivi (...). Le ricerche confermano (...) che si coinvolgono con il mezzo televisivo e il suo contenuto in vari modi attivi, tra cui gestire la propria attenzione, trarre significato dai messaggi, analizzare e criticare, e ricordare selettivamente (p. 41)". Anche i bambini e i ragazzi effettuano una loro lettura del testo televisivo, riuscendo a trarre significato attraverso quelle competenze e quegli strumenti che acquisiscono nel tempo (Valkenbrug, 2004; Van Evra, 2004). Di conseguenza, con la crescita psicologica, emotiva e fisica muta pure la loro modalità di interagire con la televisione. Questo processo porta alla considerazione che "il "significato" del contenuto televisivo non risiede né nel programma specifico, né in una creazione indipendente che si forma nella testa del bambino, ma si produce nell'interazione tra il bambino e il programma". In questo senso il forum è uno strumento utile a cogliere il significato attribuito alla serie televisiva, ad esplicitare i frutti dell'interazione tra il programma e lo spettatore, a organizzare la costruzione di un significato condiviso attraverso lo scambio di messaggi.

2.3. MILEY/HANNAH E HANNAH/MILEY: ELEMENTI DI IDENTIFICAZIONE

L'elemento di questa serie che più appassiona i suoi telespettatori è la doppia vita di Miley che può essere a un tempo una ragazzina comune, una di loro e contemporaneamente una diva, una star, un'eroina, un modello da imitare, qualcuno verso cui tendere.

I personaggi televisivi forniscono ai ragazzi dei modelli di ruolo con cui non solo identificarsi ma anche da poter imitare. Questo assunto sta alla base dei programmi educativi proposti dalla Media Education. Tant'è che già negli anni Quaranta quando padre Félix Morlion introduceva il cineforum in Italia (Rivoltella, 2001) il presupposto era proprio quello di fornire una realtà nuova, lontana dagli schemi quotidiani dei ragazzi che intorno a esso riusciva a radunare. Il processo di identificazione con il personaggio televisivo può essere indotto da molteplici aspetti, per esempio ci si può identificare con delle caratteristiche ascritte del protagonista: genere, età, razza; oppure con quelle acquisite: dolcezza dei modi, fragilità, posizione lavorativa, ecc.

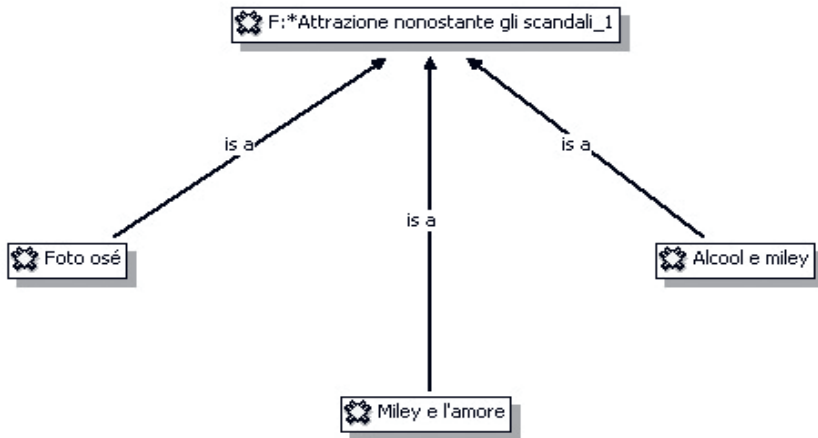


Fig. 2.2 – Elementi costituenti la categoria “attrazione nonostante gli scandali”

Nei ragazzi, quando ancora il processo di crescita non è completamente compiuto, può accadere che si innesti una forma di “identificazione ideale” (Lemish, 2007) per effetto della quale si cerca di agire, di comportarsi come il personaggio della serie, non come l'essere umano, attore che vive nella vita reale.

Questo processo ha i suoi pericoli perché come tutte le forme di idealizzazione non riesce a garantire né una percezione oggettiva dei fatti, né della realtà.

Nei nostri fans quello che si produce è una difesa di Miley nonostante i gossip di cui apprendono la notizia. Lo spazio semantico che si costruisce mettendo in relazione i codici non comuni a entrambe le famiglie “attrazione per il personaggio” e “narrazione episodi scandalistici” ci restituisce un’informazione sul permanere dell’attrazione per la star nonostante gli episodi che la vedono protagonista o di abuso eccessivo di alcool o come pin-up che posa per foto osé.

L’ammirazione per questo personaggio permane, nonostante si legga di lei che abbia ecceduto nel bere o si sia fatta ritrarre in pose poco appropriate, perché forte è il processo non solo di identificazione ma anche di ammirazione. Soprattutto per le ragazzine per le quali è quasi d’obbligo selezionare fra i vari modelli femminili proposti dalla tv quelli che spiccano anche per il loro aspetto fisico; infatti le bambine di oggi continuano a interiorizzare modelli femminili per i quali viene esaltata la bellezza esteriore.

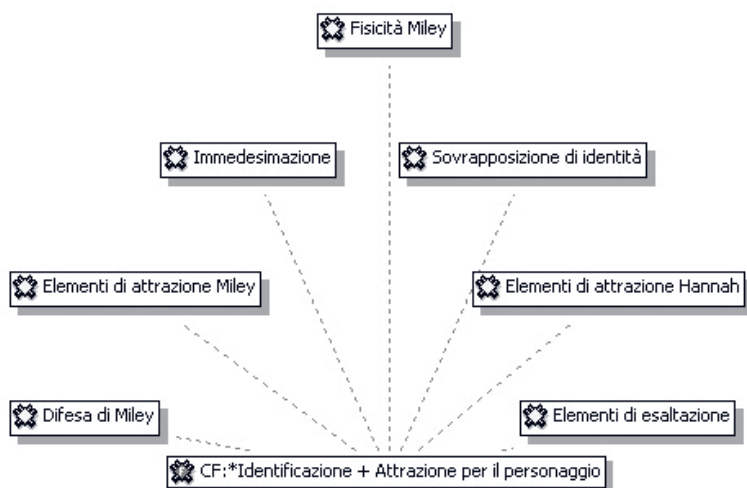


Fig. 2.3 – Mappa degli elementi di attrazione

Nel delineare l’insieme degli elementi per i quali i *tweens* impazziscono per la loro star si vede come questa passione sia frutto di un’intersezione di aspetti: fisici, caratteriali, di processi di immedesimazione nella vita che questa ragazzina conduce tanto sullo schermo che fuori dalla serie. La difesa è veramente appassionata ed esula dall’oggettivazione di quanto questa adolescente realmente vive.

2.4. I TWEENS COME TARGET DI VENDITA

Il mercato dei teenagers scoperto essenzialmente durante la seconda Guerra Mondiale ha avuto il suo decollo quando gli esperti hanno capito che era meglio “prenderli sin da piccoli” (Quart, 2003).

Danielle Hullan (2007) si interroga anche su come questi ragazzi riescano a trovare i soldi da poter spendere per l'acquisto di questi beni. Le possibili risposte vedono ora come grandi elargitori di paghetten sostanziose i genitori single che in questo modo attutiscono i sensi di colpa; ma soprattutto i *tweens* americani trovano i soldi facendo dei lavoretti domestici o svolgendo l'attività di baby sitter. Lavorare per consumare, per acquistare beni porta a definire questi ragazzi come i “nuovi proletari”, ai quali nei negozi, nei centri commerciali sono dedicati spazi, stand appositi. Inoltre, l'esposizione alla tv e ai nuovi media induce in loro il bisogno del prodotto. Un esempio è appunto la serie tv che traina la vendita delle riviste, dei libri, dei gadget. Parallelamente si cerca anche di instillare una forma di consumo in maniera più subdola: si induce il consumo prima virtuale di un bene e poi reale. Per esempio, nelle partnerships fra imprese si vede il profilarsi di questo nuovo business. Se giocando online i ragazzi per passare ai livelli successivi o per vincere devono mangiare o bere McDonald, inevitabilmente finiranno per introiettare questo comportamento e sarà per loro naturale trasferirlo nella vita di tutti i giorni.

Nel suo *Branded* Alissa Quart racconta dell'ossessione per il marchio dei giovani teenagers americani e di come questi siano diventati un target di vendita altamente redditizio per il settore dell'editoria, così la cosiddetta stampa per adulti ha creato delle riviste anche per loro “a Teen People hanno fatto seguito altri ‘derivati’ delle riviste per adulti dedicate ai teenagers, come *Elle Girl*, *Cosmo Girl!* e *Teen Vogue* (p. 27)”. Queste riviste non sono altro che il primo gradino di un processo di socializzazione allo stars system, utile per iniziare gli adolescenti al culto delle celebrità.

Più che il contenuto delle riviste fanno presa i gadget che a esse sono acclusi, ovvero le foto, i poster, le cartoline: “due belle riviste... con due mega poster” si legge in uno dei post o con “4 maxi poster” di Miley o di altri personaggi della serie.

L'aspetto legato alle vendite delle riviste e dei gadget di Hannah Montana sono di così forte impatto che sono ben 74 su 244 i segmenti di testo contenuti in entrambi i topic nei quali si parla o di riviste o di gadget connessi appunto alla serie tv (cfr. fig. 2.4). I ragazzi si scambiano informazioni sui contenuti ma anche sui costi, dove acquistare le riviste e sulle esperienze mediatiche a esse connesse. Tanti sono i messaggi di scambio di informazioni sui giochi che Disney Channel affianca alla serie tv. Il forum è proprio un'arena in cui questi piccoli fans pos-

sono strutturare relazioni di significato intorno alla fruizione dei vari episodi.

Nodes (74): QU:1.104 Ho preso Kiss n43 a 1,50 come .. (293.293), QU:1.105 oggi ho preso questa rivista .. (299.299), QU:1.102 ho comperato POPSTAR N.18 è il .. (283.283), QU:1.103 ho comperato teen 45 è il nume .. (289.289), QU:1.106 ci sono 80 pagine e 17tra post. (299.299), QU:1.109 ho comprato kiss 41, c'è un me .. (330.330), QU:1.110 Ciao raga, ho comprato Let's g. (332.332), QU:1.107 ho comprato questa rivista con .. (319.320), QU:1.108 guardate questa rivista, si ch .. (324.324), QU:1.95 ho preso queste due riviste di .. (265.265), QU:1.96 tutte e due hanno come regalo .. (265.265), QU:1.92 Ho scoperto che sta per uscire .. (247.247), QU:1.93 ho comprato questa rivista, de .. (263.263), QU:1.97 se vi interessa su Teen 46 di .. (270.270), QU:1.100 Ciao ho preso queste due rivis .. (281.281), QU:1.101 con dei fantastici gadget, in .. (281.281), QU:1.98 Ciao a tutti in edicola mi han .. (272.272), QU:1.99 ho comperato LETS GO N.32 è i .. (277.277), QU:1.111 Ciao è in edicola young pocket. (334.334), QU:2.7 Ho appena saputo che Miley sar .. (46.46), QU:2.9 La nostra adorata Miley per un .. (50.50), QU:1.126 Vorrei sapere le risposte alle .. (745.745), QU:2.5 ma nn riescono capire che senz. (40.41), QU:2.12 ma si trova su internet???per .. (63.63), QU:2.52 Un décolleté troppo in vista .. (284.284), QU:2.91 la Giovane Stella Miley Cyrus .. (533.533), QU:2.33 I pettegolezzi su una possib .. (142.142), QU:2.49 Il 19 Luglio Disney Channel US .. (275.275), QU:1.114 volevo chiedervi se hanno già .. (351.351), QU:1.115 "Quest'anno mio papà è morto n .. (373.373), QU:1.112 ho comprato pink girl 38 (343.343), QU:1.113 c'è un mega poster dei jonas e .. (343.343), QU:1.118 Ho scritto a Disney Channel e .. (489.489), QU:1.123 A ME SERVIREBBE UN AIUTO PERCH .. (646.646), QU:1.125 mi potete dire l'indirizzo di .. (730.730), QU:1.121 Per il giochino, Disney channe .. (539.539), QU:1.122 IL GIOCO "MESSAGGI PER TUTTI" .. (592.592), QU:1.44 ci sono 4 enormi poster, 2 dei .. (77.77), QU:1.45 in uno ci sno anche gli anelli .. (77.77), QU:1.42 ci sono un sacco di adesivi in .. (75.75), QU:1.43 ho comprato queste due riviste .. (77.77), QU:1.46 ho trovato in vendita questo c .. (81.81), QU:1.49 e tanti altri poster più picco .. (83.83), QU:1.50 ho comprato ieri queste due ri .. (84.84), QU:1.47 è uscito pink girl numero 44 (83.83), QU:1.48 ha due mega poster (83.83), QU:1.3 All'interno Hannah Montana e t. (4.4), QU:1.4 una è kiss 47 (8.8), QU:1.1 Ecco la copertina di Pretty Gi .. (3.3), QU:1.2 Due belle riviste (7.7), QU:1.5 con due mega poster di Patty e .. (8.8), QU:1.40 ho trovato in edicola queste d .. (75.75), QU:1.41 ci sono tanti poster e 4 maxi .. (75.75), QU:1.6 l'altra è popstar 22 (8.8), QU:1.7 tanti super poster (8.8), QU:1.51 ci sono tanti poster e 4 maxi .. (84.84), QU:1.70 Nella domanda riguardo la voce .. (171.171), QU:1.71 Ciao ragazze volevo sapere una .. (173.174), QU:1.66 A che ora e che canale è che n .. (143.143), QU:1.69 si .. in 1 puntata della prima .. (159.159), QU:1.87 ho comprato queste due rivista .. (237.237), QU:1.90 ho comprato questa rivista (243.243), QU:1.91 ci sono 2 maxi poster, 1 dei j .. (243.243), QU:1.88 ci sono 4 maxi poster ed un ga .. (237.237), QU:1.89 su Sky ci sarà il film di Hann .. (239.239), QU:1.54 se volete postate qui tutte le .. (111.111), QU:1.55 sia questa HMMania. E' un conc .. (113.113), QU:1.52 Ciao ragazze volevo sapere una .. (86.86), QU:1.53 ho risposto alla maggior parte .. (102.103), QU:1.56 si e se si vince si va in call .. (115.115), QU:1.61 Volevo partecipare però chiede .. (126.126), QU:1.64 Comunque ogni domanda si rifer .. (136.136), QU:1.59 bisogna aspettare domani e ved .. (120.120), QU:1.60 il gioco su disneychannel (124.124)

Fig. 2.4 – Citazioni estratte nel corpus mediante la query “aspetti commerciali” per i due topic (editor HTML)

2.5. LO SLANG

I partecipanti al forum si esprimono utilizzando più “formule matematiche” che vere e proprie parole. È noto che il linguaggio usato dagli utenti in ambienti virtuali, benché principalmente di notazione scritta, non presenti esclusivamente i soli tratti di tale tipo di scrittura (Passerini, Capussotti, Braunstein, 1996; Baron, 1998; Jakobs, 1998; Pistoleri, 1998).

La massima riduzione della distanza reale nello spazio della Rete, la velocità della comunicazione e il carattere dei messaggi, avvicinano la comunicazione telematica più a condizioni della comunicazione orale che a quelle proprie del testo scritto tradizionalmente (Lavinio, 1990). Non si dimentichi che in questo caso si ha a che fare con un linguaggio prettamente giovanile nello stile; molti di questi post, per esempio, hanno più il carattere dei messaggi che si inviano con il cellulare, dove per comodità si abbreviano le parole che quello della tradizionale lettera.

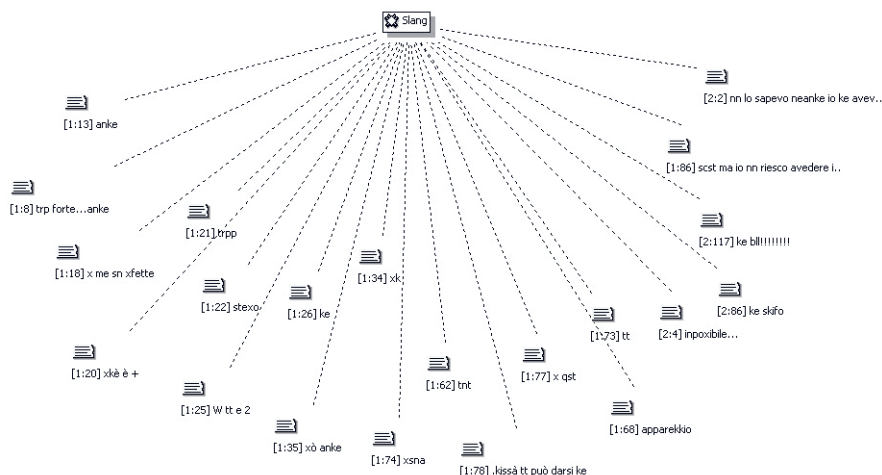


Fig. 2.5 – Network per le citazioni codificate in “slang”

L’abbreviazione più comune è l’uso del segno della *x* al posto del “per” o della doppia “esse”: “x qst”, “x me sn xfette”, “xkè è +”, “stexo”, “xk”, “xò anke”, “inpossible...”, “xsna”; allo stesso modo la *k* sostituisce il suono duro formato dall’unione della “c” con la mutolina: “kissà tt può darsi ke...”, “apparekkio”, “ke skifo”, “ke bll!!!!!!!”, “anke”, “ke”. Un altro elemento di cui si fa economia in rete sono le vocali, così “tutto” diventa “tt”, oppure l’eliminazione della “o” in “non” produce: “nn lo sapevo neanche io ke”. In questa frase, per esempio, scompaiono più vocali: “scst ma io nn riesco a vedere”; e ancora: “tnt”, “trpp”, “trp forte...anke”; finché la mancanza di vocali e l’indicazione della parola “due” con il suo numero finisce per produrre quasi una formula: “tt e 2”.

Lo schermo, e quindi il processo di disinibizione che esso permette, produce una desacralizzazione della scrittura e del testo, svanisce quel timore reverenziale proprio dei compiti in classe e associato al foglio di carta e alla penna. Come sostiene Antonelli (2007) “se il testo diventa labile, la scrittura passa nella sfera dell’effimero: scripta volant; se si scrive così spesso, scrivere diventa un gesto quotidiano, lontanissimo da quell’ufficialità e solennità di cui si era sempre ammantato” (p. 11).

Il forum è uno spazio privato per i *tweens* in cui scrivere è esso stesso un atto di costruzione della propria immagine e di comunicazione del proprio stile, pensare a una forma di italiano digitalizzato equivalente a quello che devono utilizzare a scuola è impossibile, anzi spesso ciò che avviene è l’inverso:

l'abitudine a una scrittura personalizzata fa perdere la solennità all'atto che diventa slang anche durante i compiti in classe.

2.6. DELINEANDO GLI SCENARI

L'analisi di questo particolare segmento di *tweens* fans di Hannah e Miley - che possiamo qui considerare come un'erma bifronte - ci rivela innanzitutto la specifica cognizione che questo segmento di pubblico ha delle eroine, ovvero è in grado di distinguere tra la Miley Cyrus della vita reale, la Miley Stewart protagonista della serie e il personaggio che quest'ultima interpreta quando indossa una parrucca: Hannah Montana.

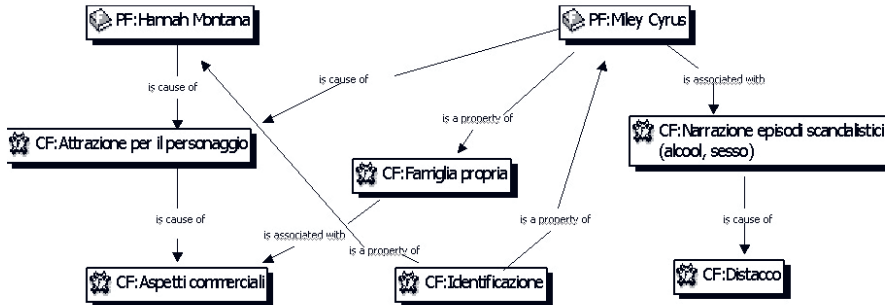


Fig. 2.6 – La rete degli elementi emersi

Questi ragazzini applicano una loro lettura al testo televisivo e sono in grado di scindere i vari elementi di cui si compone. A riprova di quanto detto c'è la specificità di alcune categorie concettuali emerse in uno dei due topic piuttosto che nell'altro. È chiaro che a incappare in episodi scandalistici è Miley Cyrus, la quindicenne che interpreta la serie tv. I commenti alle foto osé o agli episodi di abuso di alcool sono infatti legati al topic a lei dedicato e non se ne trova traccia in quello di Hannah Montana. La conseguenza è che molti di loro esprimono distacco per l'attrice protagonista e non per il personaggio interpretato nella serie. Allo stesso modo i ragazzi parlano della loro famiglia paragonandola ai genitori di Miley Cyrus, per sottolineare la condizione di privilegio di cui quest'ultima gode rispetto a loro.

I genitori di questi *tweens* appaiono anche come possibili finanziatori per l'acquisto delle riviste ponendo in questo modo in relazione i due topic. Ai genitori si chiedono i soldi per comprare le riviste che contengono poster o cartoline dell'attrice del cuore. È quindi l'attrazione e ammirazione per quest'ultima che spinge al consumo.

STRATEGIA DI ANALISI

STEP 1 – ORGANIZZARE E CODIFICARE IL TESTO

Dopo aver proceduto manualmente a scaricare i messaggi postati sul forum è necessario:

- 1) organizzare il testo per variabili/soggetti,
- 2) procedere con la codifica.

Operazione 1

Tutti i messaggi dedicati ad Hannah sono stati raccolti in un unico corpus. Ugualmente quelli dedicati a Miley hanno formato anch'essi un corpus a sé stante.

Per far riconoscere al software che si tratta di due documenti distinti è necessario, dopo averli caricati in Atlas.ti, assegnare loro le variabili nominali, secondo la procedura consentita dal **Primary Doc Family Manager**. Si avranno così due documenti:

- P1: forum Hannah
- P2: forum Miley.

Operazione 2

Caricati i testi è ora possibile procedere all'operazione di codifica. In questo caso si è scelto di operare una codifica dal basso, ovvero i codici non sono già stati individuati a priori bensì emergeranno man mano che si procederà alla lettura del testo.

Aperto il primo forum, si procede leggendo riga per riga il testo e selezionando quei segmenti ritenuti rilevanti; a essi si applica una etichetta: il codice che ha la funzione di chiudere, sintetizzare il contenuto del segmento stesso.

La creazione dei codici è un'operazione puramente concettuale, legata alla sensibilità dell'analista, poiché è lui stesso, in prima persona, che legge ed e-

strae dal testo quanto ritiene rilevante. In questa operazione il software non si può sostituire al ricercatore.

Dalla stessa porzione di testo possono emergere più codici. Per esempio, nel primo documento relativo ai post per il topic Hannah Montana si legge:

*“Ciao raga... volevo chiedervi se hanno già fatto la videodomanda numero 32 di Hannah Montana Mania su Disney Channel... nel caso... potete dirmi la risposta... perchè me la sono persa...!
grazie...!!”*

Questo segmento (*quotation*) ci suggerisce la creazione e applicazione di tre diversi codici:

- richieste di informazioni sulla serie,
- concorsi su Hannah,
- Disney Channel.

Ovvero, all'interno di questo messaggio si affrontano tre argomenti diversi che possono essere chiusi nei codici testé menzionati.

L'operazione delineata può essere resa anche graficamente.

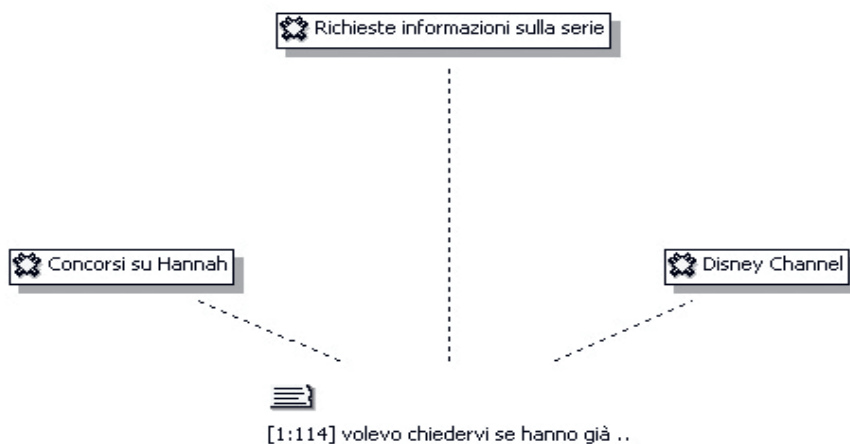


Fig. 2.7 – Network per il frammento e i suoi codici

Per farlo basta richiamare il frammento e i codici a esso applicati dalla barra degli strumenti **Quotations/Quotations Manager** selezionando il frammento in oggetto e poi **Open Network View**.

Questo network può anche essere riportato all'interno del testo per sottolineare la relazione tra la serie tv e gli aspetti di vendita e di promozione a essa legati.

STEP 2 – RAGGRUPPARE I CODICI

Per poter delineare gli argomenti discussi nei due topic è necessario ritornare a lavorare sui codici. Riportare direttamente all'interno del rapporto di ricerca o di un articolo la tabella riepilogativa dei codici (cfr. tab. 2.1) sarebbe infatti dispendioso e poco utile e ciò essenzialmente per due motivi: il primo estetico e il secondo tecnico-operativo. Si tratta, infatti, di materiali che di solito si possono inserire come allegati ma che posti all'interno del documento ne appesantiscono la lettura. Inoltre, può accadere che durante la fase di codifica si creino due codici simili, oppure due codici con il medesimo significante per i quali è opportuno operare una fusione.

Allora occorre:

- 1) verificare la “qualità” dei codici creati,
- 2) riorganizzarli in macro categorie.

Operazione 1

Mediante **Codes / Output / Codes-Primary-Documents-Table / Standard Report** si recupera l'informazione inerente a quanti codici creati e alla loro distribuzione all'interno dei due testi oggetto di analisi.

Una prima riflessione sulla tabella permette di notare che ci sono dei codici specifici per i due topic che si trovano all'interno di uno solo di essi, quale per esempio, “Elementi di attrazione Hannah” o altri che ricadono in entrambi; per esempio: “Elementi di attrazione Miley” o “Gadget”.

Inoltre, ci si accorge che vi sono codici simili per i quali può essere realizzata una fusione: “Non difesa di Miley e l'amore” e il più generale “Non difesa di Miley”; questi due possono essere uniti insieme mediante la procedura di *merge*.

Tab. 2.1 – Elenco dei codici per i topic

	Hannah Montana	Miley Cyrus	Totale
Alcool e Miley	-	12	12
Concorsi su Hannah	9	-	9
Difesa di Miley	-	12	12
Difesa di Miley e l'amore	-	3	3
Disincanto	-	4	4
Disney Channel	7	3	10
Elementi di attrazione Miley	11	4	15
Elementi di esaltazione	12	13	25
Elementi di attrazione Hannah	8	-	8
Fisicità Miley	-	23	23
Foto osé	-	8	8
Gadget	25	2	27
Genitori	-	4	4
Immedesimazione	-	11	11
Info eventi	-	2	2
Miley e l'amore	-	2	2
Musica	-	1	1
Non difesa di Miley e l'amore	-	4	4
Non difesa di Miley	-	6	6
Richieste informazioni sulla serie tv	10	-	10
Riviste	26	-	26
Scambio di informaz.	9	1	10
Espressioni "slang"	17	4	21
Sovrapposizione di identità	3	-	3
Totale	137	119	256

Operazione 2

Continuando con le procedure di analisi si lavora sul piano concettuale creando delle categorie più ampie o famiglie di codici all'interno delle quali inserire gli elementi fino a ora trovati.

Questa operazione di sussunzione è anch'essa un'operazione concettuale e manuale che il ricercatore svolge in prima persona. Infatti, si riaggregano i codici in contenitori di significato più ampi che poi verranno posti in relazione. Da **Codes/Edit Families/Open Family Manger** si realizza tecnicamente questa operazione di imputazione. La fase operativa è però preceduta da un momento di riflessione in cui si individuano i contorni delle macro categorie e si definiscono i codici da introdurre.

Qui abbiamo sei categorie:

- aspetti commerciali, all'interno della quale andranno: concorsi su Hannah, Disney Channel, gadget, info eventi, musica, richieste informazioni sulla serie tv, riviste, scambio di informazioni;
- identificazione, all'interno della quale andranno: difesa di Miley, difesa di Miley e l'amore, elementi di esaltazione, immedesimazione, sovrapposizione di identità;
- distacco, all'interno della quale andranno: disincanto, non difesa di Miley, non difesa di Miley e l'amore;
- attrazione per il/i personaggio/i, all'interno della quale andranno: elementi di attrazione Miley, elementi di attrazione Hannah, fisicità Miley;
- narrazione episodi scandalistici (alcool, sesso), all'interno della quale andranno: alcool e Miley, foto osé, Miley e l'amore;
- famiglia propria, all'interno della quale andranno: genitori;
- slang, che conterrà se stesso.

STEP 3 – LAVORARE SULL'INFORMAZIONE: LE SUPER FAMIGLIE

Non tutti gli elementi che si creano durante il processo di lavoro confluiscono nel testo che ne raccoglie e presenta i risultati. Alcuni di questi prodotti servono per poter porre in essere elaborazioni successive che li contengono e allo stesso tempo li superano. Questo vale per esempio nel caso della creazione delle super famiglie. Per poterle realizzare è necessario prima aver creato i codici, poi aver organizzato i codici in famiglie e quindi mettere in relazione le informazioni contenute all'interno delle famiglie stesse mediante operazioni che richiedono l'uso degli operatori booleani.

È il **Super Family Tool** che ci permette di lavorare con le famiglie.

Quello che ci consente di realizzare sono delle operazioni di inclusione, esclusione dei codici contenuti all'interno di esso che poi portano appunto alla costituzione delle super famiglie. Una super famiglia è quindi il frutto di una operazione fra famiglie. Utilizzando gli operatori booleani è possibile:

- creare una super famiglia sommandovi all'interno i codici di due o più famiglie selezionate (OR);
- creare una super famiglia che abbia al suo interno i codici non comuni alle famiglie che si stanno ponendo in relazione (XOR);
- creare una super famiglia mantenendo all'interno di essa solo i codici comuni a entrambe (AND);
- creare una super famiglia azzerando uno degli elementi posti in relazione e mantenendone solo uno.

Nel caso della creazione della mappa degli elementi di attrazione per la protagonista della serie si è proceduto sommando all'interno di una super famiglia tutti gli elementi di identificazione con Hannah/Miley più gli elementi di attrazione. In questo caso sono stati radunati e sintetizzati sotto un unico network tutti gli elementi utili per analizzare il processo di identificazione fra fans e personaggio.

STEP 4 – COME ESTRARRE PORZIONI DI INFORMAZIONE

Per cogliere l'incidenza degli aspetti di carattere commerciale collegati ai *tweens* e presenti nel forum è necessario selezionare i frammenti di testo pertinenti. Questa procedura si realizza mediante il **Query Tool**.

Per riuscire a estrarre tutte le *quotations* che fanno riferimento agli aspetti commerciali nel topic dedicato ad Hannah e in quello di Miley è necessario:

- scegliere la famiglia di codici appropriata;
- selezionare l'operatore AND;
- introdurre i codici che fanno riferimento a questi elementi, quindi: concorsi su Hannah, Disney Channel, gadget, info eventi, musica, richieste informazioni sulla serie tv, riviste, scambio di informazioni;
- selezionare le variabili (ovvero i due topic) in cui va ricercata la query
- e infine nominare l'operazione.

A questo punto è possibile visualizzare le porzioni di testo che soddisfano questi criteri. L'esportazione dei risultati nella versione HTML permette di visualizzare con più facilità le *quotations*.

Attraverso questa procedura potremo ritrovare tutti quei frammenti che si riferiscono agli aspetti commerciali presenti nel forum.

STEP 5 – SOLLETICARE L’IMMAGINAZIONE CREATIVA

Sebbene Atlas.ti e più in generale i CAQDAS non siano software specifici per l’analisi linguistica, l’opzione di codifica ci permette di recuperare ogni informazione sia presente nel testo. In questo modo leggendo riga per riga il testo, tutto quanto il ricercatore trovi di interesse può essere salvato per successive riflessioni. Nel caso del linguaggio utilizzato dai *tweens* si sarebbe potuto procedere:

- o codificandoli come *Invivo Codes* e quindi recuperando ogni singola informazione;
- oppure come è stato fatto qui creando già un apposito codice “slang” e salvandovi all’interno tutte le espressioni ritenute utili.

Cosa farci delle informazioni recuperate in questo modo, ma anche fra quali informazioni selezionare quelle da salvare, spetta al singolo ricercatore e a quanto la sua immaginazione gli suggerisce.

STEP 6 – RIASSUMERE E RAPPRESENTARE L’INFORMAZIONE: I NETWORK

La procedura di lettura riga per riga del testo e la successiva codifica ci conducono a un livello elevato di analisi e conoscenza dei legami presenti nei documenti. È indubbio che leggendo, codificando, riaggregando, estraendo informazioni di dettaglio il livello di penetrazione nel testo e del testo cresce nel ricercatore fino a consentirgli di porre in essere relazioni fra i dati, cui prima non aveva pensato. Tali informazioni vanno esplicitate sia in forma testuale ma anche grafica, ovvero occorre riassumere le connessioni individuate e renderle fruibili ai lettori. Per fare ciò si adoperano i network.

Nell’introduzione a questo capitolo si era detto che obiettivo di questo lavoro era:

- ricostruire l’idea che di Hannah e di Miley hanno questi ragazzi, verificandone le differenze;
- far emergere i tratti dei *tweens* come target di vendita.

La funzione del network presentato in figura 2.6 è quella di rispondere a questi quesiti iniziali. Le relazioni poste in essere fra le varie famiglie di codici esplicitano questo rapporto. In fatti si ha che:

- la relazione “*is cause of*” è di tipo transitivo e pertanto applicata a codici che hanno una relazione di tale tipo fra di loro e con la/le protagonista/e;
- la relazione “*is a property of*” è di tipo asimmetrico e quindi nell’attribuirle bisogna evidenziare che c’è un prius nella generazione di tale legame. Per esem-

pio, qui si è sottolineato mediante essa che l'identificazione parte dei ragazzi verso Hannah e Miley è quindi una loro proprietà;

- la relazione “*is associated with*” è di tipo simmetrico e si applica ai codici che hanno tale legame fra di loro e con i due testi.

Si tenga presente che alcuni elementi possono essere successivamente aggiunti dal ricercatore manualmente, ovvero esportando gli elementi presenti in Atlas.ti e continuando il lavoro o in un file word o su un power point.

RIFERIMENTI BIBLIOGRAFICI

- ANTONELLI G. (2007) *L'italiano nella società dell'informazione*, Bologna, il Mulino.
- BARON N. S. (1984) “Computer Mediated Communication as a Force in Language Change”, in *Visible Language*, vol. 18, n. 2, Spring 1984, pp. 118-141.
- BUCKINGAM D. (2004) *Né con la tv, né senza la tv*, Milano, FrancoAngeli (ed. or. 2000).
- BUCKINGAM D. (2006) *Media education. Alfabetizzazione, apprendimento e cultura contemporanea*, Gardolo, Erickson, (ed. or. 2003).
- COOK T., KAISER S. (2004) “Betwix and Be Tween. Age ambiguity and the sexualization of the female consuming”, in *Journal of Consumer Culture*, London, Sage Publication, Vol.4(2), pp. 203-227.
- GILBERT E. (1957) *Advertising and Marketing to Young People*, Pleasantville, Printers' Ink Books.
- HALL C. (1987) “Tween PowerZ: Youth's Middle Tier Comes of Age”, in *Marketing and Media Decision*, (Oct.) pp.56-62.
- HULAN D. (2007) “Mistaking Brands for Tween Identity”, in *The McMaster Journal of Communication*, The Berkeley Electronic Press, vol.4(1), pp.41-52.
- LAVINIO C. (1990) *Teoria e didattica dei testi*, Firenze, Nuova Italia.
- LEMISH D. (2008) *I bambini e la tv*, Milano, Raffaello Cortina Editore (ed. or. 2007).
- LIVINGSTONE S. (2000) *La ricerca sull'audience. Problemi e prospettive di una disciplina al bivio*, Soneria Mannelli, Rubettino (ed. or. 1998).
- PASSERINI L., CAPUSSOTTI E., BRAUNSTEIN P. (1996) “La conversazione online tra oralità e scrittura”, in Aa.Vv. (1996) *Lo strabismo telematico. Contraddizioni e tendenze della società dell'informazione*, Torino, Utet.
- PISTOLESI E. (1998) “IRC (Internet Relay Chat): una nuova tecnologia della parola. Guida storica linguistica e tecnica”, <[http://www.italica.rai.it/facolta/lingua/pistolesi/ indicehtm](http://www.italica.rai.it/facolta/lingua/pistolesi/indicehtm)> (cons. 11/12/2003).
- QUART A. (2003) *Generazione ®. I giovani e l'ossessione del marchio*, Milano, Sperlingk & Kupfer (ed. or. 2003).
- RIVOLTELLA P. C. (2001) *Media Education. Modelli, esperienze, profilo disciplinare*, Roma, Carocci.

- VALKENBURG P.M., VROONE M. (2004) “Development changes in infants’ and toddler’ attention to television entertainment”, in *Communication Research*, 31(1), pp. 288-311.
- VAN EVRA J.P. (2004) *Television and Child Development*, 3rd ed., Mahmah, NJ, Lawrence Erlbaum.
- VEBLEN T. (1971) *La teoria della classe agiata*, Torino, Giulio Einaudi Editore (ed. or. 1899).

3.

IL LESSICO DELLA GUERRA NELLA COMUNICAZIONE ONLINE

3.1. INTRODUZIONE

La comunicazione mediata dal computer (CMC) rappresenta il fenomeno socio-culturale più significativo e pervasivo degli ultimi trent'anni (Paccagnella, 2000). La comunità accademica ha reagito con un certo ritardo ai cambiamenti sociali indotti da questa rivoluzione nel mondo della comunicazione. Internet compare citato per la prima volta in un abstract del *Cambridge Sociological Abstracts* nel 1992, ma il primo articolo scientifico relativo alle scienze sociali che cita Internet nel titolo è del 1994 (Lafrance, 1994). Il termine *Computer-Mediated Communication* è presente nel titolo di un lavoro di Elaine B. Kerr e Starr R. Hiltz del 1982 (Kerr e Hiltz, 1982), poi ripreso dalla stessa Hiltz e da Murray Turoff (Hiltz e Turoff, 1985), eppure nel 2000, quindici anni dopo, c'erano solo 34 articoli che facevano riferimento a questo argomento in modo significativo.

La CMC ha favorito la formazione di aggregazioni sociali che in un primo tempo si sono identificate con il nome controverso di “comunità virtuali” (Giuliano, 1999, 2002a e 2002b; Baym, 2002; Metitieri, 2003) e poi più recentemente come “social network”. La loro caratteristica comune è la condivisione di un “territorio digitalizzato” che oggi è uno dei luoghi più frequentati del paesaggio mediatico quotidiano. Gli utenti di Internet alla data del 30 settembre 2009, secondo l'Internet World Stats (www.internetworldstats.com/), erano 1.734.000.000; il 25,4% di essi appartiene ai due grandi paesi emergenti dell'Asia (Cina e India), anche se la “penetrazione” di Internet in questi paesi è del 27% in Cina e del 7% in India contro il 52% in Europa e il 72% nel Nord America.

I “cittadini del ciberspazio” non sono solo fruitori di informazione. Con lo sviluppo delle tecnologie Web 2.0 essi stessi sono diventati sempre più dei

“produttori” di informazione, vitalizzando ambienti multiculturali in grado di spezzare le barriere linguistiche. Il termine *prosumer*, introdotto da Alvin Toffler in un libro del 1980, *The Third Wave* (Toffler, 1987), sta ad indicare una sintesi tra produttore/professionista e consumatore, una tendenza verso la forma attiva dell’utente che non coinvolge soltanto il mercato (o una sua specifica segmentazione) agli effetti del marketing ma, in generale, tutto il processo di ideazione, produzione, distribuzione e consumo dell’informazione.

La trasformazione dell’internet user da passivo ad attivo è il tratto distintivo dell’intelligenza collettiva che, seguendo De Kerckhove (1991), si può chiamare *inter-brainframe*, la messa in rete di una molteplicità di strutture cognitive che ha caratterizzato l’avvento della cultura digitale (Galimberti e Riva, 1997). Questa nuova forma di spazio antropologico, secondo Pierre Lévy (1999), delimita uno “spazio del sapere” che non è soltanto competenza tecnica e scientifica, ma è un complesso spazio semiotico fatto di emozioni, apprendimento, condivisione di esperienze, conoscenze.

Lo spazio del sapere in Internet è in gran parte sommerso, non viene nemmeno raggiunto dai motori di ricerca e rimane – nonostante il suo interesse per l’analisi sociologica – al di fuori da ogni possibilità di lettura, interpretazione e spiegazione. In questo spazio, costituito in gran parte da “parole”, è necessario avere a disposizione strumenti di orientamento e navigazione. Di fatto in Internet chiunque sia connesso alla rete ha la possibilità di scrivere e “pubblicare” tutto ciò che gli passa per la testa senza alcun criterio di selezione predeterminato. Pescare le informazioni che ci interessano, o almeno rinvenire nello spazio digitalizzato nuclei di informazione dotati di senso, è (e lo sarà ancora di più in futuro) un’impresa ardua. La tecnologia dei motori di ricerca ci viene in aiuto per alcuni aspetti che riguardano il reperimento dell’informazione in web, ma è in grado di intervenire solo a livello di *information retrieval* e non certo di analisi del contenuto semantico, nonostante gli sforzi compiuti in questa direzione dai motori di ricerca dell’ultima generazione come *Bing* e, soprattutto, *Wolfram*. Gli sviluppi più recenti della CMC, dai forum online ai servizi di social networking (dal primo fenomeno *MySpace* del 2003, sviluppato da Tom Anderson e Chris DeWolfe, ai più recenti *Facebook* di Marl Zuckerberg e *Twitter* di Jack Dorsey) hanno messo in rilievo l’importanza decisiva che ha la circolazione delle comunicazione in rete sul piano economico, politico e culturale. (Hinduja *et al.* 2008; Lewis, *et al.*, 2008; Ackland, 2009; West *et al.*, 2009; Zhang *et al.*, 2010).

I social network di oggi hanno caratteristiche diverse rispetto ai newsgroup di Usenet che fino a qualche anno fa avevano il monopolio delle “interazioni di massa” online (Resnick, 1997; Paccagnella, 1997). I newsgroup hanno iniziato la loro attività come “bacheche elettroniche” nel 1979 e, dal

2001, sono accessibili con una specifica opzione di Google. I newsgroup di Usenet, classificati per gruppi linguistici e per argomenti (più di 25.000 secondo Choi e Danowski, 2002), hanno rappresentato fino al 2004 una parte consistente della messaggeria elettronica asincrona con un traffico di circa 1.000.000 di messaggi al giorno, pari a 70-80 terabytes all'anno (Lyman *et al.*, 2003). I newsgroup sono stati in gran parte sostituiti dai forum online, prima, e dai social network, poi. Oggi, sebbene presenti per una parte residuale delle infinite nicchie di Internet, non sono più un riferimento di base per la CMC. Tuttavia per le modalità di analisi della messaggeria asincrona l'archivio di testi a disposizione (dal 1985) rappresenta un materiale prezioso che può essere utilizzato come modello da applicare ad altre basi testuali simili per gli aspetti formali, retorici e di contenuto.

3.2. IPOTESI DI LAVORO

In questo lavoro prendiamo in esame i newsgroup di argomento politico. I problemi metodologici e tecnici che si incontrano nell'analisi del contenuto di questi messaggi sono in gran parte identici a quelli che si incontrano nell'analisi della posta elettronica in genere. Lo scopo è di esplorare le potenzialità di applicazione di una strategia di text-mining e di analisi dei dati testuali al fine di individuare, con un ridotto dispendio di risorse, i contenuti generali dei messaggi, i temi trattati dagli scriventi e, possibilmente, le loro opinioni.

Le maggiori difficoltà per l'analisi del contenuto di questi materiali sono rappresentate dalla massa enorme di dati che occorre elaborare e interpretare (Whittaker *et al.*, 1998) e dalla presenza di una componente di "rumore" che è spesso difficile, se non impossibile, eliminare completamente (Giuliano, 2004). I messaggi di posta elettronica in generale (cui possono essere ricondotti come tipologia i messaggi dei newsgroup e dei forum) sono noti per il valore basso nel rapporto segnale/rumore. Nello schema di comunicazione di Shannon e Weaver, il "rumore" è qualcosa di involontario che disturba la decodifica del messaggio da parte del ricevente. Nel nostro caso il "rumore di fondo" è quella parte di informazione (volontaria o involontaria) che rappresenta un ostacolo per la corretta identificazione dei nuclei di significato presenti nel testo. Sicuramente involontaria è la componente "ambientale" del rumore che deriva dalla natura del messaggio elettronico: transcodifiche non correttamente riconosciute dai diversi sistemi operativi (sebbene oggi questo problema sia sempre più raro); presenza di indirizzi email, web-link, sequenze alfanumeriche che derivano dall'header del messaggio o da altre fonti estranee al corpo del testo che ne trasmette i contenuti. Volontario invece, ma non del tutto estraneo al contenuto del

messaggio è il rumore “retorico” che deriva dal registro stesso della comunicazione linguistica: acronimi ed espressioni gergali, abbreviazioni cifrate note solo agli utenti, citazioni di messaggi ai quali lo scrivente ha dato risposta (*quoting*), firme e motti ricorrenti (*signatures*) che segnalano l'autore del messaggio; brevi richiami pubblicitari inseriti automaticamente dai server di posta gratuita. Evidentemente si tratta, in questo caso, di una forma di disturbo che può tuttavia veicolare un contenuto e la sua eliminazione automatica deve essere attentamente valutata. Infine vi è il rumore “polemico” generato da interventi sicuramente volontari e originati da provocatori (nel gergo pionieristico di Internet erano identificati come *trolls*, creature fatate un po' sgradevoli della mitologia nordica) che con i loro messaggi cercano di sollecitare risposte emotive e litigiose da parte dei frequentatori abituali del gruppo. Il troll, se non viene ignorato, scatena frequentemente una “rissa telematica” (*flame*).

3.3. IL CORPUS IRAQ28

Il corpus Iraq28 (tab. 3.1) è costituito dai 35.492 messaggi (33 Mb) inviati nei newsgroup della gerarchia it.politica durante la guerra in Iraq, e precisamente nel mese che va dal 18 marzo al 18 aprile 2003 (Giuliano, 2006). I newsgroup della categoria <it.politica> nel momento della rilevazione erano 10, ma <it.politica.libertaria> (245 messaggi) e <it.politica.ulivo> (349 messaggi) sono stati scartati dalla rilevazione per il numero troppo esiguo di messaggi rispetto a una soglia arbitraria ma “ragionevole” di 20 messaggi in media al giorno (Step 1).

Tab. 3.1 – Newsgroup della categoria it.politica secondo i messaggi nel periodo considerato: 18 marzo – 18 aprile 2003

Newsgroup	N. messaggi	Media msg al giorno	Denominazione
it.politica	13.931	435,34	POLITICA
it.politica.internazionale	6.194	193,56	INTERNAZIONALE
it.politica.pds	2.981	136,25	PDS
it.politica.lega-nord	3.809	119,03	LEGA-NORD
it.politica.polo	4.496	112,53	POLO
it.politica.rifondazione	2.728	90,91	RIFONDAZIONE
it.politica.destra	686	21,44	DESTRA
it.politica.cattolici	667	20,84	CATTOLICI

Occorre ricordare che i messaggi sono il risultato di una modalità asincrona di interazione, con un tono implicito di discussione pubblica e con uno stile in gran parte informale (sebbene vi siano messaggi che derivano da una comunicazione molto più formale, come, ad esempio, il “copia e incolla” di articoli tratti da giornali quotidiani e da comunicati di agenzie di stampa). L’interazione è volontaria, potenzialmente globale per le modalità di accesso e anarchica dal punto di vista gestionale. Solo il newsgroup <it.politica.destra> risultava gestito da un moderatore autorizzato a tagliare e – eventualmente - filtrare i messaggi in ingresso.

Il corpus, sottoposto a normalizzazione completa in TaLTaC (Step 2), presenta le seguenti misure lessicometriche:

Tab. 3.2 – Misure lessicometriche del corpus Iraq28

Occorrenze	N	5.220.932
Forme grafiche	V	179.112
Type/Token ratio	(V/N)*100	3,43
Percentuale di hapax	(V1/V)*100	46,85
Frequenza media generale	N/V	29,49

La fascia di alta frequenza contiene 133 forme grafiche (tab. 3.3), tra le quali, se escludiamo le forme grammaticali e le forme banali che derivano dal linguaggio di Internet, troviamo le parole chiave principali che dimostrano una reattività eccezionale degli utenti di questi newsgroup agli eventi in corso.

Tab. 3.3 – Forme grafiche principali della fascia di alta frequenza nel corpus Iraq28

Forme	Occorrenze	Rango
guerra	15.605	35
Iraq	9.011	56
Saddam	6.208	81
mondo	5.413	93
Bush	5.381	97
governo	4.587	117
USA	4.276	124
Italia	4.237	126
americani	4.048	130
pace	4.023	131

Tra le forme più significative, *guerra* è la forma più frequente. Seguono i principali riferimenti al contesto (*Iraq, Saddam, Bush, USA, americani, pace*). In questa fase alcuni controlli compiuti sull'interferenza del "rumore" nell'analisi automatica del contenuto offrono dei risultati rassicuranti. In corpora con milioni di occorrenze, limitatamente all'analisi automatica del contenuto, la presenza di forme grafiche considerate "di disturbo" (soprattutto se per effetto di rumore retorico o ambientale) appare del tutto ininfluyente (Giuliano, 2004).

3.4. IL SUB-CORPUS GUERRA-IRAQ

Un modo per rendere trattabili corpora di grandi dimensioni come Iraq28 è quello di costituire corpora ridotti automaticamente sulla base di argomenti identificabili attraverso una o più parole chiave. Nel nostro caso la parola più frequente nel corpus è la forma *guerra*. Su questa base si è costruito (Step 3) un corpus ridotto di concordanze costituito di 18.698 stringhe di testo per un totale di 474.476 occorrenze e 32.945 forme grafiche distinte.

Tab. 3.4 – Misure lessicometriche del sub-corpus delle concordanze Guerra-Iraq

Occorrenze	N	474.476
Forme grafiche	V	32.945
Type/Token ratio	(V/N)*100	6,94
Percentuale di hapax	(V1/V)*100	45,18
Frequenza media generale	N/V	14,40

Le forme grafiche non riconosciute dal tagging grammaticale di TaLTaC sono pari al 26,28% delle forme grafiche distinte; alla soglia prescelta per l'analisi delle corrispondenze (soglia 20, copertura del testo 80%) le forme non riconosciute rappresentano solo il 3,69% delle forme grafiche.

3.5. ANALISI DELLE CORRISPONDENZE DEI SEGMENTI RIPETUTI

Per analizzare il contenuto dei messaggi si è scelto di fare affidamento sull'individuazione dei temi in discussione attraverso la frequenza dei nuclei semantici rappresentati dai segmenti ripetuti nelle concordanze (Step 4).

I segmenti più rilevanti per il tema in oggetto possono essere selezionati assumendo come parametri la frequenza del segmento, l'indice IS assoluto

(che mette in evidenza i segmenti ripetuti composti da un maggior numero di parole) e l'indice IS relativo (che mostra il grado di assorbimento del segmento rispetto alle parole che lo costituiscono). Nella tabella. 3.5 sono elencati alcuni dei segmenti più significativi con indice IS relativo maggiore di 0,20.

Tab. 3.5 – Segmenti ripetuti nel sub-corpus delle concordanze Guerra-Iraq

Segmenti ripetuti	Occorrenze	IS assoluto	IS relativo
Stati Uniti	315	2,88	0,72
Nazioni Unite	76	2,85	0,71
Medio Oriente	45	2,84	0,71
Casa Bianca	59	2,84	0,61
opinione pubblica	62	2,14	0,54
guerra fredda	168	1,90	0,48
Saddam Hussein	122	1,89	0,47
Unione Europea	26	1,76	0,44
guerra preventiva	327	1,69	0,42
guerra santa	88	1,67	0,42
quarta guerra	47	1,65	0,41
seconda guerra	308	1,64	0,41
guerra infinita	102	1,59	0,40
guerra mondiale	514	1,57	0,39
questa guerra	1.821	1,53	0,38
diritti umani	31	1,52	0,38
seconda guerra mondiale	262	3,25	0,36
guerra lampo	35	1,25	0,31

Nella tabella 3.5 non sono evidenziati i segmenti per ordine di occorrenze, avendo privilegiato l'indice IS relativo. Tuttavia il clima presente nei newsgroup è ben rappresentato dal segmento *contro la guerra* che è il più frequente (780 occorrenze), sebbene con un IS=0,04 dovuto alle forme “banali” delle quali è composto (*contro*=2.476; *la*=14.724; *guerra*=20.185).

L'analisi delle corrispondenze binarie (Step 5) sulla tabella segmenti con soglia 20 per newsgroup effettuata con il programma SPAD 5.0 permette di individuare i gruppi tematici caratteristici dei diversi newsgroup.

Tab. 3.6 – Analisi delle corrispondenze dei segmenti ripetuti per newsgroup nel corpus Guerra-Iraq: estrazione degli autovalori

Autovalori	% di inerzia	% cumulata di inerzia
0.0685	22,37	22,37
0.0587	19,17	41,54
0.0522	17,05	58,59
0.0415	13,57	72,16
0.0338	11,03	83,19
0.0287	9,36	92,55
0.0228	7,45	100,00

Tab. 3.7 – Analisi delle corrispondenze dei segmenti ripetuti per newsgroup nel corpus Guerra-Iraq: coordinate e contributi assoluti delle frequenze attive

Newsgroup	Coordinate		Contributi assoluti	
	F1	F2	F1	F2
CATTOLICI	0,31	0,86	6,9	59,8
DESTRA	-0,49	-0,02	13,1	0,0
INTERNAZIONALE	-0,32	0,15	39,2	10,7
POLITICA	0,25	-0,04	26,6	0,7
LEGA-NORD	0,12	-0,19	1,7	4,4
PDS	-0,11	-0,29	2,1	16,9
POLO	0,29	-0,04	9,7	0,2
RIFONDAZIONE	-0,08	-0,23	0,8	7,3

I primi due assi fattoriali spiegano il 41,54% della variabilità complessiva. Le coordinate e i contributi assoluti delle frequenze attive permettono di individuare una diversificazione attesa sul primo fattore tra i due newsgroup principali, <it.politica> e <it.politica.internazionale>, intorno a un asse di politica interna e politica estera. Non sorprende la collocazione sul versante internazionale del newsgroup <it.politica.destra> che in passato ha ospitato con una certa intensità messaggi che denotano una forte adesione ai principi della destra estrema, antisionista ed estranea ai temi della politica nazionale.

Il secondo asse fattoriale si caratterizza per la contrapposizione tra <it.politica.cattolici> e <it.politica.pds> intorno a temi che fanno riferimento al dibattito sulla opposizione o no alla guerra e alle sue motivazioni. In tabella 3.8 sono riportati solo i segmenti ripetuti che presentano un contributo significativo per la formazione di almeno uno dei due fattori estratti.

Tab. 3.8 – Analisi delle corrispondenze dei segmenti ripetuti per newsgroup nel corpus Guerra-Iraq: coordinate e contributi assoluti dei segmenti

Segmenti	Coordinate		Contributi assoluti	
	F1	F2	F1	F2
a guerra finita	-0,16	-0,35	0,1	0,8
amministrazione Bush	-0,41	0,13	1,3	0,2
Amnesty International	-0,54	0,36	1,3	0,7
concetto di guerra	0,41	0,12	0,7	0,1
Consiglio di Sicurezza	-0,62	0,15	2,6	0,2
contrario alla guerra	-0,47	0,07	3,1	0,1
contro gli americani	0,08	-0,48	0,0	0,9
contro l'Iraq	-0,31	0,06	2,7	0,1
contro tutte le guerre	0,68	0,63	2,2	2,1
criminale di guerra	-0,59	-0,17	2,6	0,2
dichiarazione di guerra	-0,56	0,62	1,4	2,1
diritto internazionale	-0,41	0,15	1,7	0,3
favorevole alla guerra	0,42	0,15	1,1	0,2
fermare la guerra	0,07	-0,47	0,1	6,3
grande guerra	0,36	0,57	0,6	1,7
guerra breve	0,05	-0,67	0,0	2,1
guerra dei trent'anni	0,81	-0,17	2,3	0,1
guerra di aggressione	-0,44	-0,28	1,5	0,7
guerra di Bush	0,23	-0,15	0,7	0,4
guerra di difesa	0,49	0,89	0,7	2,8
guerra di liberazione	0,22	0,14	0,8	0,3
guerra è guerra	0,62	-0,27	1,3	0,3
guerra finisce	0,52	-0,26	1,0	0,3

guerra giusta	0,52	1,01	4,4	19,5
guerra illegale	-0,43	0,02	1,0	0,0
guerra illegittima	-0,10	-0,41	0,1	1,1
guerra in Afghanistan	-0,39	0,18	1,0	0,3
guerra irachena	-0,07	-0,56	0,0	3,2
guerra legale	0,94	-0,16	3,0	0,1
guerra lunga	0,39	-0,45	1,7	2,8
guerra nucleare	-0,51	-0,01	1,0	0,0
guerra santa	0,26	0,17	0,8	0,4
guerra voluta	0,18	-0,42	0,1	0,8
guerre giuste	0,43	0,35	0,9	0,7
guerre mondiali	-0,15	-0,51	0,1	1,4
in caso di guerra	-0,63	-0,07	1,5	0,0
legittima difesa	0,78	0,06	2,2	0,0
liberazione dell' Iraq	0,68	0,09	1,6	0,0
movimento contro la guerra	-0,78	0,12	3,1	0,1
Nazioni Unite	-0,12	0,35	0,2	1,6
no alla guerra	0,25	-0,02	2,5	0,0
nostro paese	0,40	0,59	1,1	2,8
nuova guerra	-0,48	0,07	1,6	0,0
opposizione alla guerra	-0,31	-0,25	0,8	0,6
popolazione civile	0,42	0,96	0,7	4,3
popolo iracheno	0,38	-0,19	2,1	0,6
prigionieri di guerra	-0,31	0,18	1,6	0,6
prossima guerra	-0,58	-0,02	1,1	0,0
quarta guerra mondiale	-0,13	-0,31	0,1	0,8
ripudia la guerra	-0,03	-0,35	0,0	1,8
scoppiata la guerra	0,47	-0,48	0,7	0,9
seconda guerra del golfo	-0,28	-0,54	0,3	1,3
termine guerra	0,57	-0,20	1,1	0,2
Unione Europea	-0,25	0,05	0,7	0,0
uranio impoverito	-0,61	-0,15	1,7	0,1
vittime civili	0,57	-0,13	2,8	0,2

I contenuti emergono più chiaramente esaminando la proiezione dei segmenti ripetuti sul piano fattoriale formato dagli assi 1 e 2 (fig. 3.1). Nel primo quadrante (++), sul quale è collocato il newsgroup <it.politica.cattolici> troviamo segmenti che rappresentano posizioni diversificate intorno al tema “umanitario” della guerra: *guerra giusta*, *contro tutte le guerre*, *popolazione civile*, *guerra santa*, *concetto di guerra*, *liberazione dell'Iraq*.

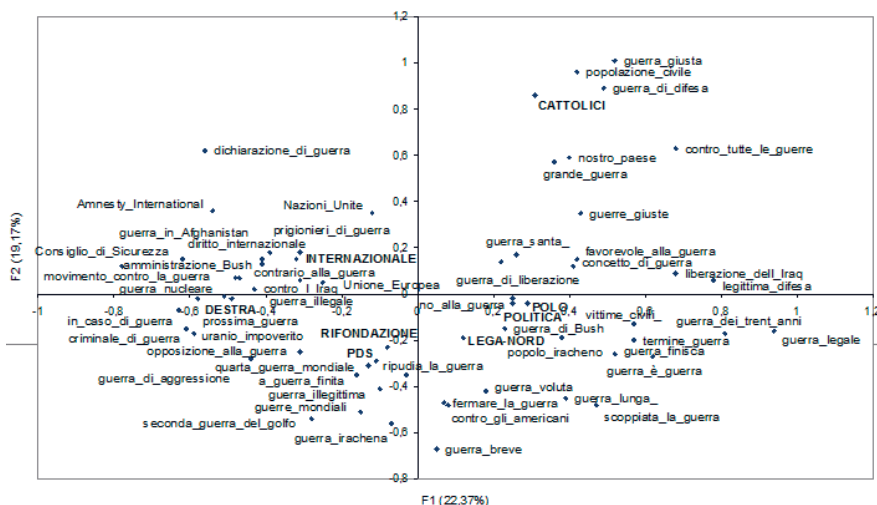


Fig. 3.1 – Analisi delle corrispondenze dei segmenti ripetuti per newsgroup nel corpus Guerra-Iraq: proiezione dei segmenti sul piano fattoriale degli assi 1 e 2

Nel quadrante opposto (--) troviamo invece temi di più marcata connotazione politica che esprimono una opposizione netta a questa seconda guerra del golfo: *guerra illegale*, *guerra illegittima*, *guerra di aggressione*. In questo quadrante si collocano i due newsgroup che fanno riferimento alla opposizione di sinistra: <it.politica.pds> e <it.politica.rifondazione>.

Sul secondo quadrante (+ -) formato dal semiasse positivo dell’asse 1 e dal semiasse negativo dell’asse 2, troviamo segmenti che rievocano le motivazioni della guerra: *guerra di Bush*, *guerra voluta*, *contro gli americani*, *popolo iracheno*, *vittime civili*. Non mancano alcuni riferimenti che sembrano giustificare l’intervento militare: *guerra legale*, *guerra è guerra*. Anche qui si individuano dei segmenti che denotano un dibattito tra favorevoli e contrari alla guerra: *fermare la guerra*, *guerra finisca*, *no alla guerra*, *guerra breve*, *guerra lunga*. In questo quadrante si colloca soprattutto il newsgroup <it.lega-nord>.

Sul quarto quadrante (– +), formato dal semiasse negativo dell’asse 1 e dal semiasse positivo dell’asse 2, troviamo segmenti che fanno riferimento agli organismi internazionali (*Nazioni Unite*, *Consiglio di sicurezza*, *Unione Europea*, *Amnesty International*) e alla collocazione della guerra nel contesto della legalità internazionale, con espliciti riferimenti a prese di posizione contro la guerra. Il newsgroup <it.politica.internazionale> si caratterizza in modo significativo per la discussione intorno a questi argomenti.

3.6. CONCLUSIONE

Come si è visto nel corso della trattazione, l’analisi del contenuto dei newsgroup presenta diversi problemi metodologici e tecnici soprattutto per l’enorme massa di dati testuali da trattare e per l’alta componente di rumore riscontrata. Pertanto si è applicata una strategia di analisi “a imbuto” con tre tappe di riduzione del dato testuale passando dal corpus complessivo (Iraq28) a un sub-corpus (Guerra-Iraq) ristretto alle stringhe di testo delle concordanze intorno alla parola pivot *guerra* e poi, successivamente, all’analisi delle corrispondenze sulla tabella lessicale dei segmenti ripetuti per newsgroup.

Questa strategia di analisi si è rivelata adeguata per l’individuazione dei temi principali del dibattito sulla guerra in Iraq. In questi messaggi tuttavia non è stato possibile individuare tendenze ideologico-politiche distinte in sintonia con le scelte tematiche dei newsgroup. Dall’analisi delle concordanze non è emerso un lessico specifico della guerra in relazione con le presunte appartenenze ideologiche degli utenti quanto, piuttosto, in relazione con i temi individuati che sono trasversali rispetto ai newsgroup stessi. Gli utenti amano intrattenersi in un dibattito politico attento, informato, vivace e caratterizzato da un elevato tono polemico. Tuttavia non emergono comportamenti verbali significativamente oltraggiosi se non in misura del tutto occasionale e frutto di provocazioni estranee al clima generale. Questa osservazione, tenuto conto anche dell’alto livello di conflittualità sociale generato da un evento come la guerra, costituisce una conferma della rilevanza delle conversazioni in gruppi di discussione telematica per la formazione delle opinioni politiche.

Le parole chiave individuate tramite l’analisi automatica del contenuto hanno messo in evidenza, complessivamente, quattro temi principali di discussione.

- 1) Il dibattito intorno alla legittimità dell’intervento militare e alla politica dell’amministrazione Bush; i riferimenti favorevoli si accompagnano a motivazioni come la “guerra di liberazione”, la “guerra santa”, la “guerra giusta”, la “guerra di legittima difesa”; i riferimenti sfavorevoli si basano sulla “illegalità” internazionale e sulla “guerra di aggressione”, sul pericolo di guerra

nucleare e di estensione del conflitto, fino a configurare scenari da terza o quarta guerra mondiale.

- 2) Il dibattito intorno alle scelte politiche del governo italiano nel quadro delle scelte compiute da altri Paesi europei; in questo ambito si evidenziano anche i riferimenti alla guerra preventiva e gli interventi del Papa, con le relative implicazioni religiose e umanitarie.
- 3) Il dibattito intorno alla polemica sul pacifismo e alle manifestazioni di piazza contro la guerra.
- 4) Il dibattito intorno alla guerra e al suo dramma, ai morti tra la popolazione civile, alle bombe, ai soldati, agli interventi di tv e giornalisti; questo tema è trasversale e non sembra connotato da orientamenti politici ed ideologici.

La metodologia adottata può avere conseguenze molto interessanti per l'analisi dell'interazione in rete, in particolare per il monitoraggio delle opinioni politiche e anche per l'analisi delle strategie di formazione dei movimenti di opinione pubblica. È solo il caso di ricordare qui che i movimenti di Porto Alegre, Seattle, Okinawa, Genova e così via, divenuti poi movimento anti-global a livello internazionale, si sono consolidati come fenomeno politico proprio attraverso la formidabile occasione fornita dalla comunicazione mediata dal computer, dai newsgroup, dai forum e, in generale, dalla possibilità di comunicare in modo molto rapido, su scala mondiale saltando tutte le tradizionali forme di mediazione.

La difficoltà, già rilevata da altri ricercatori (Fiore *et al.*, 2002), di analizzare migliaia di messaggi, o anche milioni quando il periodo di osservazione si estende nell'arco di uno o più anni, con una base di dati testuali, come si è detto, contrassegnata da un alto livello di rumore, non appare di per sé insuperabile per l'estrazione automatica dell'informazione essenziale. Nella gestione di questa massa ingente di documenti, catalogabile e indicizzabile con parole chiave che ne identificano il contenuto, le metodologie di analisi automatica dei dati testuali e di text-mining rendono possibile la selezione di ciò che è importante per l'analisi qualitativa e in profondità.

STRATEGIA DI ANALISI

STEP 1 – ORGANIZZAZIONE E NORMALIZZAZIONE DEL CORPUS

I messaggi scaricati dal news-server sono composti da una intestazione (*header*), costituita da una serie di righe di testo seguita dal “corpo” del messaggio (*body*). Nella fase di preparazione il corpus è stato ripulito dell'*header* e delle stringhe al-

fanumeriche facilmente identificabili come immagini. Il corpus è stato sottoposto a normalizzazione completa in TaLTaC con identificazione di otto partizioni rappresentate dagli otto newsgroup in analisi. Si tratta di un corpus di grandi dimensioni (oltre 5 milioni di occorrenze) che potrebbe essere adeguato per la costruzione di un lessico di riferimento.

STEP 2 – GENERAZIONE DEL VOCABOLARIO E MISURE LESSICOMETRICHE

Il secondo step comporta, in genere, l'acquisizione del corpus tramite il parsing, la generazione del vocabolario e le prime misure lessicometriche. Con TaLTaC dal menu **Analisi** – *Pre-trattamento* – *Calcolo delle sub-occorrenze* possiamo compiere alcune osservazioni sulle occorrenze secondo i newsgroup. Le misure lessicometriche e le altre operazioni sul vocabolario saranno ripetute e affinate con il procedere delle operazioni di pulizia del testo (se necessarie) e di lessicalizzazione. Le misure lessicometriche di base permettono di valutare l'adeguatezza statistica del corpus. In questo caso il rapporto forme grafiche / occorrenze (type / token ratio) è basso, infatti il corpus è molto grande e sufficientemente esteso da poter essere utilizzato per l'estrazione di un lessico del linguaggio politico nei newsgroup. Anche la percentuale di hapax, pur essendo un corpus relativamente "sporco", è nella norma.

STEP 3 – ANALISI DELLE CONCORDANZE ED ESTRAZIONE DEL SUB-CORPUS

Dal menu **Analisi** – *Analisi testuale* – *Recupero di informazione: concordanze* – *Concordanze semplici e complesse* possiamo selezionare la forma pivot *guerra* e identificare le stringhe di testo composte da 70 caratteri che precedono e 70 caratteri che seguono la forma pivot. Le stringhe di testo delle concordanze (dotate di identificatore della modalità newsgroup) si possono salvare nella cartella di lavoro per costituire il sub-corpus).

STEP 4 – INDIVIDUAZIONE DEI SEGMENTI RIPETUTI, LESSICALIZZAZIONE E COSTRUZIONE DELLA TABELLA SEGMENTI PER NEWSGROUP

Dalla barra degli strumenti di TaLTaC, cliccando sullo strumento "Individua

segmenti” (o dal menu **Analisi** – *Analisi lessicale*, selezionando *Analisi dei segmenti* e poi la voce *Individuazione dei segmenti*) si accede alla finestra di dialogo per l'esecuzione della procedura. I segmenti sono stati individuati con soglia di frequenza minima delle parole appartenenti al segmento pari a 50 e soglia dei segmenti con occorrenze pari almeno a 20.

Il calcolo dell'indice di significatività dei segmenti permette di valutare la loro rilevanza nel corpus. L'indice si ottiene dal menu **Analisi** – *Analisi lessicale*, selezionando *Analisi dei segmenti* e poi la voce *Calcolo indice IS su – Lista dei segmenti*. L'indice IS relativo assume un valore prossimo a 1 quando le forme grafiche che lo compongono sono presenti quasi esclusivamente in associazione con le altre parole del segmento. È evidente che, in questo caso, è opportuno creare una forma complessa (poliforme) da trattare come una sola forma grafica. Nel nostro caso la lessicalizzazione non è stata di tipo linguistico ma di tipo semantico. La costruzione di una forma complessa (ad esempio: *contro la guerra*), indipendentemente dal valore dell'indice IS (solo 0,04 nell'esempio), è servita a costruire una tabella di contingenze segmenti per newsgroup sulla quale è stata applicata una tecnica di analisi delle corrispondenze binarie (step 5).

STEP 5 – ANALISI DELLE CORRISPONDENZE BINARIE SULLA TABELLA LESSICALE: SEGMENTI PER NEWSGROUP

L'analisi delle corrispondenze trasforma lo spazio generato dalle variabili di partenza in uno spazio euclideo generato da nuove variabili (assi) che permettono di sostituire le vecchie coordinate con le nuove coordinate ortogonali (ottenute per combinazione lineare delle vecchie variabili). Il numero di dimensioni del nuovo spazio sarà pari al minimo tra righe (forme: segmenti ripetuti) e colonne (testi: newsgroup) -1. In questo caso, con otto testi ci saranno al massimo sette dimensioni. L'inerzia totale riprodotta (detta anche “traccia”) è la somma degli autovalori, pari qui a 0,36. Il primo asse (la migliore approssimazione alla matrice dei dati di partenza) riproduce il 22,37% dell'inerzia estratta (tab. 3.6).

In generale per una trattazione completa dell'analisi multidimensionale dei dati e per una corretta interpretazione dei risultati si dovranno consultare testi specialistici di statistica. In particolare si consiglia il manuale di S. Bolasco (2004), con un'ampia bibliografia di riferimento. Per una trattazione più agile alla portata di un utente privo di una conoscenza statistico-matematica di base, si consiglia il cap. 5 del testo di A. Tuzzi (2003), oppure il manuale di Metastasio e Cini (2009) e il capitolo dedicato all'analisi delle corrispondenze nel manuale di Cannavò e Frudà (Vardanega, 2007).

Nella tabella 3.6 è riportata la somma dell'inerzia totale spiegata e la quota di inerzia (*eigenvalue*; autovalore o valore proprio) spiegata da ciascun fattore.

Nella tabella 3.7 vediamo l'output della procedura CORBI (analisi delle corrispondenze binarie) in SPAD 5.0 con riferimento alle coordinate e ai contributi assoluti delle colonne della matrice (testi). Nella tab. 3.8 vediamo lo stesso output relativo alle righe della matrice (143 righe/segmenti). Per una valutazione dell'importanza dei contributi assoluti alla formazione degli assi fattoriali, si può calcolare il contributo medio pari a $100/V$ (dove V rappresenta il numero di righe della matrice). In questo caso il contributo medio assoluto di ciascun segmento è pari a 0,7 (i segmenti in analisi sono 143). Pertanto potranno essere considerati "significativi" i contributi assoluti superiori a 0,7.

Gli indicatori più rilevanti per la lettura dell'output e che quindi offrono maggiori informazioni per l'interpretazione degli assi fattoriali sono:

- La massa (non riportato in tabella): è il "peso relativo" di ciascuna modalità (testo/newsgroup) rispetto al complesso della variabile (corpus). Per le forme grafiche (righe/segmenti) il peso è sempre piuttosto piccolo.
- L'indice di distorsione (non riportato in tabella): è la "distanza dall'origine" di ciascuna modalità; dà indicazioni sul carattere periferico di certe modalità; un indice di distorsione molto alto si associa spesso ad una massa ridotta, il che dovrebbe indurre ad una certa cautela nel dare rilevanza a quella modalità nell'interpretazione.
- Le coordinate fattoriali: indicano la posizione della modalità sull'asse sia in termini di distanza dall'origine che in termini di "versante" (positivo o negativo) dell'asse.
- Il contributo assoluto: è il contributo che ciascuna modalità dà alla formazione dell'asse ed è proporzionale alla massa e al quadrato della coordinata fattoriale; rappresenta la parte di inerzia totale del fattore che è spiegata dalla modalità. Per ogni asse fattoriale la somma dei contributi assoluti è uguale a 100.
- Il contributo relativo (non riportato in tabella): è un valore che rappresenta il contributo che un certo fattore F_i fornisce alla variabilità della modalità. Se il valore è basso vuol dire che la modalità non è ben rappresentata sull'asse fattoriale F_i e che quindi la sua variabilità è spiegata da altri fattori.

RIFERIMENTI BIBLIOGRAFICI

- ACKLAND R. (2009) "Social Network Services as Data Sources and Platforms for e-Researching Social Networks", in *Science Computer Review*, vol. 27, no. 4 (Nov.), pp. 481-492.
- BAYM N.K. (2002). "Il posto delle comunità online nella vita offline", in *Rassegna italiana di sociologia*, 43 (1), pp. 55-71.
- BOLASCO S. (2004) *Analisi multidimensionale dei dati*, Roma, Carocci.
- CHOI J.H., DANOWSKI J. (2003) "Making a Global Community on the Net - Global Village or Global Metropolis?: A Network Analysis of Usenet Newsgroups", in *Journal of Computer Mediated Communication*, 7 (3) consultato online il 27/3/2010; < <http://jcmc.indiana.edu/vol7/issue3/choi.html>>.
- DE KERCKHOVE D. (1991) *Brainframes: technology, mind and business*, Baarn, Bosch & Keuning.
- FIORE A.T., LEE TIERNAN S., SMITH M.A. (2002) "Observed Behavior and Perceived Value of Authors in Usenet Newsgroups: Bridging the Gap", in *Proceedings of the SIGCHI conference on Human factors in computing systems*. 4 (1), Minneapolis (USA), pp. 323-330.
- GALIMBERTI C., RIVA C. a cura di (1997) *La comunicazione virtuale. Dal computer alle reti telematiche: nuove forme di interazione sociale*, Milano, Guerini e Associati.
- GIULIANO L. (1999). "Comunità nel ciberspazio. Comunicazione elettronica e spazio sociale", in *Sociologia. Rivista Quadrimestrale di Scienze Storiche e Sociali*, 33 (2), pp. 47-63.
- GIULIANO L. (2002a) "Identità e narrazione in ambienti sociali digitali", in *Rassegna italiana di sociologia*, 43 (1), pp. 7-32.
- GIULIANO L. (2002b) "G8-2001: la rivolta nel monitor. Analisi testuale dei messaggi nel newsgroup <it.eventi.g8.genova> durante gli scontri di piazza", in *6ème J-ADT 13-15 mars 2002*, Saint-Malo, IRISA, pp. 301-311.
- GIULIANO L. (2004) "L'analisi automatica dei testi ad alta componente di rumore", in E. Aureli Cutillo e S. Bolasco, *Applicazioni di analisi statistica dei dati testuali*, Roma, Editrice La Sapienza, pp. 41-54
- GIULIANO L. (2006) "Analysis of the Content of Newsgroup Messages: Methodological and Technical Issues", In: P.-L. Law, L. Fortunati, and S. Yang eds. *New Technologies in Global Societies*, Singapore, World Scientific, pp. 107-124.
- HINDUJA S., PATCHIN J.W. (2008) "Social networking and identity construction: Personal information of adolescents on the Internet: A quantitative content analysis of MySpace", in *Journal of Adolescence*, vol. 31, no. 1 (Feb.), pp. 125-146.
- HILTZ S.R., TUROFF M. (1985) "Structuring computer-mediated communication systems to avoid information overload", in *Communications of the ACM*, 28, 7 (July), pp. 680 – 689.
- KERR E.B., HILTZ S.R. (1982) *Computer-Mediated Communication Systems: Status and Evaluation*, New York, Academic Press.

- LAFRANCE J-P. (1994) "La TVI, l'autoroute electronique, Internet et les superhighways de l'avenir: vers une restructuration de l'industrie audiovisuelle", in *Technologies de l'Information et Société*, vol. 6, no. 4, pp. 401-430.
- LÉVY P. (1999) *L'intelligenza collettiva. Per un'antropologia dekl cyberspazio*, Milano, Feltrinelli.
- LEWIS K., KAUFMAN J., GONZALEZ M., WIMMER A., CHRISTAKIS N. (2008) "Tastes, ties, and time: A new social network dataset using Facebook.com", in *Social Networks*, vol. 30, no. 4 (Oct.), pp. 330-342.
- LYMAN P., SWEARINGEN K., VARIAN H.R. (2003) *How Much Information?* The School of Information Management and Systems, University of California (Berkeley). <<http://www2.sims.berkeley.edu/research/projects/how-much-info-2003/>> (cons. online 27/3/2010).
- METASTASIO R., CINI F. (2009) *L'analisi del contenuto. Procedure di analisi dei dati con il programma SPAD*, Milano, F. Angeli.
- METTIERI F. (2003) *Comunicazione personale e collaborazione in rete. Vivere e lavorare tra e-mail, chat, comunità e gruppi*, Milano, FrancoAngeli.
- PACCAGNELLA L. (1997) "Verso una sociologia del cibernazio. Uno studio di caso sulla conferenza elettronica cyber-punk", in *Quaderni di Sociologia*, 13, pp. 33-57.
- PACCAGNELLA L. (2000) *La comunicazione al computer*, Bologna, Il Mulino.
- RESNICK D. (1997) "Politics on the Internet: The Normalization of Cyberspace", in *New Political Science*, 41-42, pp. 47-67.
- ROVERSI A. (2004) *Introduzione alla comunicazione mediata dal computer*, Bologna, il Mulino.
- TOFFLER A. (1987) *La terza ondata*, Milano, Sperling & Kupfer.
- TOSONI S. (2004) *Identità virtuali. Comunicazione mediata da computer e processi di costruzione dell'identità personale*, Milano, FrancoAngeli.
- TUZZI A. (2003) *L'analisi del contenuto*, Roma: Carocci.
- VARDANEGA A. (2007) "L'analisi delle corrispondenze", in L. Cannavò, L. Frudà (a cura di), *Ricerca sociale. Dall'analisi esplorativa al data mining*, Roma, Carocci, pp. 99-113.
- WEST A., LEWIS J., CURRIE P. (2009) "Students' Facebook 'friends': public and private spheres", in *Journal of Youth Studies*, vol. 12, no. 6 (Dec.), pp. 615-627.
- WHITTAKER S., TERVEEN L., HILL W., CHEMY L. (1998) "The dynamics of mass interaction", *Proceedings of Conference on Computer Supported Cooperative Work*, pp. 257-264; <<http://social.cs.uiuc.edu/class/cs598kgk/papers/p257-whittaker.pdf>> (cons. online 27/3/2010).
- ZHANG W., JOHNSON T.J., SELTZER T., BICHARD S.L. (2010) "The Revolution Will be Networked: The Influence of Social Networking Sites on Political Attitudes and Behavior", in *Social Science Computer Review*, vol. 28, no. 1 (Feb.), pp. 75-92.

4.

FRAMMENTI DI FANTASCIENZA ONLINE E L'IMMAGINAZIONE DEL FUTURO

4.1. INTRODUZIONE

Fino a tempi molto recenti, salvo qualche rara eccezione, la funzione fondamentale dell'immaginario, inteso come estensione collettiva dell'immaginazione, è stata quella di rivolgere lo sguardo al passato cercando di rinvenire nel mito una verità da porre a fondamento del sacro e quindi del sociale. Per l'uomo pre-moderno credere nel fantastico era un dovere. L'immaginario fecondava il mondo e lo nutriva con la sua "verità". Per questo il fantastico si pone in una dimensione obliqua rispetto al "reale" (Campra, 2000, p. 16).

È solo a partire dagli inizi del XX secolo che si è sviluppata una nozione del fantastico come qualcosa in cui credere per il puro piacere di credere. Le ragioni per le quali questo è accaduto sono varie e sono connesse con l'espansione economica, la rivoluzione industriale, l'emergere delle classi subalterne e, più in generale, con quello che chiamiamo l'avvento della modernità.

Qui ci limitiamo a considerare uno solo dei nuovi genere letterari che sono apparsi in questo frangente storico: la fantascienza. D'altra parte è proprio nella ibridazione tra fantasia e scienza che si scorgono gli scenari del futuro. Tuttavia non possiamo dimenticare che quello che è considerato il primo romanzo di fantascienza, *Frankenstein or the Modern Prometheus* (1818) di Mary Wollstonecraft Shelley, è il prodotto di un clima culturale prettamente romantico che trova la sua massima espressione popolare nel romanzo gotico con il suo potere di fascinazione centrato sull'orrido e sull'ultraterreno (Giovannini e Micangeli, 1998, p. 12). Non tutti concordano, ovviamente con questa "datazione" (Sadoul, 1975, p. 20). Tuttavia, quali che siano gli autori presi come riferimento (Edgar Allan Poe, Jules Verne, Rider Haggard, Robert Louis Stevenson, Edward Bellamy, Herbert G. Wells e Jack London), è costante il punto di

vista critico con il quale questo genere letterario nell'immaginare il futuro si dispone nei confronti della contemporaneità. Per essere consapevolmente "contemporanei" è necessario conservare una certa distanza dal proprio tempo pur senza collocarsi al di fuori di esso: "Un uomo intelligente può odiare il suo tempo, ma sa in ogni caso di appartenergli irrevocabilmente, sa di non poter sfuggire al suo tempo. (...) Può dirsi contemporaneo soltanto chi non si lascia accecare dalle luci del secolo e riesce a scorgere in essere la parte dell'ombra, la loro intima oscurità" (Agamben, 2008, pp. 9 e 14).

La fantascienza si è presentata, già ai suoi esordi, come qualcosa di completamente diverso rispetto alla letteratura utopistica del Rinascimento; in essa non si esprimeva un ideale di natura o di nuovo umanesimo, quanto piuttosto ciò che della natura appariva come frutto di una trasformazione operata dal capitalismo nella sua fase più espansiva: "il tipo di letteratura che meglio esprimeva la mediazione fra cultura e natura messa in atto dalla società industriale" (Caronia, 2009, p. 6).

L'attenzione della sociologia verso questo fenomeno socio-culturale avviene significativamente proprio per questa sua particolare vocazione a percepire il buio del presente e a proiettarlo nelle anticipazioni del futuro. Così la fantascienza viene assunta da alcuni ricercatori come oggetto di studio autonomo, in grado di rappresentare modelli verosimili nei quali poter osservare le tendenze di sviluppo della società industriale e delle sue relazioni sociali (Lemieux, 1978; Clarke, 1979; Martin, 1980) oppure come vero e proprio "laboratorio sociale" (Woodcock, 1979).

In queste analisi si mette in evidenza come già a partire dalla fine degli anni '30, e poi in modo più incisivo negli anni '50 (Parrinder, 1977; Feenberg, 1978), si affaccino nella fantascienza elementi di critica sociale o di vero e proprio "catastrofismo". Il cinema, da questo punto di vista, ha popolarizzato questa tendenza distopica più di quanto non abbia fatto la letteratura stessa; si pensi, tanto per fare un esempio, a un film come *Metropolis* di Fritz Lang, del 1926, che rappresenta ancora oggi il modello ispiratore della fantascienza tecno-sociologica. Naturalmente queste narrazioni fantapolitiche anti-utopiche, che mettevano in guardia i lettori dalle degenerazioni di una società minacciata da totalitarismi e applicazioni scientifiche spregiudicate, erano già ben presenti in autori come Herbert G. Wells (*The Time Machine*, 1895), Jack London (*The Iron Heel*, 1907), Karel Čapek (*R.U.R.*, 1920), Evgénij Zamjatin (*My*, 1922), Aldous Huxley (*Brave New World*, 1932).

L'immaginario dalla metà degli anni '80 in poi diventa un tema di ricerca completamente organico con lo sviluppo della "società della comunicazione". Oggi questa locuzione appare scontata e banale. A essa si preferisce quella di "società della conoscenza" più completa e adeguata per rappresentare l'intreccio tra

informazione, sistemi organizzativi complessi, gestione dei dati e modelli decisionali. Ma non era così banale quando per la prima volta ci si rese conto che il “concetto di comunicazione di massa” era – per così dire – restrittivo e appariva necessario generalizzare la transizione dalla società della scrittura a quella allora non completamente realizzata della digitalizzazione. Nel 1987 l'immaginario è il tema prescelto da George Balandier per il n. 40 dei *Cahier internationaux de sociologie*. Il volume monografico porta il titolo: *Nouvelle images nouvelle réel*. In esso intervengono autori come Raymond Ledrut, Philippe Quéau, Jean Baudrillard e Abraham Moles. È soprattutto Raymond Ledrut nel suo saggio a mettere in rilievo il rapporto tra società reale e società immaginaria (o immaginata): la società reale non è reale senza il suo tessuto di immagini che la rendono viva (Ledrut, 1987, p. 45). Che ne sarebbe della società senza la sua rappresentazione?

Le immagini sono i prodotti, più o meno coscienti, degli immaginari dei loro autori; questi immaginari, che ricordano con ogni evidenza l'*imaginaire sociale* di Cornelius Castoriadis de *L'institution imaginaire de la société* (Castoriadis, 1975) benché esso non compaia nei riferimenti bibliografici di Ledrut, sono forme archetipiche che hanno una funzione fondamentale nella vita sociale perché la strutturano e ne organizzano il senso. Questo non significa, tuttavia, che tutti gli immaginari siano o diano vita a processi di “realizzazione”. La realizzazione esige l'azione e non tutti gli immaginari si realizzano. L'immaginario sociale può essere ricondotto a delle realtà possibili che non si sono realizzate (utopie, speranze, attese) o a realtà possibili che non si sono ancora realizzate. L'immaginario sociale ha un passato e un futuro come le persone e come le persone è sottoposto a disordini e confusioni (*troubles*). Il passato non accetta facilmente di essere “passato” e il futuro spesso si presenta come una forma di evasione, di fuga dal reale (Ledrut, 1987, p. 50). Vi è quindi una dialettica tra reale e immaginario. In questa dialettica l'immaginario è duplice: esprime la realtà del possibile e in questo senso ha una concretezza; però il possibile non è necessariamente reale, pertanto la realtà del possibile è anche una simulazione. L'immaginario è un legame tra il reale e l'irreale.

Nello stesso numero monografico Jean Baudrillard riprende il tema a lui caro della seduzione diabolica delle immagini tecnologiche e della perversione che esse stabiliscono con il referente fino a distruggere il reale e a sostituirsi ad esso (1979). L'immagine precede il reale ed esercita un fascino autonomo imponendo le proprie regole, la propria logica, la propria immoralità. L'immagine, anziché porsi come rappresentazione del reale, lo contamina, lo corrompe, lo deforma (Baudrillard, 1987, p. 141). L'immagine si pone al di là del vero e del falso, del bene e del male, fino a produrre un'implosione, un corto circuito con il reale e una negazione di senso.

Gli anni '80 rappresentano una linea di demarcazione negli studi sulla immaginazione del futuro attraverso i mondi finzionali della *science fantasy*. Ciò che conta per il sociologo non è che la realtà analizzata sia autentica o immaginaria quanto piuttosto che essa risulti credibile nella presentazione delle regolarità socio-politiche individuate. In un romanzo di fantascienza l'autore descrive in dettaglio il mondo immaginario nel quale vivono i suoi personaggi. In esso si esprimono valori, scelte, aspettative, tracce di futuri possibili che appartengono ad uno specifico genere letterario ma che sono anche espressione del clima culturale nel quale e del quale il mondo finzionale si nutre. Il testo genera un "modello di realtà" che può essere esplorato e studiato come se fosse un mondo reale (Moles, 1984; Bainbridge, 1986; Thomas, 1990; Rosalind, 1990).

Negli anni più recenti questa chiave di lettura si arricchisce con gli sviluppi fortemente "sociologici" di quella che gli studiosi di fantascienza già negli anni '60 avevano chiamato l'esplorazione dell'*inner space*, che vedeva in Philip K. Dick, Kurt Vonnegut e James G. Ballard i maggiori rappresentanti. Aniché narrare l'esplorazione dello spazio profondo questi autori preferivano registrare la trasformazione della società in senso postfordista, la crisi di un modello di sviluppo prometeico e l'avvento di scenari indesiderabili legati alla rottura dell'equilibrio ecologico, alla prefigurazione di tecnologie invasive di sorveglianza e dominio della mente, di ibridazione tra corpi e macchine (Caronia, 2009, p. 7). Su questo piano, con una particolare focalizzazione sugli spazi urbani, sulle tematiche di esclusione e inclusione e sulla polarizzazione tra tecnologia e imbarbarimento regressivo, l'attenzione dei ricercatori si focalizza sul filone cyberpunk (Burrows, 1997), sulle derive delle società telematiche (Bogard, 1996) e sugli scenari post-apocalittici che proiettano nel futuro sogni e ansietà del presente (Retzinger, 2008). La fantascienza come genere letterario appare quasi al tramonto, eppure vediamo quanto essa abbia influenzato la fiction in tutte le sue manifestazioni: cinema, televisione, pubblicità, videogiochi e tutto ciò che è apparso in questi ultimi anni come ibridazione delle tecnologie di comunicazione. "Oggi la fantascienza è il solo luogo dove sopravvive il futuro" (Ballard, 2009, p.162).

4.2. LA TECNICA DI ANALISI

Che cosa agisce, dunque, in noi (sogni, aspettative, paure) come immaginario sociale fino a condizionare le nostre scelte, le nostre congetture e ciò che del passato noi stessi stiamo scagliando al di là della soglia del presente?

Nel lavoro che qui si presenta ci si muove nella direzione di una meta-

analisi che assume come oggetto di ricerca ciò che gli utenti della Rete ritengono significativo per rappresentare la propria immagine del futuro. In questo lavoro YouTube diventa una sorta di “magazzino ideale” di questa dimensione a un tempo personale ma anche condivisa, soggettiva e collettiva, attraverso la quale i “prosumer” di immagini selezionano frammenti di fantascienza e di futuro dall’immenso universo della Rete per riprodurli all’infinito dei click di tutti coloro che volontariamente scelgono di far parte di questa rete di relazioni immaginali.

La riflessione su tali immagini, scene o selezioni di frammenti di film si inserisce all’interno della sociologia sulle immagini (Faccioli e Losacco, 2006), che ci aiuta a interpretare e spiegare cosa si celi dietro o dentro di esse allo scopo di cogliere con esattezza il senso del testo: “Interpretare (...) non significa soltanto dispiegare un’attenzione caparbia per l’oggetto, ma anche interagire esplicitamente con esso; non semplicemente passare in rassegna, ma anche riattivare, ascoltare, dialogare” (Casetti e di Chio, 1990, p. 13).

La selezione delle scene, dei trailer o dei montaggi proposti dai prosumer è avvenuta attraverso l’inserimento su YouTube di parole chiave, quali: fantascienza, *science fiction*, *science fiction movies* e *future predictions*. È ovvio che questo tipo di “esca” porti alla superficie un *mare magnum* di informazioni. Su questo primo filtro di ricerca ne va, allora, innestato un secondo: dettato dalla pertinenza delle immagini rinvenute con l’obiettivo proprio del ricercatore. I materiali così selezionati possono essere analizzati applicando a essi la tecnica della *content analysis*; secondo quanto suggerisce anche Gillian Rose (2007) in merito all’utilizzo di tale tecnica per interpretare la rappresentazione degli aborigeni sulle riviste di settore. È la stessa autrice a sottolineare che una selezione delle immagini non costituisce una distorsione nella creazione del dataset perché comunque all’interno di esso ricadranno elementi rappresentativi di una data visione. Il nostro intento coincide con quanto sostiene la Rose, poiché nostro obiettivo è quello di individuare – mediante questi frammenti – quale sia la costruzione odierna del futuro e quindi come quest’ultimo venga rappresentato dai prosumer su YouTube.

Successivamente, le scene selezionate secondo i parametri individuati mediante la costruzione e osservazione condotta con la scheda di analisi (Krippendorff, 1980; Losito, 1993), vengono elaborate utilizzando i CAQDAS (*Computer Qualitative Data Analysis Software*). L’utilizzo di questi pacchetti ci permette di esplicitare e porre in relazione i tratti odierni rappresentativi del futuro; in questo modo si può trasformare una conoscenza tacita (anche se visiva) in esplicita. Come già argomentato altrove (La Rocca, 2009), Polanyi in *La conoscenza personale* chiarisce che il “capire non è un atto né arbitrario né

un'esperienza passiva, ma un atto responsabile che aspira alla validità universale" (1958, trad. it. 1990, p. 70). Polanyi distingue fra conoscenza tacita ed esplicita, la prima definibile come personale e specifica di un dato contesto e pertanto difficilmente formalizzabile e comunicabile; la seconda invece codificata e trasmissibile mediante l'uso di un linguaggio formale e sistemico. Questa distinzione è fatta propria da due studiosi di dinamiche organizzative Ikujiro Nonaka e Hirotaka Takeuchi (1995). Sulla scia del pensiero epistemologico di Polanyi, il quale sostiene che gli esseri umani acquisiscono conoscenze organizzando le loro esperienze, i due autori mettono a fuoco il processo che porta a una conversione della conoscenza tacita in quella esplicita; nel nostro caso si tratta quindi di individuare ed esplicitare i tratti con i quali viene rappresentato il futuro. Le modalità di conversione della conoscenza, stando ai due autori, sono quattro: socializzazione, esteriorizzazione, combinazione, interiorizzazione (fig. 4.1).

La socializzazione è descritta come un processo di condivisione di esperienze e di creazione di forme di conoscenza tacita, come accade ad esempio all'apprendista che impara a lavorare il legno guardando lavorare il maestro. Nell'esteriorizzazione si ha il passaggio di una conoscenza tacita mediante concetti espliciti; per esempio quando mediante il linguaggio cerchiamo di esprimere una sensazione o un'immagine.

		a	
		Conoscenza tacita	Conoscenza esplicita
da	Conoscenza tacita	Socializzazione	Esteriorizzazione
	Conoscenza esplicita	Interiorizzazione	Combinazione

Fig. 4.1 – Quattro modalità di conversione di conoscenza
(fonte: Nonaka e Takeuchi, 1995; trad. it. 1997, p. 102)

Nella combinazione i concetti vengono organizzati in un sistema di conoscenza; le persone si scambiano continuamente informazioni di questo tipo mediante, per esempio, il passaggio di documenti o dati. Con l'interiorizzazione la conoscenza esplicita viene tradotta in conoscenza tacita. In quest'ultimo passaggio si coglie la circolarità del modello a spirale che i due autori hanno ipo-

tizzato per la conoscenza (Nonaka, Takeuchi 1995, trad. it. 1997, pp. 103-114); per cui mediante l'interiorizzazione le esperienze — maturate attraverso la socializzazione, l'esteriorizzazione e la combinazione — vengono fatte proprie dagli individui che le hanno esperite.

		a	
		Conoscenza tacita	Conoscenza esplicita
da	Conoscenza tacita	Socializzazione	Decostruzione
	Conoscenza esplicita	Conferma	Costruzione

Fig. 4.2 – Le quattro fasi della creazione della conoscenza etnografica
(fonte: La Rocca, 2009, p. 147)

Il processo di conversione della conoscenza, presentato in questi passaggi, permette di estendere la riflessione alla metodologia di analisi e interpretazione qui utilizzata. Si opera qui, in prima battuta, una socializzazione del lavoro svolto sulle immagini, cui segue una decostruzione del testo. La sussunzione all'interno di categorie dei concetti etichettati nella fase precedente caratterizza la costruzione. Infine, avviene la conferma di quanto trovato; a questo punto il testo che si viene a creare può essere – non appena lo si è trascritto – oggetto di una nuova analisi.

L'uso dei CAQDAS che qui si propone, e in special modo di Atlas.ti5, per l'analisi delle immagini consente il passaggio da una conoscenza tacita a una esplicita, organizzando e ponendo in relazione le informazioni da noi raccolte.

4.3. LE CARATTERISTICHE DEL FUTURO

Il futuro nel passato è a tinte fosche, ma ciò che più si nota è una sorta di preparazione al futuro: siamo all'esterno di una possibile navicella spaziale, che è ancora in via di costruzione, la proiezione verso il futuro è una tensione verso un viaggio non ancora compiuto che ha come meta sicuramente il cielo ma non ancora lo spazio. Al cielo si anela perché sulla terra dilaga la guerra: il futuro è senz'altro la fine di quanto si conosce.

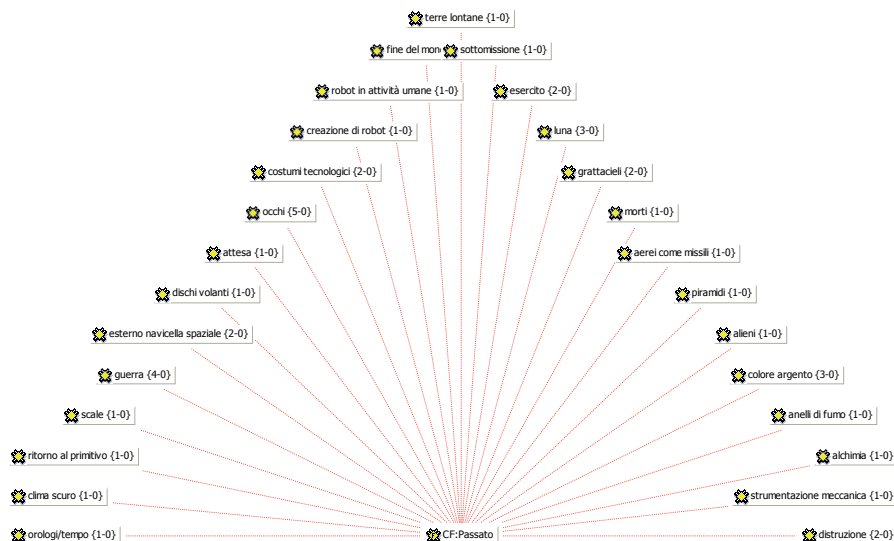


Fig. 4.3 – Network dei codici per “passato”

Nel gruppo di immagini che definiscono il futuro nel presente si è, per così dire, fatto un passo avanti: il futuro è nell’universo, c’è quindi una tensione verso l’ignoto. Se si guarda al cielo non si pensa più solo alla luna ma a tutti i pianeti – che già si conoscono – e alle loro possibili forme di vita.





Fig. 4.4 – Immagini tipo per le unità d'analisi (passato, presente e futuro)

L'uomo è “sopravvissuto” alla distruzione/guerra del futuro passato. Dopo essere stati sulla luna e aver guardato all'universo e ai suoi pianeti, nelle predizioni future il futuro è sulla terra e nell'acqua, nell'integrazione tra natura e tecnologia. Il futuro è luminoso.

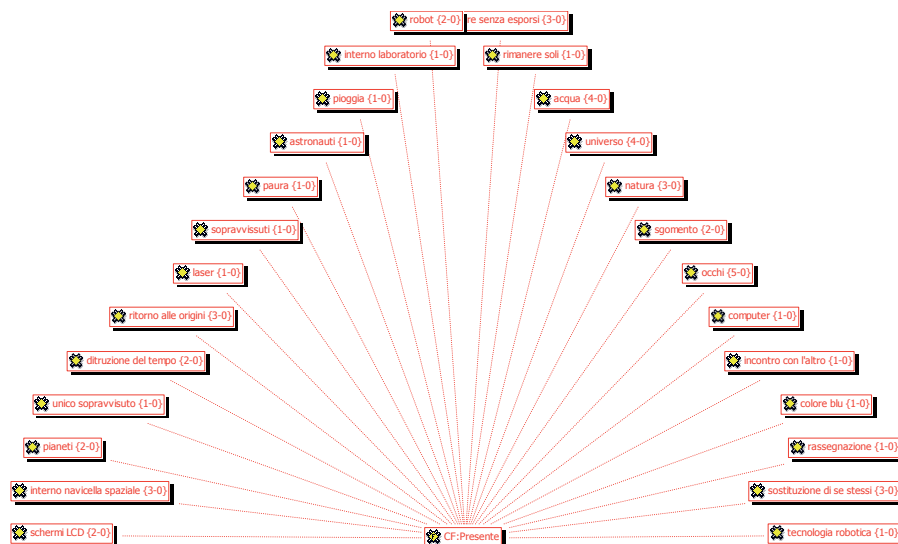


Fig. 4.5 – Network dei codici per “presente”

4.4. LA COSTRUZIONE DI UNA REALTÀ

William Isaac Thomas scrisse “se gli uomini definiscono le situazioni come reali, esse saranno comunque reali nelle conseguenze”. Parafrasando: se il futuro è rappresentato a tinte fosche esso nel suo divenire sarà tale? Parimenti: se lo si immagina luminoso e vivibile sarà questa la forma che assumerà? L’unica certezza è il riconoscere che quello appena formulato è, per adesso, un quesito senza risposta, o meglio se risposta c’è essa è affidata “ai posteri”.

Interrogarsi sul futuro è, però, una delle attività insite nella natura umana: è la leva che ci spinge a implementare la conoscenza. Le rappresentazioni dei futuri qui proposte riassumono le tensioni, che nella loro dimensione presente, si trovano a vivere gli uomini; per esempio: nelle unità di immagini “passato” l’altrove immaginato è la luna; nel “presente” sono i pianeti e l’universo; nelle “predizioni future” il regno scelto è la terra.

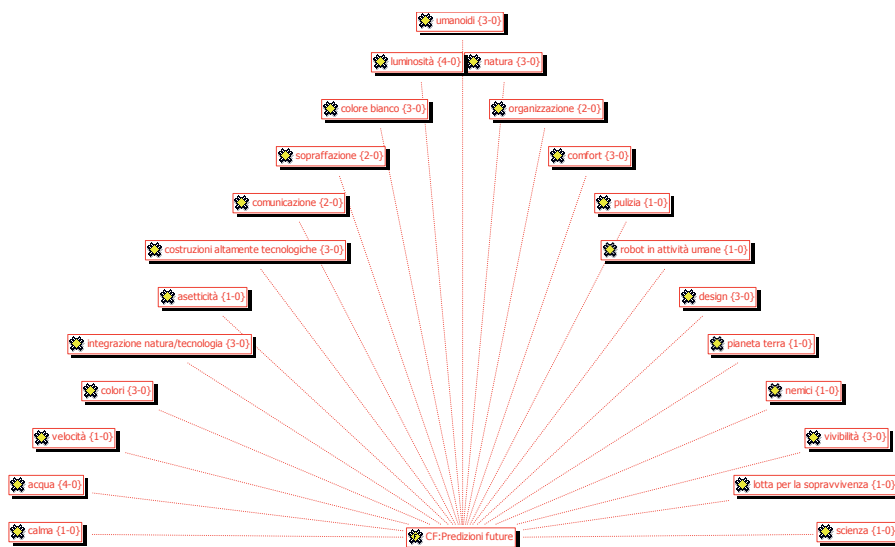


Fig. 4.6 – Network dei codici per “predizioni future”

La terra – nel gruppo di immagini “predizioni future” – è però qualificata sotto una nuova veste, si ipotizza un suo restyling: è luminosa, si è raggiunta un’integrazione fra la natura e l’uomo, è vivibile, confortevole, i tempi scorrono calmi ma anche veloci. Si tratta del pianeta terra, ma di certo non come noi oggi lo conosciamo (cfr. fig. 4.7).

4.5. (CHI)OSARE

Il lavoro qui condotto è consistito nell’osservare la rappresentazione del futuro così per come gli utenti di YouTube l’hanno estrapolata, riorganizzata e fatta circolare in Rete, operando un’ulteriore selezione sulla rappresentazione del futuro creata e veicolata da un mass media, quale il cinema. In questo processo di osservazione e riproduzione, noi saremmo il terzo osservatore. In linea con quanto fatto prima dagli autori cinematografici e poi dai prosumer, qui si sarebbe prodotta un’osservazione di secondo livello. La rappresentazione della realtà futura viene elaborata attraverso un’attribuzione di senso a essa: creandole un frame.

Prima di estendere gli obiettivi e le conseguenze del nostro percorso di lavoro, riprendiamo le linee argomentative fornite da Niklas Luhmann (2002) in merito alla realtà prodotta dai mass media. L’autore sostiene che i mass media, in

termini kantiani, generino un'illusione trascendentale.

“In questa accezione l'attività dei mass media non viene considerata semplicemente una sequenza di operazioni ma una sequenza di osservazioni, o più precisamente di operazioni che osservano. Per arrivare a questa accezione di mass media dobbiamo quindi osservare il loro osservare. Per la prima accezione di cui abbiamo parlato è sufficiente un'accezione di primo ordine, come se si trattasse di fatti. Per la seconda possibile accezione bisogna invece assumere la posizione di un osservatore di secondo ordine, di un osservatore di osservatori” (2002, p. 18)”.

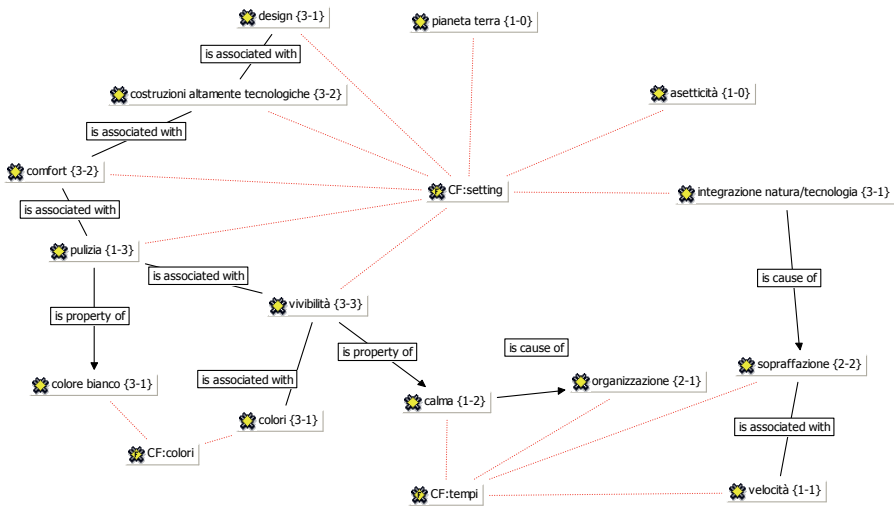


Fig. 4.7 – Il futuro sulla terra e le sue caratteristiche

Stando a quanto testé citato, i mass media devono, in quanto sistemi che osservano, essere in grado di distinguere tra autoreferenza ed eteroreferenza, ovvero sono chiamati a costruire a un tempo, una realtà e, parimenti, produrne un'altra che sia distinta dalla propria. Il meccanismo secondo il quale si procede è ispirato a Luhmann da Spencer Brown e dalla sua *Laws of Form* (1979), nella quale Spencer Brown parla della re-entry definendola come “l'operazione limite di un calcolo che rimane sul piano dell'osservazione di primo ordine e nell'ambito di distinzioni bivalenti”. Quindi se i mass media costruiscono la realtà e a loro volta devono essere in grado di tracciare una linea di demarcazione fra autoreferenza ed eteroreferenza, ciò comporta che il sistema (inteso

come mass media) deve essere in grado di distinguere se stesso da ciò che osserva. Per farlo è necessario avere uno schema che sia costruito in modo da poter generare prima una differenza (in termini di sistema e ambiente) e poi ancora una distinzione.

Nelle parole di Luhmann “il sistema, allora, deve in primo luogo operare e proseguire nelle proprie operazioni (deve cioè ad esempio poter vivere e comunicare), prima di poter utilizzare interamente la differenza così generata come distinzione, e quindi come schema per le proprie osservazioni” (2002, p. 25).

Sebbene il focus in Luhmann sia la costruzione della realtà operata dai mass media, il suo percorso, che lo porta a identificare come necessario uno schema che si differenzi dall’oggetto osservato per poter poi stabilire quanto da esso si distingua, è lo strumento che si è cercato qui di realizzare.

STRATEGIA DI ANALISI

STEP 1 – SELEZIONARE LE IMMAGINI

Il primo processo applicato ai video è stata l’analisi del contenuto sulle immagini. La scomposizione e codifica dei frammenti è stata eseguita utilizzando una scheda d’analisi, intesa come una serie ordinata di voci alle quali siamo chiamati a rispondere visionando i materiali video. Le voci della scheda sono da intendersi come delle domande cui occorre trovare risposta al fine di decodificare il testo. Le aree di osservazione hanno riguardato:

- il setting,
- i personaggi,
- i colori,
- i tempi.

In questo modo si è cercato di tirar fuori, dalla prima selezione, delle immagini “piene” di significato per l’analisi in oggetto. Sono state scartate quelle scene, che seppur emerse dalla prima ricerca poco o nulla avevano a che fare con l’immagine di un possibile futuro.

STEP 2 – ORGANIZZARE E CODIFICARE LE IMMAGINI

Dopo aver proceduto manualmente a scaricare le immagini presenti su YouTube è necessario:

- 1) organizzare il testo per variabili,
- 2) procedere con la codifica.

Operazione 1

I materiali raccolti sono stati organizzati utilizzando tre chiavi: passato, presente e predizioni future. Si tratta pur sempre di un passato del futuro, di un presente del futuro e di un futuro futuribile. Nel gruppo “passato” sono stati raccolti i frammenti video che si riferiscono a un futuro che ricade nell’arco temporale che va dal 1900 al 1999; nel “presente” quelli riferibili dal 2000 a oggi (2010) e nelle predizioni future tutti quelli che disegnano un futuro oltre il 2010. In totale si tratta di 114,51 minuti di video; ripartiti come indicato in tabella 4.1.

Tab. 4.1 – Numerosità dei video e dei minuti per unità di analisi

	N. video	Minuti
Passato	6	41,36
Presente	7	25,33
Predizioni future	6	47,82
Totale	19	114,51

La sistematizzazione dei materiali secondo questa *timeline*, consente di organizzare il contenuto delle immagini del nostro dataset mediante una scansione temporale convenzionale.

Per far riconoscere al software che si tratta di tre documenti distinti è necessario dopo averli caricati assegnare loro le variabili nominali, secondo la procedura consentita dal **Primary Doc Family Manager**. Si avranno così tre documenti:

- P1: Passato,
- P2: Presente,
- P3: Predizioni future.

Operazione 2

Caricate le immagini è ora possibile procedere all’operazione di codifica, ovvero all’applicazione di un’etichetta. In questo caso si è scelto di operare una codifica dal basso, ovvero i codici non sono già stati individuati a priori bensì emergeranno man mano che si procederà alla lettura del testo.

Si ricorda che quella effettuata è un’analisi sulle immagini che non tiene conto del verbale contenuto nei frammenti.

Si è poi proceduto a una codifica, dove per codifica si intende l’applicazione di un’etichetta, un codice appunto, alle immagini.

STEP 3 – LAVORARE SUI CODICI

Per poter delineare i tratti delle immagini presenti in Rete è necessario ritornare a lavorare sui codici. Riportare direttamente all'interno del rapporto di ricerca o di un articolo la tabella riepilogativa dei codici (cfr. tab. 4.2) sarebbe infatti dispendioso e poco utile e ciò essenzialmente per due motivi: il primo estetico e il secondo tecnico-operativo. Si tratta, infatti, di materiali che di solito si possono inserire come allegati ma che posti all'interno del documento ne appesantiscono la lettura. Inoltre, può accadere che durante la fase di codifica si creino due codici simili, oppure due codici con il medesimo significante per i quali è opportuno operare una fusione. Allora occorre:

- 1) verificare la “qualità” dei codici creati;
- 2) riorganizzarli in macro categorie.

Tab. 4.2 – Elenco dei codici per i topic

	Passato	Presente	Predizioni Future	Totale
Acqua	-	3	1	4
Aerei come missili	1	-	-	1
Alchimia	1	-	-	1
Alieni	1	-	-	1
Anelli di fumo	1	-	-	1
Asetticità	-	1	-	1
Astronauti	-	-	1	1
Attesa	1	-	-	1
Calma	-	1	-	1
Clima scuro	1	-	-	1
Colore argento	3	-	-	3
Colore bianco	-	3	-	3
Colore blu	-	-	1	1
Colori	-	3	-	3
Comfort	-	3	-	3

Computer	-	-	1	1
Comunicazione	-	2	-	2
Costruzioni altamente tecnologiche	-	3	-	3
Costumi tecnologici	2	-	-	2
Creazione di robot	1	-	-	1
Design	-	3	-	3
Dischi volanti	1	-	-	1
Distruzione	2	-	-	2
Distruzione del tempo	-	-	2	2
Esercito	2	-	-	2
Esterno navicella spaziale	2	-	-	2
Fine del mondo	2	-	-	2
Grattacieli	2	-	-	2
Guerra	4	-	-	4
Incontro con l'altro	-	-	1	1
Integrazione natura/tecnologia	-	3	-	3
Interno laboratorio	-	-	1	1
Interno navicella spaziale	-	-	3	3
Laser	-	-	1	1
Lotta per la sopravvivenza	-	1	-	1
Luminosità	-	4	-	4
Luna	3	-	-	3
Morti	1	-	-	1
Natura	-	2	1	3
Nemici	-	1	-	1
Occhi	2	-	3	5

Organizzazione	-	2	-	2
Orologi/tempo	1	-	-	1
Paura	-	-	1	1
Pianeta terra	-	1	-	1
Pianeti	-	-	2	2
Pioggia	-	-	1	1
Piramidi	1	-	-	1
Pulizia	-	1	-	1
Rassegnazione	-	-	1	1
Rimanere soli	-	-	1	1
Ritorno al primitivo	1	-	-	1
Ritorno alle origini	-	-	3	3
Robot	-	-	2	2
Robot in attività umane	-	1	-	1
Scale	1	-	-	1
Schermi LCD	-	-	2	2
Scienza	-	1	-	1
Sgomento	-	-	2	2
Sopraffazione	-	2	-	2
Sopravvissuti	-	-	1	1
Sostituzione di se stessi	-	-	3	3
Sottomissione	1	-	-	1
Strumentazione meccanica	1	-	-	1
Tecnologia robotica	-	-	1	1
Terre lontane	1	-	-	1
Umanoidi	-	3	-	3
Unico sopravvissuto	-	-	1	1

Universo	-	-	4	4
Velocità	-	1	-	1
Vivere senza esporsi	-	-	3	3
Vivibilità	-	3	-	3
Totale	40	48	44	132

Operazione 1

Mediante **Codes/Output/Codes-Primary-Document-Table/Standard Report** si recupera l'informazione inerente a quanti codici creati e alla loro distribuzione all'interno delle unità di analisi create.

Una prima riflessione sulla tabella permette di notare che ciascuna unità contiene elementi specifici in merito all'idea del futuro. Su questa informazione è possibile strutturare delle riflessioni.

I codici prodotti dall'analisi sulle immagini sono in totale 132, di questi 40 sono serviti per scomporre l'unità definita come "passato", 44 sono quelli necessari a etichettare il "presente" e 48 quelli attribuiti ai frammenti video selezionati mediante la chiave "predizioni future".

L'esplosione dei codici per unità di frammenti (figg. 4.3, 4.5, 4.6) ci permette di identificare dei primi tratti salienti per ciascun insieme di immagini.

Operazione 2

Continuando con le procedure di analisi si lavora sul piano concettuale creando delle categorie più ampie o famiglie di codici all'interno delle quali inserire gli elementi fino a ora trovati.

Questa operazione di sussunzione è anch'essa un'operazione concettuale e manuale che il ricercatore svolge in prima persona. Infatti, si riaggregano i codici in contenitori di significato più ampi che poi verranno posti in relazione. Da **Codes/Edit Families/Open Family Manger** si realizza tecnicamente questa operazione di imputazione. La fase operativa è però preceduta da un momento di riflessione in cui si individuano i contorni delle macro categorie e si definiscono i codici da introdurre.

In questo lavoro le aree o macro categorie in base alle quali raggruppare i diversi codici sono le medesime utilizzate nella creazione della scheda per l'analisi del contenuto manuale. Ovvero, se in un primo momento: il setting, la presenza di personaggi, i colori, i tempi sono stati dei parametri in base ai quali considerare delle scene come "vuote" o "piene" di significato per la nostra a-

nalisi adesso diventano delle categorie che vanno riempite di contenuti.

Quindi:

- all'interno di setting sono stati inseriti i codici che descrivono l'ambientazione presentata (la guerra, i pianeti, le navicelle spaziali);
- nella macro categoria personaggi sono stati raggruppati i protagonisti delle scene, siano essi umani, robot o umanoidi;
- all'interno della macro categoria colori sono stati inseriti tutti quei codici inerenti la descrizione che il ricercatore ha prodotto dei colori presenti nelle varie scene;
- all'interno della macro categoria tempi sono stati raggruppati i codici creati per delineare l'uso e la presenza del tempo all'interno dei frammenti video.

STEP 4 – RIASSUMERE E RAPPRESENTARE L'INFORMAZIONE

La procedura di lettura delle immagini e la successiva codifica ci conducono a un livello elevato di analisi e conoscenza dei legami presenti nei testi. È indubbio che guardando, codificando, riaggregando, estraendo informazioni di dettaglio il livello di penetrazione nel testo e del testo cresce nel ricercatore fino a consentirgli di porre in essere relazioni fra i dati, cui prima non aveva pensato. Tali informazioni vanno esplicitate sia in forma testuale ma anche grafica, ovvero occorre riassumere le connessioni individuate e renderle fruibili ai lettori. Per fare ciò si adoperano i network.

Nell'introduzione a questo capitolo si era detto che obiettivo di questo lavoro era:

- ricostruire l'idea che del futuro viene veicolata in Rete dai prosumer.
- La funzione del network presentato in figura 4.7 è quella di rispondere a questi quesiti iniziali. Le relazioni poste in essere fra le varie famiglie di codici esplicitano questo rapporto.

Infatti si ha che:

- la relazione “*is cause of*” è di tipo transitivo e pertanto è applicata a codici che hanno una relazione di tale tipo fra di loro e con il setting selezionato: il pianeta terra;
- la relazione “*is a property of*” è di tipo asimmetrico e quindi nell'attribuirle bisogna evidenziare che c'è un *prius* nella generazione di tale legame. Per esempio, qui si è sottolineato mediante essa che la pulizia è rappresentata dal colore bianco;
- la relazione “*is associated with*” è di tipo simmetrico e si applica ai codici che hanno tale legame fra di loro.

RIFERIMENTI BIBLIOGRAFICI

- AGAMBEN G. (2008) *Che cos'è il contemporaneo*, Roma, Nottetempo.
- BAINBRIDGE W.S. (1986) *Dimensions of Science Fiction*, Cambridge, Harvard University Press.
- BALLARD J.G. (2009) *I miracoli della vita*, Milano, Feltrinelli.
- BAUDRILLARD J. (1979) *Simulacre et simulations*, Paris, Galilée.
- BAUDRILLARD J. (1987) "Au-delà du vrai ed tu faux, ou le malin génie de l'image", in *Cahiers Internationaux de Sociologie*, vol. 34, n. 82, pp. 139-145.
- BOGARD W. (1996) *The Simulation of Surveillance: Hypercontrol in Telematic Societies*, Cambridge: Cambridge Un. Press.
- BURROWS R. (1997) "Virtual Culture, Urban Social Polarisation and Social Science Fiction", in R. Burrows, *The Governance of Cyberspace*, London, Routledge, pp. 38-45.
- CAMPRA R. (2000) *Territori della finzione. Il fantastico in letteratura*, Roma, Carocci.
- CARDANO M. (2003) *Tecniche di ricerca qualitativa*, Rom.: Carocci.
- CARONIA A. (2009), *Universi Quasi Paralleli. Dalla fantascienza alla guerriglia mediatica*, Roma, Cut-up edizioni.
- CASSETTI F., DI CHIO F. (1990) *Analisi dei film*, Milano, Bompiani.
- CASTORIADIS C. (1995) *L'istituzione immaginaria della società*, Torino, Bollati Boringhieri.
- CLARKE I. F. (1979) *The Pattern of Expectation: 1644-2001*, New York, Basic Books.
- COME A.B. (1995) *Kierkegaard as Humanist. Discovering My Self*. Montreal, Buffalo, McGill-Queen's University Press.
- FACCIOLI P., LOSACCO G. (2006) *Manuale di sociologia visuale*, Milano, FrancoAngeli.
- FEENBERG A. (1978) "The Politics of Survival: Science Fiction in the Nuclear Age", in *Alternative Futures*, vol. 1, n. 2, pp. 3-23.
- GIOVANNINI F., MINICANGELI M. (1998) *Storia del romanzo di fantascienza* Roma, Castelvecchi.
- GOBO G. (2001) *Descrivere il mondo. Teoria e pratica del metodo etnografico in sociologia*, Roma, Carocci.
- HALTSTEAD M.N. (1976) "Social Science Fiction and the Sociological imagination", in *The Cornell Journal of Social Relations*, vol. 11, n. 1, pp. 59-64.
- KRIPPENDORFF K. (1980) *Content Analysis: An introduction to its methology*, Beverly Hills, CA, Sage.
- LA ROCCA G. (2009) "Organizzare l'informazione etnografica. Il ruolo dei CAQDAS", in *Rassegna Italiana di Sociologia*, anno L – n.1/2009, pp. 133-160.
- LEDROUT R. (1987) "Société réelle et société imaginaire", in *Cahiers Internationaux de Sociologie*, vol. 34, n. 82, pp. 41-56.
- LEMIEUX J. (1978) "Determinismes du present et conceptions du futur: rapports de classes et relations ethniques dans la science-fiction américaine et française", in *Communication et Information*, vol. 2, n. 3, pp. 171-177.
- LIVINGSTON D. (1971) "Science Fiction Models of Future World Order Systems", in *In-*

- ternational Organization, vol. 25, no. 2 (Spring), pp. 254-270.
- LOSITO G. (1993) *L'analisi del contenuto nella ricerca sociale*, Milano, FrancoAngeli.
- LUHMANN N. (2002) *La realtà dei mass media*, Milano: FrancoAngeli.
- MARTIN W.C.(1980) "Science Fiction and Futures Research as Cultural Production and Reproduction: The Problematic of Knowledge", in *Capitalist Society, Quarterly Journal of Ideology*, vol. 4, n. 2, pp. 15-23.
- MOLES A.A. (1984), "Utopie, science-fiction, oeuvre catastrophe et conditionnement social", in *Degres*, vol. 12, n. 39-40, pp. I-113.
- NONAKA I., TAKEUCHI H. (1997), *The Knowledge-Creative Company. Creare le dinamiche dell'innovazione*, Milano, Guerini & Associati (ediz. orig. 1995).
- PARRINDER P. (1977) "The Black Wave: Science and Social Consciousness", in *Modern Science Fiction. Radical Science Journal*, vol. 5, pp. 37-61.
- POLANYI M. (1990) *La conoscenza personale. Verso una filosofia post-critica*, Milano, Rusconi (ediz. orig. 1958).
- POPPER K.R. (1983), *Conoscenza oggettiva. Un punto di vista evoluzionistico*, Roma, Armando Editore.
- POSTMAN N. (2003). *Come sopravvivere al futuro*, Milano, Orme editori.
- RETZINGER J.P. (2008) "Speculative Visions and Imaginary Meals: Food and the environment in (post-apocalyptic) science fiction films", in *Cultural Studies*, vol. 22, n. 3, pp. 369-390.
- ROSALIND W. (1990), *Notes on the Underground: an Essay on Technology, Society, and the Imagination*, Cambridge, MIT Press.
- ROSE G. (2007) *Visual Methodologies*, London, Sage.
- SADOUL J. (1975) *La storia della fantascienza*, Milano, Garzanti.
- SPENCER BROWN G. (1994), *Laws of Form*, London, Paperback (ediz. orig. 1979).
- SPERBER D. (1984) *Il sapere degli antropologi*, Milano, Feltrinelli (ediz. orig. 1982).
- SUVIN D.R. (1970) "De la tradition utopique dans la Science Fiction russe", in *Archives Internationales de Sociologie de la Cooperation et du Developpement*, vol. 27, pp.
- THOMAS L.-V. (1990) "Utopie, science-fiction et fantasmes", in *Recherches Sociologiques*, vol. 21, n. 1, pp. 5-24.
- VOLPI G. (2003) *Guida alla formazione di una cineteca*. Roma, Dino Audino Editore.
- WOODCOCK J. (1979) "Science Fiction and the Real Future", in *Alternative Futures*, vol. 2, n. 2, pp. 25-37.

5.

L'IMMAGINE DELLA CINA NELLA STAMPA QUOTIDIANA DAL 2000 AL 2008

5.1. INTRODUZIONE

Secondo alcuni osservatori, da qualche anno la Cina è particolarmente sensibile alla propria immagine sul piano internazionale. I Giochi Olimpici del 2008 sono stati presentati (e vissuti dai cinesi stessi della *middle class*) come un'occasione unica per accreditare definitivamente la nuova Cina Popolare come paese “normale”.

Dal 1972, con i primi contatti di Nixon e Kissinger, è iniziato un processo di avvicinamento della Cina ai paesi occidentali che ha trovato un punto di svolta nell'apertura dei primi fast food a Pechino nel 1987 (Yan, 2000) e poi successivamente di altri famosi marchi commerciali statunitensi ed europei.

“Immagine” è un parola di immediata comprensione, ma è anche un concetto ambiguo in cui confluiscono contenuti diversi che sono sollecitati da una commistione di esperienze dirette e indirette, stereotipi e pregiudizi. In alcuni casi l'immagine rimanda a qualcosa di esteriore e da non prendere troppo sul serio, come se ci fosse un'apparenza, l'immagine appunto, e poi una realtà soggiacente che attende di essere rivelata.

Con più rigore e precisione, Serge Moscovici ha introdotto negli anni '70 il concetto di “rappresentazione sociale” che si rifà alle brillanti intuizioni di Durkheim e Tarde degli inizi del Novecento (Farr e Moscovici, 1989). Ricostruire e analizzare le rappresentazioni sociali significa mettere a fuoco le idee del senso comune, le nostre convinzioni più profonde, che si formano sia nelle conversazioni quotidiane, nei luoghi di incontro e di lavoro, che nell'opinione pubblica, soprattutto con l'esposizione ai mass media. Le rappresentazioni sociali diventano costruzioni convenzionali della realtà, orientano il nostro modo di vedere il mondo e di operare nel mondo.

L'immagine della Cina e dell'Oriente è sempre stata controversa nella cultura occidentale. G. Rozman (2002) individua cinque fasi storiche principali che rappresentano modi diversi di osservare e interpretare la cultura cinese, soprattutto da parte degli europei:

- la fase di totale ammirazione che ha coinciso con le prime missioni dei gesuiti e con la nascita dell'Illuminismo, in cui la Cina era presentata come un modello da imitare (1600-1750);
- la fase di disincanto che è seguita subito dopo, ad opera degli illuministi "rivoluzionari" che vedevano la Cina come un impero corrotto e in decadenza (1750-1800);
- la fase in cui appare la sinologia come disciplina accademica e il romanticismo salda l'infatuazione per la Cina con le note resistenze all'affermarsi della modernizzazione (1800-1890);
- la fase coloniale-razzista in cui il darwinismo sociale diventa l'ideologia dominante anche come interpretazione scientifica dei rapporti con le culture non europee (1890-1945);
- la fase dei *China Studies* in cui l'ideologia colonialista, trapiantata negli Stati Uniti come modernizzazione capitalistica, si confronta con un corrispondente romanticismo politico di estrema sinistra che si afferma soprattutto tra i giovani in Europa (1945-1968).

In questi ultimi anni sembrerebbe che tutte queste "rappresentazioni" convivano insieme in una sorta di competizione tra "etichette" contrastanti:

- la Cina come mercato potenziale ad alto sviluppo economico e tale da attrarre forti investimenti finanziari;
- la Cina come mercato del lavoro a basso costo, senza le garanzie acquisite nelle società occidentali, e quindi vissuta come minaccia imminente per l'occupazione;
- la Cina come paese esportatore di prodotti non controllati e che non rispondono agli standard di qualità fissati dall'Unione Europea;
- la Cina come sistema politico autoritario, responsabile del massacro di Piazza Tiananmen e delle repressioni verso leader politici e religiosi (Tien e Nathan, 2001).

La ricerca intende verificare i contenuti e i confini di queste rappresentazioni in un paese come l'Italia che, negli ultimi anni, ha esercitato una forte attrazione economica e migratoria nei confronti della Cina. In generale l'immigrazione cinese in Europa (e l'Italia da questo punto di vista non rappresenta certo un'eccezione) è ritenuta da molti studiosi come caratterizzata da una forte componente di clandestini dei quali è molto difficile quantificare la consisten-

za e i percorsi (Pickie, 2004; Pedone, 2008). Questo alimenta l'immagine di una regia occulta che non ha equivalenti in altre correnti migratorie e ha conseguenze decisive sul modo in cui nell'opinione pubblica si costruisce il senso generale della presenza cinese nella nostra società.

5.2. IPOTESI DI LAVORO

L'ipotesi di fondo di questo lavoro è che la rappresentazione sociale della Cina sia fortemente condizionata da stereotipi e pregiudizi, indipendentemente dalla attendibilità o meno delle fonti di conoscenza dei problemi legati all'immigrazione clandestina o alla presenza sul mercato di prodotti illegali di importazione. L'immigrazione cinese in Italia non è tra le più consistenti in Europa: la stima ISTAT per il 2008 è di circa 170.000 cinesi residenti (demo.istat.it; Caritas e Migrantes, 2010), mentre per il Regno Unito e la Francia le stime superano le 500.000 unità (www.statistics.gov.uk; www.mondialisme.org).

Tuttavia l'immigrazione cinese viene presentata spesso come molto problematica. Ciò che colpisce è il rapido incremento in questi ultimi anni e la provenienza degli immigrati da province diverse da quelle della prima ondata (Zhejiang e Fujian). I cinesi immigrati non sono più i fuggiaschi da un paese emarginato, sebbene in cammino verso la modernizzazione, ma i rappresentanti più attivi di una potenza mondiale emergente. La Cina ha conquistato ormai una posizione di primo piano nel commercio mondiale. Lo stato della bilancia commerciale europea vede la Cina Popolare balzare ai primi posti tra i paesi fornitori dell'area euro. In generale, il suo peso economico e culturale è oggi tale da rimettere in discussione gli stessi equilibri della politica internazionale.

Nello specifico, i contenuti della rappresentazione sociale della Cina sono delineati attraverso l'immagine veicolata dalla stampa. Infatti la comunicazione giornalistica si presenta come una sorta di mediazione culturale, un filtro e un mezzo in grado di trasformare gli eventi in notizie così da essere poi fruite come rappresentazioni di esse dall'opinione pubblica (Bonazzi, 1998).

L'arco temporale preso in considerazione va dal 2000 al 2008, dalle prime aperture diplomatiche di Jiang Zemin, che portarono alle visite del presidente Clinton nel 1999 e di Bush nel 2000, fino ai Giochi Olimpici del 2008 che hanno rappresentato una sorta di legittimazione della Cina sul piano internazionale. L'intensità di crescita delle notizie in Google News dal 2000 al 2008 (fig. 5.1) rende conto – complessivamente – della rilevanza di questo periodo:

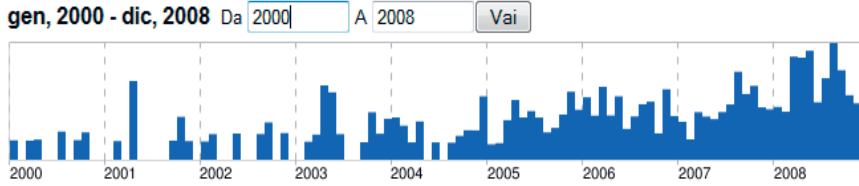


Fig. 5.1 – Intensità degli articoli nell’archivio in lingua italiana di Google News dal 2000 al 2008

5.3. IL CORPUS CINA2008

Il corpus Cina2008 (tab. 5.1) è costituito da 2.521 articoli (8.700 Mb) pubblicati tra il 1 gennaio 2000 e il 31 dicembre 2008 sui giornali quotidiani *Il Corriere della sera*, *La Repubblica* e *La Stampa* (Step 1).

Tab. 5.1 – Articoli pubblicati dai quotidiani con il nome “Cina” nel titolo tra il 1 gennaio 2000 e il 31 dicembre 2008

Anni	Corriere della sera	La Repubblica	La Stampa	Totale
2000	79	32	59	170
2001	86	57	76	219
2002	45	30	83	158
2003	111	38	79	228
2004	113	58	86	257
2005	221	79	94	394
2006	203	72	92	367
2007	203	68	96	367
2008	197	88	76	361
Totale	1.258	522	741	2.521

Il corpus, sottoposto a normalizzazione completa e lessicalizzazione in TaLTaC2 (Step 2), presenta le seguenti misure lessicometriche:

Tab. 5.2 – Misure lessicometriche del corpus CINA2008

Occorrenze	N	1.327.153
Forme grafiche	V	66.802
Type/Token ratio	(V/N)*100	5,03
Percentuale di hapax	(V1/V)*100	42,65
Frequenza media generale	N/V	19,87

La fascia di alta frequenza contiene 74 forme grafiche (tab.5.3), tra le quali, se escludiamo le forme grammaticali, troviamo esclusivamente parole banali (*Cina*, *Pechino*, *cinese*, *cinesi*) che tuttavia denotano la forte omogeneità del corpus rispetto al focus della ricerca. È certamente da sottolineare che complessivamente sono gli Stati Uniti a dominare la scena internazionale dei rapporti con la Cina. Nel primo decile delle medie frequenze si affacciano alcune delle parole chiave più rilevanti: i rapporti con l'Italia e l'Europa, gli altri paesi, il mercato internazionale e quella parola *presidente* che indica una forte caratterizzazione “ufficiale” di questi rapporti.

Tab. 5.3 – Forme grafiche principali appartenenti alla fascia di alta frequenza e al primo decile della fascia di media frequenza nel corpus Cina2008

Forme	Occorrenze	Rango	Fascia
Cina	11.759	15	Alta
Pechino	5.048	30	Alta
cinese	3.635	37	Alta
cinesi	3.285	39	Alta
anni	2.630	45	Alta
USA	2.045	51	Alta
milioni	1.736	67	Alta
mondo	1.716	68	Alta
governo	1.501	78	Media
Italia	1.481	79	Media
Paese	1.263	95	Media
presidente	1.204	98	Media
mercato	1.151	102	Media
paesi	1.076	111	Media
Europa	954	118	Media

Come forma isolata *presidente* conta in effetti 1.716 occorrenze (il che la andrebbe a collocare nella fascia di alta frequenza). Un approfondimento sulla segmentazione di questa forma rende conto meglio del suo contesto (tab. 5.4) che, come si è detto, delinea un campo semantico in gran parte “politico-governativo”.

Tab. 5.4 – Segmentazione della forma *presidente* nel corpus Cina2008: segmenti con occorrenze > 5

Segmento	N	Segmento	N
p. Hu Jintao	129	p. francese Jacques Chirac	10
p. Bush	41	p. della Fondazione Italia-Cina	9
p. Jiang Zemin	34	p. della Confindustria	8
p. di turno	33	p. della Provincia	8
p. cinese Hu Jintao	32	p. dell' istituto	8
p. russo	25	p. del Senato	7
p. cinese Jiang Zemin	24	p. della Commissione militare	7
p. Ciampi	24	p. del consorzio	7
p. della Camera di Commercio	20	p. taiwanese Chen Shui-Bian	7
p. del gruppo	16	p. e amministratore delegato	6
p. USA	15	p. della Fiat	6
p. Mao	14	p. della Bce	6
p. della Commissione europea	13	p. russo Putin	6
p. dell'associazione	11	p. russo Vladimir Putin	6
p. Chen	10	p. di turno dell' Unione Europea	6

Altrettanto interessante è l'approfondimento sulla forma *mercato* che conta 1.151 occorrenze (tab. 5.5) e dalla quale emerge con tutta evidenza che l'argomento dominante sulla stampa quotidiana è quello economico con riferimenti importanti ai rapporti tra i mercati internazionali dei prodotti e solo marginalmente agli aspetti critici rappresentati dal mercato del lavoro (19) e dal mercato nero (6).

Tab. 5.5 – Segmentazione della forma *mercato* nel corpus Cina2008:
segmenti con occorrenze ≥ 5

Segmento	N	Segmento	N
m. cinese	147	m. globale	9
m. interno	43	m. domestico	7
m. mondiale	27	m. internazionale	7
m. del lavoro	19	m. automobilistico	7
m. europeo	12	m. di sbocco	7
m. locale	12	m. del futuro	6
m. americano	10	m. immobiliare	6
m. asiatico	10	m. nero	6
m. italiano	9	m. dell'auto	6

5.4. ANALISI DELLE CORRISPONDENZE SULLE FORME PECULIARI E SPECIFICHE DEL CORPUS CINA2008

Lo scopo di questo lavoro è di ricostruire un'immagine di sintesi della Cina sui giornali quotidiani. L'analisi multidimensionale della matrice “parole per testi”, in particolare l'analisi delle corrispondenze binarie, risponde in modo ottimale a questo obiettivo: gli “individui-riga” della matrice sono rappresentati dalle parole mentre i “testi-colonna” sono rappresentati dalle 9 modalità della partizione “anno” e dalle tre modalità della partizione “testata”. Tuttavia una matrice di 66.000 righe produrrebbe un output numerico e grafico di difficile lettura e interpretazione. Si rende necessaria una selezione delle forme grafiche per ridurre la dimensione della matrice. La selezione sulla base di una soglia di occorrenze, per esempio una soglia 10 (che comporterebbe la scelta delle forme con occorrenze ≥ 10), non appare soddisfacente: la copertura del testo è elevata (90% delle occorrenze) ma le parole distinte sono comunque eccessive ($V = 12.229$).

In questi casi – di solito – è preferibile confrontare il vocabolario con un lessico di riferimento ed estrarre le forme peculiari, cioè quelle forme che sono “tipiche” della tematica trattata e che sono sovra-rappresentate nel corpus (Step 3). Le forme peculiari che sono significativamente utilizzate in Cina2008 rispetto al lessico di Rep90 presente in TalTac, costituito da 270 milioni di occorrenze di

10 annate del quotidiano *La Repubblica*, sono 659; a queste vanno aggiunte le 456 forme lessicalizzate nello Step 1 che ovviamente non possono essere sottoposte a confronto con un lessico esterno in quanto create all'interno del corpus stesso. Il confronto appare particolarmente adeguato trattandosi di un lessico giornalistico con una larghissima base testuale in termini di argomenti e di cronologia. Le forme peculiari più frequenti ci danno una conferma del contenuto prevalentemente politico-economico del corpus (tab. 5.6).

L'estrazione delle forme peculiari ci permette di ridurre la matrice a 1.115 forme. Potremmo considerare il risultato abbastanza soddisfacente. Tuttavia, se teniamo presente che l'analisi delle corrispondenze binarie si basa fondamentalmente su un confronto simultaneo tra i profili riga e i profili colonna, assumendo come riferimento la sovra (o la sotto-utilizzazione) delle forme in una partizione del corpus, possiamo ulteriormente ridurre la matrice se applichiamo al vocabolario delle forme peculiari e lessicalizzate una ulteriore analisi di specificità che individua le forme caratteristiche di ciascuna partizione (Step 4). La tabella 5.7 riporta, a titolo di esempio, le forme caratteristiche più frequenti nelle tre testate del corpus.

Tab. 5.6 – Parole peculiari e lessicalizzate più frequenti nel corpus Cina2008 rispetto al lessico giornalistico di riferimento Rep90

Forme	N	Forme	N	Forme	N
cinese	3.635	mondiale	630	ufficiale	387
cinesi	3.285	nuovi	602	americana	383
milioni	1.736	americani	594	viaggio	383
mondo	1.716	sviluppo	578	internet	376
Paese	1.263	rapporti	542	Repubblica Pop.	367
mercato	1.151	regime	527	occidentali	362
paesi	1.076	aziende	523	investimenti	361
USA	895	imprese	521	regione	360
città	883	americano	495	visita	355
euro	855	produzione	486	China	352
prodotti	719	Hu Jintao	459	militare	350
economia	716	diritti umani	403	popolazione	350
crescita	696	Olimpiadi	402	mercati	348
autorità	670	provincia	400	Occidente	337
India	638	libertà	389	premier	335

Tab. 5.7 – Forme peculiari e lessicalizzate specifiche più frequenti (soglia N > 150) nella partizione testata del corpus Cina2008 rispetto al lessico giornalistico di riferimento Rep90 (specificità positiva; soglia 10 e p-value < 0,025)

Corriere della sera	La Repubblica	La Stampa
milioni	cinesi	cinese
Paese	USA	cinesi
euro	città	Paese
aziende	autorità	mercato
imprese	americani	città
popolo	regime	prodotti
imprenditori	americano	crescita
industria	Hu Jintao	economia
Repubblica popolare		produzione
		sviluppo
		rapporti
		aziende

Con questa doppia selezione la matrice da sottoporre ad analisi delle corrispondenze binarie per la partizione “anno” è composta da 789 forme (311 lessicalizzate-specifiche e 478 peculiari-specifiche); la matrice della partizione “testata” è composta da 539 forme (184 lessicalizzate-specifiche e 355 peculiari-specifiche).

Tab. 5.8 – Analisi delle corrispondenze binarie sulla matrice delle forme grafiche per anni del corpus Cina2008: estrazione degli auto valori e inerzia (traccia: 0,5321)

Fattori	Autovalori	% inerzia	% cumulata
1	0.1123	21.11	21.11
2	0.0990	18.60	39.71
3	0.0814	15.31	55.02
4	0.0560	10.53	65.54
5	0.0534	10.03	75.57
6	0.0515	9.67	85.24
7	0.0457	8.59	93.84
8	0.0328	6.16	100.00

Complessivamente le forme grafiche esprimono una variabilità cronologica del 50%; il 18,6% di questa variabilità è spiegata dal secondo fattore che più propriamente rappresenta l'andamento temporale in modo lineare (fig. 5.2).

I riferimenti cronologici sono chiari: l'attentato alle Twin Towers; l'incidente diplomatico del 2 aprile 2001, durante il quale un quadrimotore militare americano si scontrò con un aereo cinese e fu costretto ad un atterraggio forzato; la polemica Cina-Usa sullo scudo spaziale; la lotta dei tibetani contro il regime intensificata durante l'anno di preparazione delle Olimpiadi; il virus della polmonite atipica (Sars) del 2003 e l'influenza aviaria del 2005. L'immagine della Cina è fortemente intrecciata con queste vicende. È significativo osservare come sul semiasse positivo del primo fattore i problemi sanitari va-

dai pregiudizi, qualora fosse confermata, non si associa necessariamente all'orientamento politico-culturale dei giornali quotidiani. Tuttavia si possono riscontrare alcune differenze, sebbene non marcate.

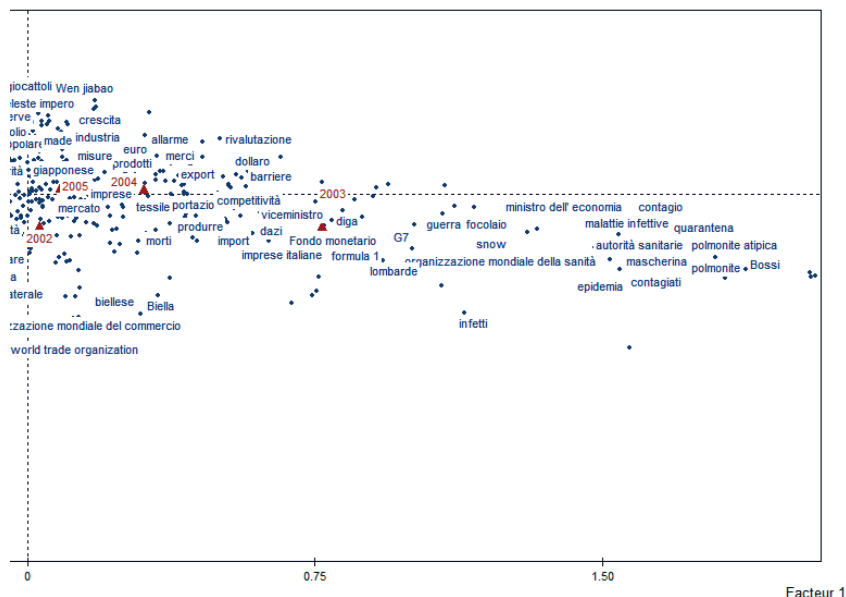


Fig. 5.2b – Analisi delle corrispondenze binarie della matrice forme grafiche per anni del corpus Cina2008: semipiano fattoriale positivo degli assi 1 e 2

Tab. 5.9 – Analisi delle corrispondenze binarie sulla matrice delle forme grafiche per testate del corpus Cina2008: estrazione degli auto valori e inerzia (traccia: 0,1539)

Fattori	Autovalori	% inerzia	% cumulata
1	0.0996	64.72	64.72
2	0.0543	35.28	100.00

Il primo fattore (fig. 5.3) spiega il 64,72% della varianza ed è rappresentato dalla contrapposizione tra le forme grafiche del primo semiasse negativo (fig. 5.3a) che caratterizzano prevalentemente le tematiche politiche (*regime cinese, disegualianza, esercito popolare, superpotenza, regime comunista*) e le forme del primo semiasse positivo (fig. 5.3b) che caratterizzano le tematiche economiche (*join*

venture, prodotti italiani, imprese italiane, tessile, Lingotto, marchio, commerciale, mercato, produrre, tecnologia, crescita, sviluppo).

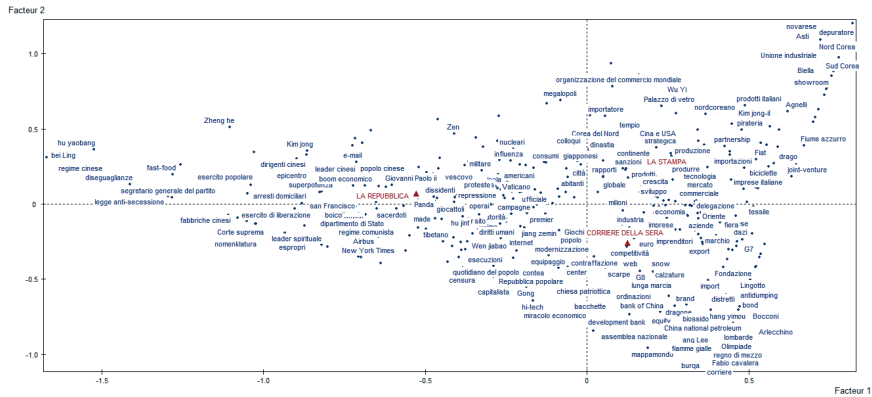


Fig. 5.3 – Analisi delle corrispondenze binarie della matrice forme grafiche per testate del corpus Cina2008: piano fattoriale degli assi 1 e 2

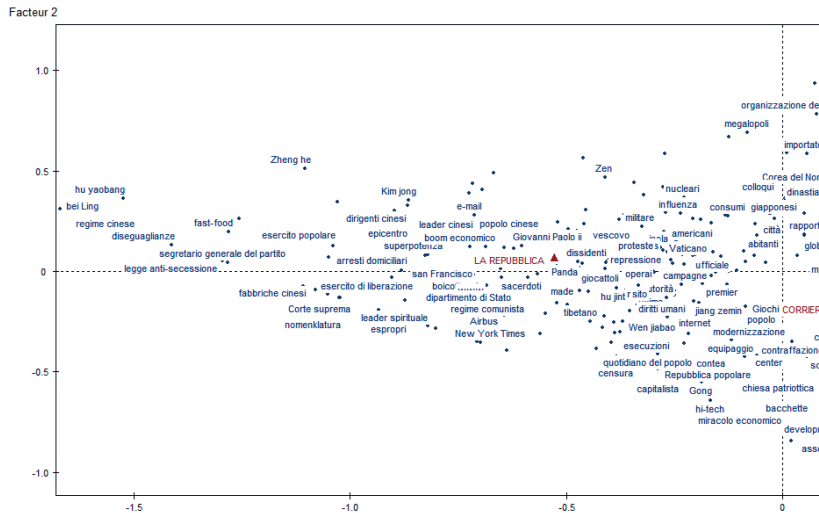


Fig. 5.3a – Analisi delle corrispondenze binarie della matrice forme grafiche per testate del corpus Cina2008: semipiano fattoriale negativo degli assi 1 e 2

È *La Repubblica* a delineare con più convinzione una rappresentazione critica delle scelte politiche compiute dal governo cinese (fig. 5.3a): in prossimità del punto in cui è rappresentato il profilo di questa testata giornalistica troviamo forme come *dissidenti*, *repressione*, *diritti umani*, *sacerdoti*, *regime comunista*, ecc.). Su questo semipiano troviamo anche alcuni riferimenti narrativi ai processi di trasformazione e riforma di cui la Cina è stata protagonista in questi anni (*Hu Yaobang*, segretario generale del Partito Comunista nei primi anni Ottanta).

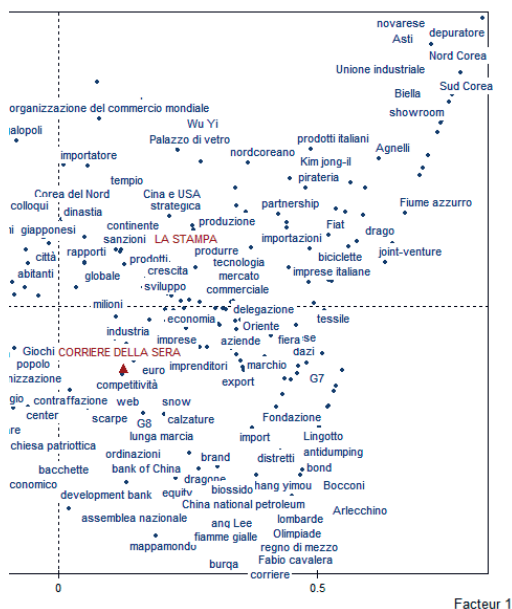


Fig. 5.3b – Analisi delle corrispondenze binarie della matrice forme grafiche per testate del corpus Cina2008: semipiano fattoriale positivo degli assi 1 e 2

La Stampa e il *Corriere della sera* insistono piuttosto sugli aspetti economici e finanziari. Sul secondo fattore (fig. 5.3b) possiamo notare una diversificazione tra queste due testate, con la prima (*La Stampa*) che mette in evidenza il successo dei prodotti italiani (*Biella*, *Asti*, *Unione industriali*, *Agnelli*, *Fiat*, *partnership*) indicando solo marginalmente le problematiche legate all'ingresso della Cina nei mercati internazionali (*pirateria*, *sanzioni*), e la seconda (*il Corriere della sera*) che affianca ad un forte interesse per le imprese un maggior senso critico (*antidumping*, *biossido*, *Fiamme Gialle*, *contraffazione*) con qualche prossimità ai temi

più caratterizzanti della Repubblica: *diritti umani, esecuzioni, internet, censura*). Nel caso della Stampa i rilievi critici, quando sono presenti, sembrano invece legati prevalentemente ai rapporti internazionali (*Cina e Usa, Corea del Nord, Palazzo di Vetro*).

5.5. IL TONO POSITIVO E NEGATIVO DEGLI ARTICOLI NELLA STAMPA

L'interrogativo finale sull'immagine della Cina nella stampa dal 2000 al 2008 si concretizza in una domanda: in sintesi l'immagine si presenta come positiva o negativa? La risposta, seguendo un percorso metodologico affidato all'analisi automatica, si può ottenere misurando il tono positivo o negativo degli articoli attraverso la presenza in essi degli aggettivi che esprimono positività o negatività (Step 6). Tuttavia, tale risposta, non può che essere approssimata e condizionata in modo significativo dal contesto nel quale l'aggettivo si colloca. Nella tabella 5.10 sono indicati i primi venti lemmi di aggettivi le cui forme flesse sono classificate con significato positivo o negativo nel database di TaLLaC.

Alcuni lemmi hanno un contenuto palese di ambiguità: <economico>, <pubblico>, <religioso>, <responsabile>, <pronto>, per i lemmi positivi; <passato>, per i lemmi negativi. Tuttavia va rilevato che tale ambiguità è più accentuata per il versante positivo che per il negativo. Il lemma <economico> in questo contesto ha un significato di per sé privo di connotazioni, ma il lemma <religioso> si associa più spesso con le proteste dei monaci tibetani o con gli attacchi alla Chiesa di Roma piuttosto che con accenti di positività legati agli aspetti spirituali. Sull'altro versante, se è vero che il lemma <passato>, con una forte ambiguità semantica e grammaticale, non può essere ascritto al campo negativo senza tenere conto del contesto, è altrettanto vero che gli altri 19 lemmi coprono inequivocabilmente da soli il 37% del campo semantico negativo.

Complessivamente l'indice di negatività, tenendo conto della *Pollyanna hypothesis* di cui si è già detto (Cap. 1), si colloca nella norma con uno 0,44 che – considerato che si tratta di un linguaggio giornalistico – non presenta particolari connotazioni negative. Certamente è significativa la diversificazione del punteggio rispetto alle testate giornalistiche, il che dimostra la sensibilità dell'indicatore nel rilevare il tono positivo o negativo del testo: gli articoli de *La Repubblica*, che nell'analisi delle corrispondenze si era qualificato come il giornale più critico nel rappresentare l'immagine della Cina (fig. 5.3), presenta un indice di negatività decisamente connotato sul versante negativo: 0,56. Al contrario *La Stampa* è il giornale che offre l'immagine più positiva della Cina.

Tab. 5.10 – Aggettivi positivi e negativi nel corpus Cina2008
(primi venti lemmi in ordine decrescente per occorrenze)

Aggettivi positivi	N	Aggettivi negativi	N
nuovo	2.697	morto	715
economico	814	vecchio	496
forte	534	passato	400
vero	513	difficile	347
aperto	498	duro	288
pubblico	482	povero	259
possibile	396	grave	190
ricco	395	pericoloso	188
importante	355	sospetto	173
buono	282	ferito	110
esperto	276	malato	105
religioso	269	debole	100
bello	262	cattivo	94
deciso	260	severo	92
responsabile	259	illegale	81
enorme	258	negativo	78
libero	236	falso	70
chiaro	195	contrario	69
pronto	192	condannato	67
superiore	179	contraffatto	66

*Tab. 5.11 – Occorrenze di aggettivi negativi e positivi (lemmi)
secondo la testata giornalistica e l'indice di negatività nel corpus Cina2008*

Testata	Aggettivi negativi	Aggettivi positivi	Negativi/Positivi
Corriere della sera	4.030	9.637	0,42
La Repubblica	2.769	4.979	0,56
La Stampa	2.914	7.513	0,39
Totale	9.713	22.129	0,44

5.6. CONCLUSIONE

L'analisi condotta sul corpus Cina2008, per quanto riguarda le testate giornalistiche esaminate, non offre un sostegno significativo all'ipotesi di lavoro secondo la quale stereotipi e pregiudizi condizionano l'immagine della Cina offerta dalla stampa. Certamente la stampa ha messo in evidenza, con accentuazioni diverse, gli eventi politici e sociali che caratterizzano le profonde trasformazioni vissute da questo paese senza, tuttavia, metterne artificiosamente in risalto solo gli aspetti più critici. L'analisi delle corrispondenze ha messo in evidenza quanto l'immagine della Cina – nell'opinione pubblica – sia condizionata dagli allarmi sanitari che hanno colpito il Paese, soprattutto dal 2003 al 2005. Se consideriamo che la maggior parte degli allarmi epidemici di questi ultimi anni, compresa la Sars e l'influenza aviaria, sono stati definiti recentemente “bufale apocalittiche” (Kerbaker, 2010) possiamo comprendere quanto la massiccia copertura mediatica data a questi eventi abbia influenzato – e forse danneggiato – l'affacciarsi dell'economia cinese sui mercati internazionali.

A queste catastrofi, in gran parte frutto di una “costruzione mediatica”, si affiancano nell'immagine, le catastrofi naturali (terremoti e inondazioni), che evidentemente, sono una parte rilevante della cronaca giornalistica, e quelle che appaiono come un effetto secondario dello sviluppo economico: inquinamento, contraffazioni e tensioni sociali. In questo ambito, sul piano più strettamente politico, hanno pesato le Olimpiadi, sulle quali, nonostante gli immensi sforzi compiuti dal governo cinese, è rimasta la macchia indelebile delle proteste associate all'indipendenza del Tibet e le controverse polemiche sulla censura e sui diritti umani ancora oggi di grande attualità.

STRATEGIA DI ANALISI

STEP 1 – ORGANIZZAZIONE, NORMALIZZAZIONE DEL CORPUS E LESSICALIZZAZIONE

Gli articoli del corpus Cina2008 sono stati selezionati dagli archivi online dei giornali quotidiani *Il Corriere della sera*, *La Repubblica* e *La Stampa*⁴ dal 1 gennaio 2000 al 31 dicembre 2008 secondo un “criterio di rilevanza” per il quale sono stati considerati solo gli articoli con il nome “Cina” nel titolo.

I testi sono stati sottoposti a normalizzazione completa in TaLTaC2 con identificazione di tre partizioni per i quotidiani e nove partizioni per gli anni. Il corpus, ad una prima analisi del vocabolario, presenta numerosi poliformi per i quali si richiede una lessicalizzazione: *diritti umani*, *human rights*, *low cost*, *anidride carbonica*, *polmonite atipica*, *rivoluzione culturale*, *prodotto interno lordo*, ecc.

Il processo di lessicalizzazione modifica radicalmente la struttura linguistica del corpus ma permette di conseguire risultati più attendibili per l'analisi del contenuto.

La lessicalizzazione di 456 “entrate” di vocabolario non è stata effettuata per i nomi propri che potevano comportare una perdita di informazione: *Giovanni Paolo II* è stato lessicalizzato perché l'identificazione del pontefice avviene sempre con il nome completo; Vladimir Putin non è stato lessicalizzato perché a fronte di 32 occorrenze di *Vladimir Putin* vi sono nel corpus 143 occorrenze *Putin* che, in caso di lessicalizzazione, si sarebbero ridotte (con duplicazione di forma grafica) a 111 occorrenze.

La lessicalizzazione avviene in quattro passaggi: 1) estrazione dei segmenti dal menu **Analisi** – *Analisi lessicale* – *Analisi dei segmenti* – *Individuazione dei segmenti*; 2) calcolo dell'indice IS per la selezione dei segmenti da lessicalizzare sulla base, soprattutto, dell'IS relativo; esportazione della lista di lessicalizzazione dal menu **File** – *Esporta* – *In un file di testo* – *Lista di lessicalizzazione/tematizzazione*, completamento della procedura dal menu **Analisi** – *Analisi lessicale* – *Analisi dei segmenti* – *Lessicalizzazione*.

⁴ Corpus predisposto con la collaborazione delle dott.sse Alice Deias e Maria Grazia Guerri (tesi di laurea specialistica in Comunicazione della conoscenza per le imprese e le organizzazioni).

STEP 2 - GENERAZIONE DEL VOCABOLARIO E MISURE LESSICOMETRICHE

Il secondo step, in questo caso, prende in considerazione il vocabolario solo a seguito della lessicalizzazione. Le misure lessicometriche di base permettono di valutare l'adeguatezza statistica del corpus. Il rapporto forme grafiche/occorrenze (type/token ratio) denota un'ottima estensione del vocabolario; anche la percentuale di hapax presenta un ottimo rapporto. Con TaLTaC2 dal menu **Analisi** – *Pre-trattamento* – *Calcolo delle sub-occorrenze* possiamo compiere alcune osservazioni sulle occorrenze secondo le testate e secondo gli anni.

STEP 3 – ESTRAZIONE DELLE FORME PECULIARI

Un primo passo nella riduzione di complessità in una matrice di dati testuali consiste nella estrazione delle forme grafiche che sono caratteristiche di un particolare lessico. Questa operazione richiede un lessico di riferimento con il quale effettuare il confronto e individuare le forme che sono sovrautilizzate (o sottoutilizzate) nel corpus. Le forme che sono utilizzate proporzionalmente in egual misura possono essere considerate influenti sulla caratterizzazione del corpus.

Tra le risorse di TaLTaC2, per il caso qui in esame, è presente il lessico *Rep90*, identificato anche come *Linguaggio comune – FG con uso >50*, che è una lista costituita da 60.489 forme grafiche tratte da una raccolta di 270 milioni di occorrenze di 10 annate del quotidiano *La Repubblica* (1990-1999). Si tratta pertanto di un lessico del linguaggio giornalistico particolarmente adeguato a essere utilizzato per individuare le parole caratteristiche presenti negli articoli a stampa sulla Cina.

Dal menu **Analisi** – *Analisi lessicale*, selezioniamo il comando *Linguaggio peculiare – Confronto con un lessico di frequenza* e seguiamo le opzioni offerte dalla finestra di dialogo effettuando il confronto sulla lista di intersezione, cioè sulle forme grafiche comuni tra i due lessici. Il risultato del confronto viene inserito nella lista *Vocabolario* della sessione come “scarto standardizzato” di ciascuna forma grafica rispetto alle forme del lessico di riferimento. Lo scarto avrà un segno positivo se la forma è sovrautilizzata nel corpus rispetto al lessico del quotidiano *La Repubblica*, negativo se è sottoutilizzata. Sono da ritenere significative e “peculiari” le forme con uno scarto in valore assoluto maggiore di 3,84 che corrisponde al valore del χ^2 con 1 grado di libertà e $p\text{-value}=0,05$.

STEP 4 – ESTRAZIONE DELLE FORME SPECIFICHE

L'estrazione delle forme specifiche permette di identificare le parole caratteristiche di una partizione del corpus rispetto all'insieme del corpus stesso. Si tratta, pertanto, di individuare le parole sovrautilizzate (o sottoutilizzate) negli articoli di una testata rispetto alle altre.

Dal menu **Analisi** – *Analisi Lessicale* – *Analisi delle specificità*, dopo aver effettuato il calcolo delle sub-occorrenze nel corpus per la variabile “Testata”, otteniamo per ciascuna forma grafica del *Vocabolario* il valore di probabilità $\alpha \leq 0,025$ (“p-value”) che la forma nella parte in esame sia sovra o sotto-dimensionata rispetto alle attese per il solo effetto del caso (legge di distribuzione ipergeometrica; Tuzzi, 2003, pp. 131-134).

STEP 5 – ANALISI DELLE CORRISPONDENZE BINARIE

Nello Step 5 del capitolo 3 abbiamo già illustrato i principi generali dell'analisi delle corrispondenze binarie. In questo caso nella matrice “parole per testi”, la variabile “testi” è costituita dalle 9 modalità “anno” (dal 2000 al 2008), pertanto l'analisi delle corrispondenze permetterà di individuare al massimo otto dimensioni. Nella seconda applicazione i “testi” sono rappresentati dalle tre modalità della variabile “testata”: *Il Corriere della sera*, *La Repubblica* e *La Stampa* e permettono di individuare due dimensioni.

STEP 6 – CALCOLO DELL'INDICE DI NEGATIVITÀ

La presenza del dizionario degli aggettivi positivi e negativi tra le risorse statistico-linguistiche di TaLTaC2 permette di “etichettare” le forme grafiche del corpus e di effettuare una valutazione d'insieme, sebbene approssimativa, sul “tono” positivo o negativo dei testi sottoposti ad analisi.

Dal menu **Analisi** – *Analisi Lessicale* – *Confronto con un lessico di frequenza* mettiamo a confronto il Vocabolario della sessione con il modello di riferimento *Dizionario aggettivi positivo-negativo*, utilizzando l'opzione “Intersezione”; nei campi da inserire nella lista risultato selezioniamo “Forma grafica” e “Occorrenze” nella finestra della Lista da confrontare, “CAT-SEM” e “Lemma” nella finestra del Modello di riferimento.

Il conteggio del totale delle occorrenze si ottiene con una “lemmatizzazione” sulla tabella di intersezione selezionando la colonna “CAT-SEM” e, dal menu **Calcola**, il comando *Fusioni*. Sulla tabella riassuntiva del risultato che riporterà il numero di unità lessicali (forme flesse degli aggettivi)

con cui la categoria (aggettivi positivi e negativi) è stata formata e il totale delle occorrenze per ciascuna categoria si potrà calcolare l'indice di negatività che è dato dal rapporto tra il totale delle occorrenze negative e il totale delle occorrenze positive (Occ Neg / Occ Pos): da una serie di prove effettuate su liste di riferimento risulta che un valore superiore a 0,40 indica un testo con una connotazione tendenzialmente negativa (Cap. 1).

Per un confronto dell'indice di negatività secondo le testate è necessario riportare l'informazione contenuta all'interno della tabella di intersezione nella tabella *Vocabolario* della sessione. Dal menu **Analisi** si seleziona il comando *Analisi Lessicale – Tagging Semantico* e poi *Vocabolario*, seguendo le istruzioni fino a “Includi le tabelle dal DB della sessione: Intersezione di Vocabolario con TAG grammaticale” e “POSNEG”. In seguito a questa operazione la tabella *Vocabolario della sessione* mostrerà gli aggettivi etichettati con le modalità negativo-positivo e permetterà di calcolare l'indice di negatività del testo per ciascuna modalità con la consueta procedura di fusione delle forme grafiche.

RIFERIMENTI BIBLIOGRAFICI

- BONAZZI F. (1998) *Itinerari di sociologia della comunicazione*, Milano, Franco Angeli, 1998.
- CARITAS E MIGRANTES (2010) *Dossier statistico 2009. XIX Rapporto sull'immigrazione*. Roma, Edizioni Eidos.
- FARR R. M., MOSCOVICI S. (1989) *Rappresentazioni sociali*, Il Mulino, Bologna.
- KERBAKER A. (2010) *Bufale apocalittiche*, Firenze, Ponte delle Grazie.
- PEDONE V. (2008) *Il vicino cinese*. Roma, Nuove Edizioni Romane.
- PIEKIE F. (2004) *Chinese Globalisation and Migration in Europe*, San Diego, CA, Center of Comparative Immigration Studies.
- ROZMAN G. (2002) “China's Changing Images of Japan, 1989-2001: The Struggle to Balance Partnership and Rivalry”, in *International Relations of the Asia-Pacific*, 2, 1, pp. 95-129.
- TIEN C., NATHAN J.A. (2001, “Trends: American Ambivalence toward China”, in *The Public Opinion Quarterly*, 2001, 65, 1, pp. 124-138.
- TUZZI A. (2003) *L'analisi del contenuto. Introduzione ai metodi e alle tecniche di ricerca*, Roma, Carocci.
- YAN Y. (2000) “Of Hamburger and Social Space: Consuming McDonald's in Beijing”, in D. S. Davis, [Ed], *The Consumer Revolution in Urban China*, Berkeley, CA: Univ. California Press, pp. 201-225.

6.

ISTANTANEE DELLA FAMIGLIA IN PUBBLICITÀ: TRADIZIONALE E RICOMPOSTA

6.1. INTRODUZIONE

Nell'introduzione a *Le nuove famiglie* A. L. Zanatta (2008) cita E. Durkheim per sostenere che “non esiste un modo di essere e di vivere che sia il migliore per tutti (...) La famiglia di oggi non è né più né meno perfetta di quella di una volta: è diversa, perché le circostanze sono diverse (p. 7)”. Cambiamenti che hanno visto la famiglia passare da un modello tradizionale basato sul matrimonio a una famiglia ricostituita o ricomposta in cui la scelta di due individui di stare insieme dà vita a una nuova forma familiare.

Si tratta di due passaggi epocali di cui si può dar conto prendendo a riferimento ora l'opera di M. Barbagli *Sotto lo stesso tetto* (1984) ora il lavoro già citato di A. L. Zanatta (2008) *Le nuove famiglie*.

Delineando i principali mutamenti avvenuti nella famiglia centro-settentrionale, Barbagli argomenta che la famiglia nucleare è comparsa qui molti secoli prima di quella coniugale intima. Nella sua analisi lo studioso distingue tra quanto accade nelle zone rurali e quanto si verifica, invece, nei centri urbani. È in città che – tra il XIV e il XV secolo – la forma nucleare si afferma prima ancora che nelle campagne, ovviamente con alcune eccezioni. Le motivazioni non sono – sostiene ancora l'autore – da ricercare nell'alto tasso di mortalità proprio di questi secoli. Infatti, gli artigiani delle città, pur avendo ancora i genitori in vita, creavano delle famiglie nucleari; il medesimo comportamento veniva adottato anche dagli strati più poveri della popolazione urbana. Gli unici che continuavano a risiedere nella

casa del padre – in città – sono stati gli strati appartenenti ai ceti più elevati. Diversa la situazione nelle campagne, dove le famiglie complesse erano più frequenti in tutti gli strati sociali della popolazione con alcuni incrementi in quelle che vivevano nei poderi. Questa struttura si modificò tra il Settecento e l'Ottocento, probabilmente a causa dei cambiamenti avvenuti nelle regole di trasmissione della proprietà da una generazione all'altra fra i ceti più agiati (p. 24).

Con Barbagli ci fermiamo intorno al 1970 e ricorriamo a Zanatta per un'analisi che dia conto dei cambiamenti più recenti. L'autrice ci dice che con la dizione famiglia ricostituita “si intende una coppia convivente, sposata o non sposata, con o senza figli, in cui almeno uno dei due partner proviene da un precedente matrimonio interrotto per morte, separazione o divorzio (Zanatta, 2008, 76)”.

Il passaggio dalla famiglia tradizionale a quella ricostituita deriva – secondo Zanatta (2008) – da un lato da fattori di natura socioeconomica, come l'industrializzazione avanzata e l'ingresso delle donne nel mercato del lavoro e dall'altro da fattori di natura culturale, come il declino dei valori religiosi tradizionali, l'affermazione dell'autonomia individuale e dell'ideale dell'amore romantico. Se il matrimonio deve essere basato sull'amore e non più sugli interessi economici delle famiglie, scegliere di interromperlo e contrarne un altro diventa quasi una necessità di vita; poiché si sa l'amore non è eterno e soprattutto non ha età.

Questo slittamento comporta il passaggio da una famiglia moderna o nucleare, formata da una sola unità coniugale – che può essere intesa sia nella forma completa, ovvero composta da marito, moglie con o senza figli oppure incompleta, dove viene a mancare uno dei due coniugi –, a una ricomposta in cui è palesemente mutata la struttura famigliare.

La conseguenza di questi diversi processi è il sorgere e il diffondersi di nuovi modi di “fare famiglia”, di nuovi tipi di famiglie che vanno dalle convivenze alle famiglie ricomposte, dalle coppie omosessuali alle coppie multiculturali; cui si affiancano le famiglie nucleari, che oggi diventano quelle “tradizionali”. Il passaggio da una forma all'altra non è né esclusivo né esaustivo delle forme di legame famigliare che sono oggi presenti nella nostra società, anzi esse convivono l'una accanto all'altra. Ad aver intuito bene questa identità multipla della famiglia è senza dubbio il mondo pubblicitario, che a seconda del prodotto che deve reclamizzare e il target verso cui deve promuoverlo utilizza ora uno stereotipo di struttura famigliare ora l'altro.

Obiettivo del presente lavoro è analizzare due spot pubblicitari in cui il soggetto sia la famiglia ora nella veste di famiglia ricomposta ora nella forma di famiglia tradizionale. Gli spot scelti sono quello di *Ernestomeda 2009*, produttore di cucine, e la *réclame* della *Renault Scénic 2009*.

6.2. IMMAGINI DELLA FAMIGLIA TRADIZIONALE

Gli elementi costitutivi di una famiglia “tradizionale” sono essenzialmente individuabili nella coppia, nella casa (il loro nido), nei figli. La ricerca di questi elementi all'interno di uno spot può avvenire ricorrendo al *décompagè*⁵ delle sequenze, in modo da suddividere in quadri scenici il filmato da analizzare. Non appena si ha coscienza di cosa contenga lo spot si può iniziare a interrogarsi sul perché dell'inserimento di quegli elementi piuttosto che di altri.

L'uso della coppia, della casa, dei figli all'interno di uno spot pubblicitario induce nello spettatore il richiamo immediato dell'idea di una famiglia nucleare tradizionale. Nelle immagini proposte dallo spot *Ernesto Meda 2009* questi tre elementi sono presenti secondo un ordine temporale ben preciso e scanditi dall'alternarsi di momenti di luce e di buio, che sussurrano la metafora della ciclicità e dei giorni che, inevitabilmente, trascorrono.

L'immagine della famiglia nucleare tradizionale è mostrata nella sequenza⁶ di apertura, in cui fanno capolino la madre, il padre e un bambino; quest'ultimo è ritratto nell'atto di giocare con della pasta (fig. 6.1) per poi essere nuovamente posto al centro della scena durante un'altra attività di gioco: l'atto di saltare dentro una pozzanghera.



Fig. 6.1 – Ritratto di bambino che gioca

⁵ Si veda il dettaglio nello step 1.

⁶ Formata dalle inquadrature 1, 2, 3, 4.

Nella seconda ⁷ sequenza muta lo scenario: la cucina che fa da sfondo non è più la stessa, mutano anche i personaggi oltre che il tempo del racconto. Al centro della sequenza è il rapporto padre-figlia: un uomo spazzola i capelli a una bambina: un gesto intimo, tanto intimo da lasciar intendere che si tratti del padre della stessa (fig. 6.2).

Nella terza sequenza ⁸ si chiarisce come abbia avuto origine la famiglia: i protagonisti questa volta sono due innamorati: un uomo e una donna che aspettano l'alba insieme ⁹.



Fig. 6.2 – Immagini del rapporto padre-figlia

Il setting e gli oggetti presenti in questo spot assumono una duplice valenza. Si tratta, infatti, di mettere in scena rimandi continui alla vita che si può svolgere all'interno di una cucina, che è vero che rappresenta il prodotto reclamizzato, ma soprattutto il focolare domestico, l'unione familiare, la serenità.

Il tema della serenità e/o della felicità è un elemento indispensabile nella promozione di un prodotto, poiché questo deve essere prima di tutto desiderabile.

⁷ Formata dalle inquadrature 5, 6, 7, 8, 9.

⁸ Formata dalle inquadrature 10, 11, 12, 13, 14.

⁹ Per completezza è bene precisare che c'è un'ultima sequenza formata dalle inquadrature 15,16 dove si abbandonano i racconti per lasciare visibilità al brand.

Non stupisce che tutti i personaggi sembrano essere felici: il bambino e i suoi genitori nella prima sequenza, così anche per il padre e la figlia protagonisti della seconda sequenza e, infine, lo stesso dicasi per la coppia dell'ultima sequenza, che ci regala anche una nota di intimità.

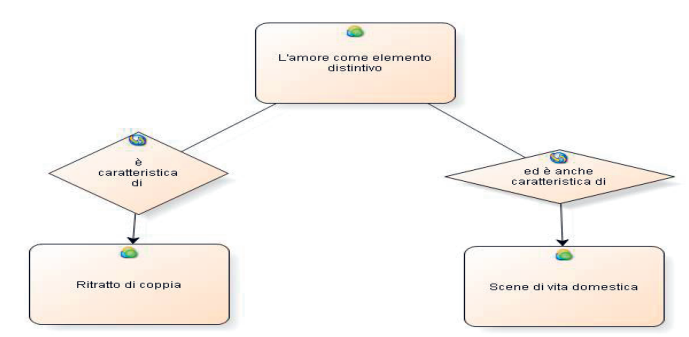


Fig. 6.3 – Elementi costitutivi del nucleo familiare

6.3. IMMAGINI DELLA FAMIGLIA RICOMPOSTA

Il frame narrativo dello spot *Renault Scénic 2009* è il tragitto in auto che un padre compie tornando a casa assieme ai suoi cinque figli. Siamo in presenza di uno spaccato di vita quotidiana: si tratta di un genitore che, come molti altri, spende parte del pomeriggio accompagnando i suoi figli a lezione di nuoto, di tennis, di danza e poi li va a riprendere, per poter così tornare, finalmente, tutti insieme a casa.

Per sottolineare che si tratta di una famiglia ricomposta il regista dello spot si avvale dell'artificio del flashback. Ci mostra il momento in cui l'uomo, protagonista dello spot, ha contratto il primo matrimonio; ci narra della dissoluzione dello stesso per poi avvicinarci temporalmente all'evento del secondo matrimonio e condurci alla situazione attuale.

In virtù dell'ampio scarto temporale in cui è articolata la narrazione è possibile individuare due macrosequenze, che qui saranno chiamate “prima” e “adesso”; queste ultime sono composte da ulteriori sequenze.

Della sequenza “prima” fanno parte le scene in cui si dà conto della na-

scita del figlio concepito con la prima moglie ¹⁰, del matrimonio contratto con la stessa ¹¹; si passa poi agli elementi definitori del secondo matrimonio segnalato dal primo piano della seconda moglie ¹², cui segue la presentazione del figlio concepito con la stessa ¹³.

La sequenza denominata “adesso” si snoda lungo un arco temporale presente al protagonista dello spot: lui è in macchina e poco alla volta introduce i suoi cinque figli. In ordine è possibile assistere: alla presentazione del primo figlio Daniele, nato dal matrimonio con la prima moglie ¹⁴; alla presentazione dei figli acquisiti dal protagonista, i due gemelli nati dal precedente matrimonio della seconda moglie ¹⁵. Segue la presentazione della seconda figlia, Sofia nata dal matrimonio con la seconda moglie ¹⁶. Si ha ancora la presentazione di un figlio che il protagonista ha recentemente scoperto di avere ¹⁷. E finalmente si arriva a casa, dove la seconda moglie, sull'uscio, sta aspettando sorridente marito e figli ¹⁸. Nello spot è presente anche una nota di multiculturalità, infatti appena giunto sul vialetto di casa, l'uomo si accorge che in macchina c'è anche Arturo, il figlio del vicino che ha dimenticato di accompagnare a casa ¹⁹.

È interessante notare le differenze che intercorrono tra le due sequenze di flashback che rappresentano i due momenti principali della prima e della seconda unione familiare del protagonista.

Apparentemente le due sequenze sembrano essere uguali. In entrambe si ritrovano le scene del giorno del matrimonio e del giorno della nascita di uno dei figli. I flashback riferiti alla prima e alla seconda unione coniugale sono realizzati mediante delle riprese di tipo amatoriale, è identificabile la trama della pellicola, che ci fa capire che si tratta di un filmato realizzato “in casa”. Tale

¹⁰ Trattasi delle inquadrature: 2, 3, 4, 5.

¹¹ Trattasi delle inquadrature: 6, 7.

¹² Trattasi delle inquadrature: 13, 14.

¹³ Trattasi delle inquadrature: 19, 20, 21.

¹⁴ Il protagonista accompagna il figlio a lezione di nuoto. Il riferimento è alle inquadrature: 1, 8, 9, 10.

¹⁵ Il protagonista va a prendere i gemelli alla lezione di tennis. Il riferimento è alle inquadrature: 11, 12, 15, 24, 25.

¹⁶ Il protagonista va a prendere la bambina dalla lezione di danza classica appena terminata. Il riferimento è alle inquadrature: 16, 17, 18, 22, 23.

¹⁷ Trattasi dell'inquadratura: 26.

¹⁸ Trattasi delle inquadrature: 27, 28, 29.

¹⁹ Trattasi delle inquadrature: 30, 31. Per amore di completezza è bene segnalare anche l'ultima sequenza, composta da tre frame all'interno dei quali si presenta il nome del prodotto, lo slogan dello spot *facciamo posto a tutte le famiglie* e le informazioni tecniche e commerciali del prodotto pubblicizzato. Il riferimento è alle inquadrature: 32, 33, 34.

similitudine sembra implicitamente affermare che le due esperienze familiari hanno lo stesso grado d'importanza nella vita del protagonista e il loro ricordo ha per lui la stessa valenza.

Tuttavia, alcuni dettagli presenti nelle immagini lasciano percepire delle differenze. Le differenze più rilevanti sono connesse alle scene del matrimonio e sono legate principalmente all'abbigliamento degli sposi e al setting.

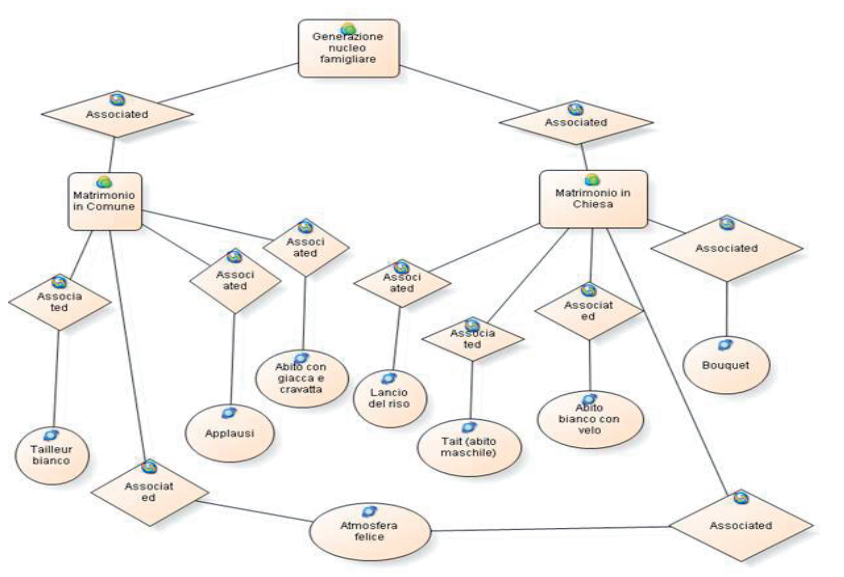


Fig. 6.4 – Elementi di cui si compone il gruppo del set “modalità di generazione del nucleo familiare”

Nel riannodare le fila della creazione dei due nuclei familiari, è utile visualizzare quanto appena detto mediante il modello generato dall'esplicitazione del contenuto del set “modalità di generazione del nucleo familiare” (fig. 6.4).

Nelle scene del primo matrimonio, la sposa indossa il tradizionale abito nuziale (abito bianco con velo e *bouquet*), lo sposo indossa un vestito scuro elegante con *papillon*. Gli sposi sono felici, sorridenti, entusiasti; gli ospiti che li circondano lanciano riso e coriandoli bianchi in un'atmosfera di festa. L'edificio dal quale escono sembra essere una chiesa.



Fig. 6.5 – Il primo matrimonio

Nelle scene del secondo matrimonio, invece, la sposa indossa un *tailleur* bianco, semplice e sobrio; lo sposo un vestito chiaro e una cravatta. Sullo sfondo sembra esserci, non più una chiesa ma il Comune. Gli ospiti applaudono ma senza lanciare riso o coriandoli. L'evento viene rappresentato, dunque, come un momento più sobrio rispetto al precedente.



Fig. 6.6 – Il secondo matrimonio

Mentre nella sequenza del primo matrimonio le immagini rimandano a una dimensione ludico-festosa, la sequenza del secondo matrimonio presenta un'atmosfera di felicità più "matura", più sobria.

6.4. ALCUNE CONSIDERAZIONI

Negli spot qui presentati sono ritratti due tipi distinti di famiglia, c'è tuttavia da segnalare che trattandosi di una pubblicità commerciale tanto i toni, che i colori sono in entrambi simili e soprattutto sereni e spensierati. Scopo dello spot non è far riflettere l'*audience* sulle implicazioni insite nell'uno o nell'altro modello, quanto piuttosto di fornire una identificazione e indurre il desiderio dell'acquisto.

È nell'identificazione che gli spot stimolano che si inserisce la nostra riflessione sui modelli famigliari.

Il primo spot ci narra la storia della vita familiare articolandola nello spazio di una cucina. Il setting scelto – il focolare – e i soggetti ritratti nelle sequenze inducono in noi l'idea che si tratti di una rappresentazione sociale di un tipo di famiglia nucleare tradizionale.

Il secondo spot analizzato è la storia di un padre che accompagna i suoi cinque figli (tre naturali e due acquisiti) alle loro attività sportive quotidiane, per poi andarli a riprendere. Questo spot non vuole creare una identificazione soltanto nelle famiglie numerose, ma fa leva sulle famiglie numerose di nuova generazione. Ovvero quelle che sono numerose non perché due coniugi di primo matrimonio hanno dato la luce a un numero elevato di figli, bensì perché a seguito di più e nuove unioni sono nati più figli che adesso si trovano a dividere un medesimo spazio: l'abitacolo di un'auto. All'interno di questo spot non si legge una discriminazione tra il modello tradizionale e il nuovo modello di famiglia (poiché per il venditore anche una famiglia numerosa di tipo tradizionale potrebbe rappresentare un possibile cliente). La famiglia ricomposta non assume connotazioni negative, anzi in un meccanismo simile a quello delle *sliding doors*: apre e chiude la scena su declinazioni di famiglie tradizionali e ricomposte. Si vedono, infatti, gli stessi volti felici e sorridenti (l'espressione degli sposi in entrambi i matrimoni, l'espressione dei neogenitori il giorno della nascita dei figli, la moglie sorridente che aspetta a casa il ritorno del marito e dei figli) e le stesse dimostrazioni d'affetto (il marito vicino alle mogli durante il parto, le carezze del padre ai neonati) che dominano tanto nello spot *Ernesto Meda 2009* qui esaminato, tanto nella lunga lista di spot rappresentanti famiglie nucleari tradizionali.

Renault, in questo spot, provoca lo spettatore proponendo come "naturale" un modello di famiglia agli antipodi rispetto a quello tradizionale. In questo modo riesce a colpire il pubblico, presentando un modello di famiglia non ancora am-

piamente diffuso a livello pubblicitario. Allo stesso tempo, però, tutela il suo rapporto col destinatario mediante il ricorso a immagini e schemi sociali ampiamente riconosciuti e accettati. Vengono infatti mantenuti, qui, quegli stereotipi che assicurano l'accettabilità dello spot agli occhi del target e che cancellano il rischio di rifiuto del messaggio. La famiglia che lo spot presenta è una famiglia non nucleare, non tradizionale, ma felice, alla stregua di qualsiasi altra famiglia "pubblicitaria" tradizionale.

STRATEGIA DI ANALISI

STEP 1 – IL DÉCOUPAGE DELLE IMMAGINI

All'interno di uno spot pubblicitario si condensano diversi riferimenti culturali, sociali, economici elaborati in modo che possano avere una loro presa sul pubblico. Sebbene in NVivo8 possono essere inseriti dei filmati, per chi non è avvezzo a lavorare sugli spot pubblicitari e a discettare sul loro sincretismo, diventa più utile realizzare il *découpage* di uno spot scomponendolo in alcuni dei suoi elementi: inquadrature, elementi visivi ed elementi sonori. Tale procedura consentirà di impadronirsi degli elementi di cui esso si compone.

Nel lavoro con NVivo può risultare interessante affiancare al *découpage* l'analisi delle immagini ritenute maggiormente esplicative del rapporto e dell'ideale di famiglia che la pubblicità in oggetto vuole trasmettere.

Nell'effettuare il *découpage* la prima scomposizione da compiere è quella in inquadrature. Che secondo C. Bianchi (2007) "può essere utile per comparare le diverse sostanze espressive costituenti il testo in esame (p. 51)".

Lo strumento scelto per facilitare la scomposizione è una tabella, composta da una colonna visiva e una sonora: all'interno della colonna visiva si descrive l'inquadratura con riferimento ai personaggi coinvolti e all'azione; all'interno della colonna sonora si dà conto del verbale e degli elementi musicali corrispondente a ogni inquadratura.

Alla fase di scomposizione, segue una fase di ricomposizione: le inquadrature, omogenee tra loro in termini di significato e unità d'azione, vengono raggruppate in sequenze.

Questo tipo di esercizio permette anche a chi non si occupa di pubblicità di poter lavorare su tali testi ricorrendo alla scomposizione in sequenze come mappa per muoversi all'interno degli spot.

Nelle tabelle seguenti si propone la scomposizione di due spot: *Ernesto Meda* del 2009 di 41 secondi e quello mandato in onda – sempre nel 2009 – dalla *Renault* per promuovere la *Scénic* della durata di 46 secondi.

Scomposizione 1 – Ernesto Meda 2009 ²⁰

Inquadrature		Colonna visiva	Colonna Sonora	
N°	Durata	Descrizione	Verbale	Musica (e sonoro)
1	4"	Un bambino, seduto su un ripiano della sua cucina, disegna una faccia sul piatto usando i cereali	“ridere...”	Canzone spot
2	2"	L'inquadratura si allarga e viene visivamente introdotta la famiglia del bambino	//	Canzone spot
3	2"	In cucina si fa buio e le ante degli armadi si trasformano in schermi nei quali è proiettata l'immagine di due innamorati	“...innamorarsi...”	Canzone spot
4	3"	Cambia il modello di cucina e mentre una donna, della quale non si vede il viso, lava l'insalata, su una superficie viene proiettata l'immagine di un bimbo che salta in una pozzanghera	“...una bambino che salta nella pozzanghera...”	Canzone spot (Risata del bambino e cane che abbaia lontano)
5	2"	Cambia nuovamente il modello di cucina. Nessun personaggio è presente mentre nella penombra del-	“...un prato di lillà...”	Canzone spot (Cinguettio di uccelli)

²⁰ Parte del materiale qui utilizzato proviene dal seminario “Sociologia visuale” svolto dalla dott.ssa G. La Rocca all'interno del corso *Teorie e tecniche della comunicazione sociale e d'impresa*, a.a. 2009-2010, Facoltà di Scienze della Formazione, Università degli Studi di Palermo. La scomposizione in sequenze è stata realizzata dalle studentesse Simona Campo per lo spot Ernesto Meda e da Maria Stella Palumberi per Renault Scénic come prova didattica all'interno del seminario.

		p'ambiente si fa chiara la proiezione di un prato di lillà		
6	2"	Si fa giorno e un ragazzo cerca qualcosa negli scompartimenti dell'armadio mentre un uomo spazzola i capelli ad una bambina	"...qualcuno che ti spazzola i capelli..."	Canzone spot
7	2"	La bambina mangia qualcosa seduta su uno sgabello della cucina	//	Canzone spot
8	3"	Nella stessa cucina in penombra viene inquadrata una pentola che bolle, sullo sfondo un dettaglio della cucina (sembrerebbero tre forni), alcune proiezioni di gocce sulla superficie dei forni e sulle ante degli armadi adiacenti	"...il ticchettio della pioggia..."	Canzone spot (Ticchettio pioggia)
9	2"	Il contesto è il medesimo della precedente inquadratura. Una volta cessata la pioggia l'attenzione si concentra su una pentola che bolle	//	Canzone spot (Una pentola che bolle)
10	2"	La cucina inquadrata cambia; è giorno, l'inquadratura dall'alto riprende i piani di lavoro.	"...la radio..."	Canzone spot con audio radio
11	3"	Appaiono delle proiezioni di numeri che indicano le stazioni di frequenza delle radio sui soggetti dell'inquadratura quali piatti.	"...sulla tua canzone preferita..."	Canzone spot con audio radio
12	2"	Cambia il modello di cucina. Due innamorati stanno	"...aspettare l'alba..."	Canzone spot (gab-

		insieme mentre una luce si fa sempre più chiara		biani)
13	2"	Luce artificiale sulle superfici degli armadi	//	Canzone spot (gab-biani)
14	2"	I due innamorati si abbracciano	"...capirsi..."	Canzone spot (ancora un po' gab-biani)
15	3"	Scritta bianca su sfondo nero: "non sempre le cose preziose sono irraggiungibili"	"...una cucina Ernesto Meda..."	Canzone spot
16	5"	Scritta bianca su sfondo nero: "Ernesto Meda. La bellezza non è un privilegio"	"...la bellezza non è un privilegio..."	Canzone spot

Scomposizione 2 - Renault Scénic 2009

Inquadratura		Colonna visiva	Colonna sonora	
N°	Durata	Descrizione	Verbale	Musica
1	1.5"	Un bambino con le cuffie si gira per guardare fuori dal finestrino dell'auto in movimento.	(voice off) Uomo: "Lui è Daniele, mio figlio.	Inizio colonna sonora. Brano esclusivamente musicale che accompagna l'intero spot.
2	1"	(Flashback) All'interno di una sala ospedaliera, una donna è distesa su un letto. Un uomo, in piedi accanto a lei, le tiene la mano.	... È nato dal matrimonio ...	
3	1"	La donna sta per partorire. Mentre tiene la mano al-	... con Elena, ...	

		L'uomo, urla e ha un viso sofferente.		
4	1"	La donna guarda sorridente il neonato (avvolto in una coperta) che tiene in braccio. L'uomo, accanto a lei, carezza il neonato.	... la mia prima moglie ...	
5	1"	La donna distesa sul letto tiene in braccio il neonato, accanto a lei l'uomo carezza il neonato. Un'infermiera sistema dei fiori sul comodino.	... Ah ...	
6	1"	Due sposi escono da una chiesa e scendono la scalinata mentre gli invitati applaudono e lanciano coriandoli bianchi.	... è stato un ...	
7	1"	I due sposi sorridono. (Fine flashback)	... matrimonio strepitoso!	
8	2"	L'uomo che guida l'auto, si gira per guardare dallo specchietto retrovisore il figlio seduto nel sedile posteriore.	... Adesso lo porto in piscina ...	
9	2"	L'auto in movimento, rallenta e si ferma davanti a un edificio.	... per il corso di nuoto.	
10	1"	L'auto riparte. Il bambino, già sceso dall'auto, cammina verso l'edificio.	... Poi riparto ...	
11	2"	Due gemelli aspettano in piedi sul marciapiedi, davanti ad un edificio. Ognuno dei due tiene in mano una racchetta da tennis. A terra ci sono due borsoni.	... e vado a prendere i gemelli Marco e Luca.	

12	1”	Arriva l’auto che rallenta e si accosta davanti ai gemelli che, intanto prendono i borsoni da terra.	... Gli altri due miei figli, ...	
13	1.5”	(Flashback) Due sposi escono da un edificio; stanno per scendere da una scalinata circondati dagli invitati.	... cioè di mia moglie, ...	
14	0.5”	Donna sorride. (fine flashback)	... la mia nuova moglie.	
15	1”	Gemelli aprono lo sportello posteriore dell’auto, mettono dentro i borsoni e salgono.	Li prendo e li porto subito ...	
16	1”	Auto in movimento.	... a prendere la mia ...	
17	1”	Diverse bambine in tutù rosa a lezione di danza classica	... terza figlia, ...	
18	1”	Bambina in tutù.	... la piccola Sofia. ...	
19	1”	(Flashback) Una donna sta per partorire, distesa su un letto in una stanza d’ospedale. Urla e ha un viso sofferente.	Ah... Sofia.	
20	1”	La donna distesa sul letto tiene in braccio il neonato, accanto a lei l’uomo che carezza il neonato.	È mia figlia, ...	
21	1”	La donna guarda sorridente il neonato (avvolto in una coperta) che tiene in braccio. Si gira poi sorridente verso l’uomo che, in	... è la figlia ...	

		piedi accanto a lei carezza il neonato. (Fine flashback)		
22	1"	L'uomo sta guidando l'auto. Viene ripreso il dettaglio delle sue mani sul volante.	... della mia attuale ...	
23	1"	L'auto rallenta e accosta davanti all'edificio dal quale esce la bambina in tutù. La bambina cammina verso l'auto.	... moglie.	
24	1"	L'auto in movimento. A un certo punto comincia ad andare a marcia indietro.	Insieme andiamo a recuperare i due gemelli ...	
25	1"	I due gemelli sul marciapiedi litigano tra loro.	... che avevo dimenticato.	
26	1.5"	Bambini in auto: i gemelli e la bambina nei sedili posteriori. Nel sedile anteriore, accanto al conducente, c'è un nuovo soggetto. Un ragazzino che si gira verso il conducente e lo guarda.	Poi prendo Mattia, che ho appena saputo che è mio figlio.	
27	1"	Auto in movimento	Insieme ...	
28	1"	La donna aspetta davanti casa l'auto che arriva. La donna saluta sorridendo.	... torniamo ...	
29	1.5"	Auto parcheggiata davanti alla casa. La donna si avvicina all'auto dalla quale scendono i bambini.	... tutti a casa, prima di ...	

30	1"	L'uomo in auto guarda lo specchietto retrovisore.	... riaccompagnare Arturo, ...	
31	1"	La mano dell'uomo sposta lo specchietto retrovisore. Lo specchio riflette l'immagine di un altro bambino seduto nel sedile posteriore che sbadiglia.	... il figlio del vicino."	
32	3"	L'auto di nuovo in movimento. In alto appare l'headline: "New Renault Scènic. Facciamo posto a tutte le famiglie."	Donna (voice over): "New Renault Scènic. Facciamo posto a tutte le famiglie."	
33	2.5"	Sfondo nero sul quale appare il logo Renault.		Brano musicale s'interrompe. Breve jingle accompagna l'apparizione del logo.
34	5"	Sfondo nero. Informazioni tecniche e commerciali relative all'auto pubblicizzata.	Donna (voice over): "Tua con 7 posti, clima, cerchi in lega e radio mp3. A 19.950 euro."	Riprende il brano musicale dello spot, fino alla fine.

L'utilità tecnica di questa scomposizione in sequenze è rappresentata dalla possibilità di selezionare ed escludere le scene relative alla mera pubblicizzazione del prodotto. Inoltre in NVivo è possibile introdurre le due schede di scomposizione e lavorare su queste anziché sul video degli spot.

STEP 2 – L'IMPORTANZA DI CIO' CHE NON SI VEDE

Sebbene solo alcuni degli output realizzati con NVivo entrino a far parte dell'articolo o della comunicazione in oggetto, ciò non vuol dire che alcuni step si possano saltare. La codifica e quindi la creazione di etichette all'interno delle quali racchiudere i concetti che emergono dalle immagini è essenziale e inevitabile per la costruzione e l'esplicitazione del senso racchiuso in esse.

A tal proposito si ricorda che l'analisi concettuale – ovvero il momento di emersione dei concetti dal testo o dalle immagini - si concentra sui temi emersi, prevedendo:

- la creazione e l'applicazione dei codici,
- il tipo di struttura da attribuire a essi, ovvero se gerarchica o libera,
- la fusione di più codici all'interno di uno solo,
- l'estrapolazione della porzione di testo cui il codice è stato applicato,
- l'individuazione della frequenza dei codici.

Ultimati tali passaggi è possibile dedicarsi a un'analisi di tipo relazionale.

STEP 3 – LE RELAZIONI E I MODELLI

Un ulteriore livello di analisi è rappresentato dallo studio di relazioni tra i concetti. Questo è possibile anche mediante il ricorso a i network: supporti grafici che esplicitano le relazioni poste in essere. I concetti e/o le categorie emerse durante il lavoro di analisi possono trovare una loro collocazione all'interno di un network, il quale riproduce una relazione fra gli elementi selezionati. Un network sintetizza graficamente un insieme di concetti e le relazioni che emergono tra gli stessi. Rappresenta, quindi, una mappa dei contenuti di un testo o di una collezione di testi posti in analisi.

I concetti, le relazioni fra gli stessi e i paragrafi o le porzioni di testo dai quali sono estrapolati rappresentano gli elementi base di composizione del network concettuale. Nella rappresentazione grafica il concetto identifica il cuore del network, è quindi l'elemento centrale nella visualizzazione.

La relazione tra due concetti è espressa attraverso un link, spesso una linea o una freccia. Nel network concettuale le relazioni rappresentano una sorta di «mi-

sura» della forza, dell'intensità di un concetto; indicando il grado di enfasi posta su un aspetto piuttosto che su di un altro.

Per creare una relazione con NVivo occorre attivare da **Classifications/Relationships Types** quindi dalla barra degli strumenti **New/Relationship Types in This Folder** e dalla finestra *New Relationship Type* attribuiamo il nome alla relazione che si sta per strutturare.

Nel caso delle relazioni create nei paragrafi precedenti si tratta di relazioni fra **Sets**. Un sets è un insieme di nodi di codici raggruppati all'interno di una unica categoria, il set appunto.

Per esempio, nell'analisi del primo spot è stato creato il set "l'amore come elemento distintivo", questo tratto è caratteristico e rappresenta il legame fra gli altri due sets all'interno dei quali sono stati raggruppati i restanti codici. Si tratta dei sets "ritratto di coppia" e "scene di vita domestica".

Le operazioni svolte in questo caso sono consistite nel:

- raggruppare i codici in sets,
- individuare la natura della relazione,
- generare la relazione.

La relazione creata secondo questi passaggi può essere resa graficamente attraverso i **Models**, ovvero mediante l'esplicitazione grafica della stessa.

Se, infatti, le procedure di generazione della relazione sono state eseguite correttamente è possibile in pochi passaggi giungere al modello.

Da **Models/New Model in This Folder** si crea il modello. Nel pannello di navigazione in basso si apre una finestra di disegno, mediante l'utilizzo del tasto destro del mouse si seleziona *Add Project Items* e si richiama la relazione che si vuole visualizzare.

RIFERIMENTI BIBLIOGRAFICI

- ARDONE R.G. (1990) *Rappresentazioni familiari: la famiglia nelle teorie scientifiche e nel senso comune*, Roma, Borla.
- BAGNASCO A., BARBAGLI M., CAVALLI A. (2004) *Elementi di sociologia*, Bologna, il Mulino.
- BARBAGLI M. (1984), *Sotto lo stesso tetto. Mutamenti della famiglia in Italia dal XV al XX secolo*, Milano, Il Mulino.
- BERNARDINI I. (1995) *Finché vita non ci separi*, Milano, Rizzoli.
- BIANCHI C. (2007) *Analisi semiotica dell'audiovisivo pubblicitario*, Roma, Carocci.
- FRANCESCATO D. (1992) *Quando L'amore finisce*, Bologna, Il Mulino.
- FRANCESCATO D. (1994) *Figli sereni di amori smarriti: ragazzi e adulti dopo la separazione*, Milano, Arnoldo Mondadori Editore.

- FLOCH J.-M. (2002) *Semiotica Marketing e Comunicazione. Dietro i Segni, le Strategie*, Milano: FrancoAngeli.
- MAZZONI S. (1995) “Le famiglie ricostituite: considerazioni generali e proposte di intervento”, in M. MALAGOLI TOGLIATTI, G. MONTANARI G (a cura di), *Famiglie divise*, Milano, Franco Angeli.
- MEULDERS-KLEIN M. T., THÉRY I. (1993) *Les recompositions familiales aujourd'hui*, Paris, Nathan.
- SARACENO C., NARDINI M. (2001) *Sociologia della famiglia*, Bologn.; Il Mulino.
- ZANATTA A.L. (2008) *Le Nuove Famiglie*, Bologna, Il Mulino.

7.

LE CANZONI DEL FESTIVAL DI SANREMO: LESSICO E ANALISI DEL CONTENUTO

7.1. INTRODUZIONE

Dal 1951 al 2010 nella selezione ufficiale del Festival di Sanremo sono state cantate più di 1.750 canzoni. I testi di queste canzoni rappresentano al meglio il codice della tradizione canzonettistica popolare. In sessant'anni di storia le parole delle canzoni rispecchiano i mutamenti profondi della società italiana. Basta scorrere anche velocemente i titoli delle canzoni per rendersene conto: da *Al mercato di Pizzighettone* (1951) a *Su questa panchina* (2010). Non si tratta di cambiamenti superficiali. Dal 1951 ad oggi sono cambiati i valori sociali, i comportamenti, le strutture socio-economiche, le istituzioni politiche e, di conseguenza, anche i linguaggi, gli strumenti espressivi, gli apparati simbolici.

Le canzoni di Sanremo, rispetto al complesso delle canzoni italiane, risentono più lentamente di questi cambiamenti: si tratta pur sempre di una manifestazione nazional-popolare, con un forte impatto mediatico monopolizzato dalla RAI (con alterne fortune di pubblico), sulla quale si esercitano, talvolta, censure o auto-censure che frenano la risonanza immediata dei movimenti sismici che avvengono nel mondo della canzone e nel Paese (Borgna, 1998; Vesigna, 2000; Gennaccari e Maffei 2008). Tuttavia, quando uno studioso attento e documentato come Giuseppe Antonelli deve sottolineare i punti di svolta di “mezzo secolo di italiano cantato” non può evitare di citare *Nel blu, dipinto di blu* di Domenico Modugno (1958), *L'italiano* di Toto Cotugno (1983) e *La paranza* di Daniele Silvestri (2007), non senza dimenticare la voce critica di Luigi Tenco e il tragico cono d'ombra del suo suicidio del 1967 (Antonelli, 2010, pp. 9-29).

Lorenzo Coveri, seguendo le indicazioni di Gianni Borgna, storico della canzone italiana (Borgna, 1985), in uno studio del 1992 individuava tre periodi distinti di sviluppo linguistico della canzone: il periodo che precede *Nel blu, dipinto di blu* (1958); il secondo periodo che si conclude con il successo dei cantautori alla fine degli anni Settanta (in particolare la coppia Mogol-Battisti); il terzo periodo degli anni Ottanta (Coveri, 1996, p. 15; Jachia, 1998, p. 12).

7.2. IPOTESI DI LAVORO

Questo lavoro è rivolto a un'analisi essenzialmente quantitativa dei testi delle canzoni e assume come riferimento le osservazioni più significative compiute da Giuseppe Antonelli (2010), per quanto riguarda il lessico, e da Giovanni Di Franco e Stefano Nobile (1995), per i contenuti.

Una prima pista da seguire è quella della periodizzazione. I tre periodi di sviluppo della canzone italiana individuate da Coveri sono ritracciabili nella struttura dei profili lessicali che emerge da un'analisi automatica? In caso di risposta affermativa, quali sono le linee interpretative di questi periodi?

Una seconda pista, sempre basata su un'analisi del lessico, si ispira ai vincoli fonetici imposti dalla musica che richiedono spesso, in conclusione di una frase musicale che termina in battere, una parola tronca con l'accento sull'ultima sillaba: libertà, caffè, tivvù, ma anche i dittonghi noi, avrai, ecc. Questo comporta da parte dei parolieri delle canzoni, come vedremo, scelte lessicali che hanno dei risvolti sul contenuto perché nella lingua italiana le forme tronche sono meno frequenti che nel francese e nell'inglese.

Una terza pista riguarda la scomparsa di parole molto in voga, la comparsa di parole nuove e la persistenza di parole che sopravvivono quasi solo nelle canzoni indipendentemente dalla loro utilizzazione nella linguaggio quotidiano.

Infine vi sono i contenuti più propriamente veicolati dal cambiamento sociale, dai cambiamenti di mentalità, dei rapporti interpersonali e del "clima" culturale complessivo.

7.3. IL CORPUS SANREMO2010

Il corpus Sanremo2010 (Step 1) è costituito da 1.758 testi di canzoni (1.792 Kb) presentate al Festival dalla prima edizione del 1951 fino al 2010.

Il corpus, sottoposto a normalizzazione in TaL-TaC, presenta le misure lessico metriche riportate in tabella 7.1. La fascia di alta frequenza contiene 55

forme grafiche (tab. 7.2), tra le quali, se escludiamo le forme grammaticali, troviamo le due parole più tradizionali della canzone italiana in rima: *cuore* e *amore*. Complessivamente i primi venti sostantivi ci permettono subito di delineare un quadro di sostanziale fedeltà della canzone italiana alla tradizione.

Tab. 7.1 – Misure lessicometriche del corpus Sanremo2010

Occorrenze	N	341.753
Forme grafiche	V	16.757
Type/Token ratio	(V/N)*100	4,90
Percentuale di hapax	(V1/V)*100	44,02
Frequenza media generale	N/V	20,40

Massimo Arcangeli e Luca Piroddi (2010), in un articolo della *Repubblica* che precedeva il Festival 2010, ricordavano la canzone di Valeria Rossi del 2001 (ma non presentata al Festival) dal titolo *Dammi tre parole: sole, cuore, amore*. Le parole delle canzoni cambiano, naturalmente, ma non quelle di maggiore frequenza.

Tab. 7.2 – Primi dieci sostantivi appartenenti alla fascia di alta e media frequenza nel corpus Sanremo2010

Forme	Occorrenze	Rango	Fascia
amore	2.835	21	Alta
cuore	1.192	43	Alta
vita	3.635	37	Alta
mondo	888	59	Media
occhi	790	72	Media
cosa	714	76	Media
sole	704	77	Media
giorno	698	78	Media
notte	687	80	Media
cielo	628	86	Media
tempo	603	88	Media
mare	567	91	Media
amor	429	115	Media
sera	368	127	Media
vento	347	132	Media

uomo	334	138	Media
parole	325	140	Media
mani	311	141	Media
sogni	311	141	Media
casa	303	146	Media

I due autori rilevavano anche la mancata presenza, tra le forme più significative, del lemma <corpo> (67), imputandola ad una sorta di idiosincrasia del Festival verso la fisicità. Vero, ma sarebbe bene considerare anche la presenza molto massiccia delle “parti del corpo” (con uso anche metaforico e metonimico) tra le quali spicca una parola di uso recente come *culo*, attestata per la prima volta nel 1992 (tab. 7.3).

Tab. 7.3 – Sostantivi che si riferiscono a parti del corpo nel corpus Sanremo2010

Forme	Occorrenze	Rango	Fascia
cuore	1.192	43	Media
occhi	790	72	Media
mani	311	141	Media
mano	281	162	Media
viso	134	305	Media
bocca	128	325	Media
braccia	114	350	Media
capelli	102	385	Media
pelle	98	397	Bassa
piedi	79	475	Bassa
labbra	67	545	Bassa
corpo	59	617	Bassa
cuori	58	629	Bassa
dita	43	810	Bassa
spalle	39	865	Bassa
gola	36	902	Bassa
sangue	33	973	Bassa
naso	31	1.025	Bassa
petto	29	1.073	Bassa

seno	25	1.196	Bassa
denti	19	1.451	Bassa
vene	19	1.451	Bassa
braccio	18	1.510	Bassa
culo	16	1.658	Bassa
dito	14	1.482	Bassa
gambe	15	1.743	Bassa
seni	11	2.228	Bassa

7.4. LE FASI DI SVILUPPO DEL LESSICO DELLE CANZONI DI SANREMO

In sessant'anni i cambiamenti nel lessico delle canzoni sono del tutto scontati. Meno scontato è riscontrarne le forme e i contenuti. Inoltre ci si può chiedere se questi cambiamenti, rilevati spesso dalla critica musicale o da osservazioni puntuali compiute da studiosi di linguistica su base qualitativa, trovano una conferma o meno attraverso un'analisi puramente quantitativa che non sia influenzata in alcun modo dalle scelte soggettive dell'analista. L'analisi delle corrispondenze sui profili lessicali (Step 2) può ben rispondere a questo obiettivo mettendo in evidenza, nella proiezione sui piani fattoriali delle frequenze delle parole, le somiglianze e le differenze tra i profili lessicali di ciascun anno. L'unica condizione che si pone a questo trattamento è che le occorrenze di ciascuna parola della matrice individui (parole) per testi (anni del Festival) risponda a un criterio di soglia minima sulla somma marginale di riga pari ad almeno 12 occorrenze (la soglia indicata come ottimale da TaLTaC nella fase di calcolo delle misure lessicometriche che corrisponde a una copertura del testo pari all'88,91%).

*Tab. 7.4 – Analisi delle corrispondenze sulla tabella lessicale
(soglia $N \geq 12$) del corpus Sanremo2010: traccia 1,26 (primi tre fattori)*

Fattore	Autovalore	% di inerzia	% cumulata
1	0,0724	5,71	5,71
2	0,0402	3,17	8,87
3	0,0342	2,70	11,57
...	...	...	...

Tab. 7.5 – Analisi delle corrispondenze sulla tabella lessicale
del corpus Sanremo2010: coordinate e contributi delle frequenze attive

Anno	Coordinate		Contributi assoluti		Contributi relativi	
	F1	F2	F1	F2	F1	F2
1951	0,77	0,31	5,2	1,5	0,22	0,04
1952	0,82	0,29	7,6	1,8	0,23	0,03
1953	0,70	0,28	5,7	1,6	0,18	0,03
1954	0,77	0,41	7,4	3,9	0,25	0,07
1955	0,71	0,22	5,5	1,0	0,21	0,02
1956	0,82	0,25	8,4	1,3	0,28	0,02
1957	0,74	0,15	9,2	0,7	0,30	0,01
1958	0,78	0,15	7,8	0,5	0,29	0,01
1959	0,76	-0,01	6,0	0,0	0,20	0,00
1960	0,66	0,00	4,3	0,0	0,15	0,00
1961	0,55	0,00	3,7	0,0	0,15	0,00
1962	0,49	0,00	3,4	0,0	0,11	0,00
1963	0,44	-0,18	2,0	0,6	0,06	0,01
1964	0,38	-0,48	2,2	6,2	0,07	0,10
1965	0,12	-0,44	0,2	4,8	0,01	0,10
1966	0,01	-0,41	0,0	4,8	0,00	0,10
1967	0,02	-0,26	0,0	2,4	0,00	0,05
1968	0,08	-0,22	0,1	1,3	0,00	0,03
1969	0,17	-0,30	0,5	2,8	0,02	0,06
1970	0,15	-0,57	0,4	11,4	0,01	0,16
1971	0,16	-0,36	0,5	4,5	0,01	0,06
1972	0,11	-0,30	0,3	3,5	0,01	0,07
1973	0,02	-0,24	0,0	2,3	0,00	0,04
1974	0,05	-0,20	0,1	1,4	0,00	0,03
1975	0,07	-0,23	0,1	2,1	0,00	0,05
1976	-0,02	-0,18	0,0	1,1	0,00	0,03
1977	-0,04	-0,23	0,0	0,8	0,00	0,02
1978	-0,08	-0,18	0,1	0,5	0,00	0,01
1979	0,04	-0,06	0,0	0,1	0,00	0,00
1980	-0,04	-0,14	0,0	0,9	0,00	0,02
1981	-0,15	-0,09	0,4	0,3	0,01	0,01
1982	-0,10	0,03	0,2	0,0	0,01	0,00
1983	-0,09	-0,03	0,2	0,1	0,01	0,00
1984	-0,04	0,05	0,0	0,1	0,00	0,00

1985	-0,13	0,04	0,5	0,1	0,02	0,00
1986	-0,09	0,10	0,2	0,5	0,01	0,01
1987	-0,09	0,04	0,2	0,1	0,01	0,00
1988	-0,12	0,08	0,5	0,4	0,02	0,01
1989	-0,18	0,12	1,3	1,1	0,03	0,02
1990	-0,14	0,23	0,6	2,9	0,02	0,05
1991	-0,13	0,30	0,5	5,2	0,02	0,08
1992	-0,16	0,21	1,0	3,1	0,03	0,06
1993	-0,20	0,46	1,6	15,5	0,04	0,20
1994	-0,15	0,21	0,9	3,1	0,03	0,05
1995	-0,15	-0,03	1,0	0,0	0,03	0,00
1996	-0,16	0,03	0,9	0,1	0,04	0,00
1997	-0,18	-0,06	1,3	0,3	0,05	0,01
1998	-0,22	0,05	1,3	0,1	0,04	0,00
1999	-0,20	-0,13	1,0	0,7	0,03	0,01
2000	-0,11	-0,02	0,3	0,0	0,02	0,00
2001	-0,14	-0,01	0,6	0,0	0,02	0,00
2002	-0,14	-0,04	0,7	0,1	0,02	0,00
2003	-0,16	0,04	0,8	0,1	0,02	0,00
2004	-0,14	-0,08	0,4	0,2	0,02	0,00
2005	-0,17	-0,07	0,8	0,3	0,03	0,00
2006	-0,17	0,02	0,7	0,0	0,03	0,00
2007	-0,12	0,15	0,5	1,3	0,02	0,02
2008	-0,10	0,04	0,3	0,1	0,01	0,00
2009	-0,08	-0,06	0,2	0,2	0,01	0,00
2010	-0,08	0,06	0,2	0,1	0,01	0,00

Il risultato dell'analisi delle corrispondenze (fig. 7.1) conferma solo in parte la periodizzazione proposta. La proiezione sul piano fattoriale delle frequenze attive che risultano dalla partizione per anno mette bene in evidenza la presenza di tre periodi, ma il punto di passaggio innovativo dal primo al secondo periodo non sembrerebbe affatto essere il 1958 quanto piuttosto 1964.

D'altra parte, se è vero che il 1958 è l'anno di *Nel blu dipinto di blu* (Modugno e Migliacci) considerata da tutti come la canzone italiana più significativa di quegli anni, tra il 1961 e il 1964 si affacciano sulla scena di Sanremo i cantanti e i cantautori che faranno la fortuna della canzone italiana negli anni tra la fine degli anni Sessanta e i primi degli anni Settanta: Adriano Celentano, Giorgio Gaber, Gino Paoli, Sergio Endrigo e, soprattutto, quel grande paroliere che è stato Giulio Rapetti, meglio noto come Mogol, che tra il 1964 e il 1967 firma le parole di ben nove canzoni, di cui quattro nel 1965.

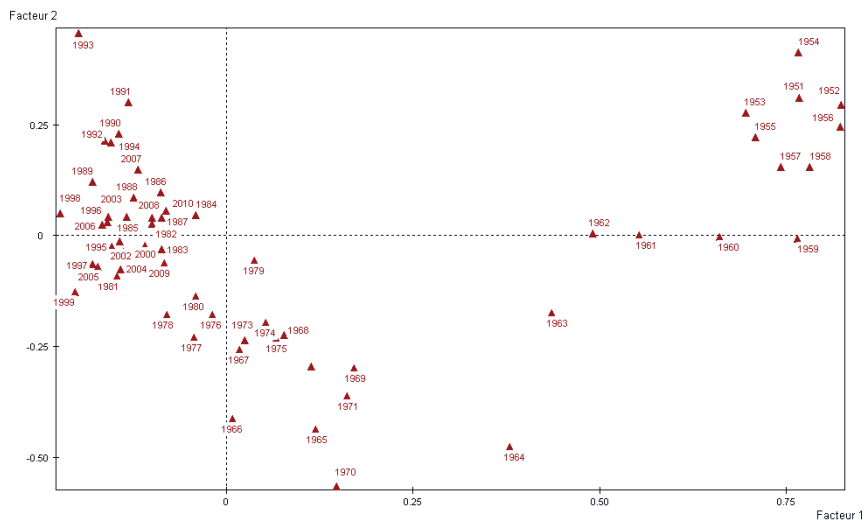


Fig. 7.1 – Analisi delle corrispondenze della tabella lessicale Sanremo2010 forme per testi: rappresentazione delle frequenze attive sul piano fattoriale F1-F2.

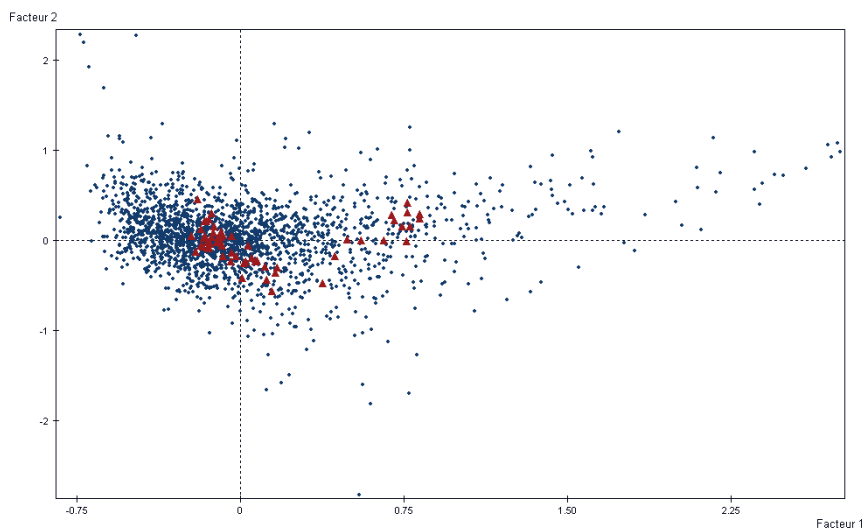


Fig. 7.2 – Analisi delle corrispondenze della tabella lessicale Sanremo2010 forme per testi: rappresentazione delle forme (soglia $N \geq 12$) sul piano fattoriale F1-F2

Sul piano fattoriale è lungo il primo asse che si delinea un andamento temporale che rende molto simili i profili lessicali degli anni Cinquanta in netta contrapposizione con gli anni Novanta (fig. 7.2). Possiamo anche osservare come l'ultimo decennio non presenti caratteristiche autonome che non siano già definite in precedenza.

Il cambiamento più significativo nel corso degli anni sembrerebbe quello imposto dalla fonetica per rispettare le modalità ritmiche della musica: i periodi musicali si concludono spesso con un tempo “forte” e questo impone al verso una terminazione tronca (Antonelli, 2010, p. 36). Negli anni Cinquanta è tradizionalmente accettato il troncamento delle parole in *l*, *m*, *n* e soprattutto in *r* (*amore/amor*, *cuore/cuor*).

Tab. 7.6 – Parole troncate nel corpus Sanremo2010 secondo il periodo

Forme	1951-1964 Occorrenze	1965-1980 Occorrenze	1981-2010 Occorrenze
sol	43	0	7
ciel	35	1	0
siam	4	12	7
battiam	10	0	0
pian	10	4	7
canzon	15	0	0
lontan	15	0	0
amor	341	70	18
cuor	169	15	5
fior	35	1	0
dir	21	6	8
sognar	26	2	5
mar	24	1	3
cantar	24	1	2
parlar	16	7	1
allor	19	3	0
dolor	18	4	0
amar	19	0	0
morir	11	5	1

dimenticar	11	3	0
scordar	4	7	0
passar	11	0	0
guardar	9	0	2
tornar	10	0	0
color	6	2	2

Amor e *cuor*, come si vede, resistono bene ancora nel secondo periodo ma cadono quasi del tutto in disuso gli ultimi anni. I parolieri trovano altre soluzioni per l'accento obbligato utilizzando parole straniere (*your, forever, shalom, down*), onomatopee (*boom, tam, zum*) e nomi propri (*Jesabl, Omar*). Dalla metà degli anni Sessanta si preferisce fare ricorso alle parole accentate.

Tab. 7.7 – Sostantivi con accento finale nel corpus Sanremo2010 secondo il periodo (soglia N ≥ 50)

Forme	1951-1964 Occorrenze	1965-1980 Occorrenze	1981-2010 Occorrenze
città	23	40	188
felicità	58	29	74
verità	5	53	91
età	24	16	106
libertà	3	40	84
realtà	13	21	53
metà	0	13	66
eternità	13	8	47
pietà	5	11	39
caffè	4	10	48
dì	53	4	10

Una soluzione più persistente nel tempo per chiudere il verso con una sillaba tonica (Step 3; tab. 7.8) è quella di utilizzare alcune voci del futuro (*verrà, passerà, finirà, cambierà, sarò, farò, vivrò, andrò, piangerò, morirò*) e del passato remoto (*sentì, capì, morì, passò, trovò, tornò, lasciò, cantò*). L'uso del passato remoto, come rileva lo stesso Antonelli, è piuttosto datato; a partire dagli anni Settanta si affaccia l'uso del condizionale (*vorrei, avrei, potrei, farei*) che con l'uscita di dittongo è trattata come una forma tronca (Antonelli, 2010, p. 37).

Tab. 7.8 – Imprinting dei verbi nel corpus Sanremo2010 secondo il periodo (valori % sulle forme verbali non ambigue del periodo)

Modalità	1951-1964 Occorrenze	1965-1980 Occorrenze	1981-2010 Occorrenze
1-2-3 p. s.	67,5	74,5	67,8
cond. 1 p. s.	0,6	2,3	2,2
fut. 1 p. s.	4,1	4,3	4,1
fut. 2 p. s.	2,1	3,2	2,1
fut. 3 p. s.	3,9	4,9	4,2
pass. rem. 3 p.	3,4	1,1	0,6
passato	12,1	9,3	5,9
presente	77,1	77,5	82,2
futuro	10,5	12,9	11,6
Totale forme	6.627	11.321	36.026

La prima e la seconda persona singolare sono più frequenti, soprattutto nel secondo periodo. In parte anche questo risultato prospetta una soluzione al problema ritmico della musica: le forme pronominali tornano utili per chiudere il verso con l'accento: *te, tu, me*. In parte c'è sicuramente una accentuazione, marcata nel tempo, nell'uso del pronome di prima persona *io* che, sebbene non sia obbligatorio in italiano, nelle canzoni ricorre con insistenza (tab. 7.9).

Tab. 7.9 – Forme pronominali nel corpus Sanremo2010 secondo il periodo (valori ‰ sulle occorrenze del periodo)

Modalità	1951-1964 Occorrenze	1965-1980 Occorrenze	1981-2010 Occorrenze
mi	8,3	12,7	10,3
te	11,6	10,2	9,9
tu	11,3	13,9	9,0
ti	8,2	11,3	9,3
io	5,2	13,0	9,8
me	8,7	9,7	7,3
ci	4,9	3,3	6,5
c'	2,5	5,2	5,5
noi	1,2	3,7	4,9

lei	1,9	2,6	1,7
lui	1,2	2,1	0,6
t'	2,1	0,7	0,4
Totale occ.	43.606	69.336	228.811

Tuttavia va rilevata anche una crescente presenza dei pronomi personali plurali *ci* (*c'*) e *noi* che complessivamente passano da 8,6‰ a 16,9‰.

7.5. PAROLE VECCHIE E NUOVE

Parole che restano, parole che passano di moda e parole nuove che si affacciano sulla scena. Il ricambio è più ampio di quanto non sembri a prima vista. Su 16.757 forme 1.762 (il 10,5%) sono presenti solo nel periodo 1951-1964; 1.383 (pari all'8,3%) sono presenti solo nel periodo 1965-1980; 8.696 (pari al 60%) sono presenti solo nel periodo 1981-2010. La persistenza delle stesse forme grafiche per almeno due periodi è solo del 21,2% e per tutti e tre i periodi è del 12,2%.

Se escludiamo le forme troncate (*amar*, *lontan*, *canzon*, *passar*), i nomi propri, le forme flesse dei verbi e le onomatopее, le forme più significative esclusive di ciascun periodo sono indicate in tabella 7.10.

Tab. 7.10 – Parole presenti nel corpus Sanremo2010 esclusivamente nel periodo indicato

1951-1964		1965-1980		1981-2010	
Forme	Occ.	Forme	Occ.	Forme	Occ.
pepe	17	hippy	24	femmene	54
piccolina	16	talismano	13	femmina	40
trenino	13	forestiero	12	down	36
mandolino	12	miscugli	12	lazzo	28
magnifico	11	buena	11	darling	28
usignuolo	11	dios	11	mea	26
casetta	11	bluejeans	7	distante	25
parasole	9	tramontana	7	road	24
patatina	9	nocciolino	7	minore	22
laghetto	8	salsa	6	diversi	22
piccino	8	banane	6	marinai	22

lume	7	elenco	6	beguine	20
tamburino	7	sassofono	6	straniero	19
anellino	7	buchi	6	forever	18
cicogna	7	citofono	5	veri	18
trulli	7	trottola	5	sud	18
piccolissimo	7	beato	5	pianeta	17
omino	7	sacrosanta	5	maledetta	17
solitaria	6	zingara	4	capitano	17
furia	6	arca	4	possibilità	17
libellule	6	reparto	4	generale	16
fantastica	6	denari	4	bellissimo	16
mammina	6	filosofia	4	lividi	16
sfortuna	6	cesto	4	single	16
notturmo	6	frottole	4	culo	16
pallide	6	ragioniere	4	energia	16
papera	5	stupidamente	4	core	16
mandorlo	5	indifeso	4	rispetto	16
pianola	5	scarafaggi	4	swing	15
patata	5	fidanzata	3	benvenuto	15
paperina	5	molecole	3	avventure	15
cappellino	5	ottone	3	specialmente	14
chiesetta	5	malate	3	animale	14
firmamento	5	idolo	3	hands	14
torrente	5	mandorla	3	battiti	14
villa	5	rape	3	duri	14

Naturalmente in questa esplorazione dobbiamo tenere conto delle parole nuove che sono apparse dal 1965 in poi, anche se non esclusive di un periodo (tab. 7.11). Si può osservare come in questa lista emergano con forza parole che rimandano a contenuti sociali o comunque che denotano una forma di riflessione sull'impegno civile (*fame, pane, malattia, fatica*) e parole che rinviano a momenti critici o addirittura negativi della vita quotidiana (*rumore, dura, sbagliato, noia, paure, problemi, inferno*, ecc.)

Tab. 7.11 – Parole presenti nel corpus Sanremo esclusivamente dopo il 1964

Forme	1965-1980	1981-2010
pelle	12	86
idea	5	70
rumore	3	59
liberi	4	56
dura	13	45
futuro	4	53
sbagliato	24	33
muro	4	58
odore	1	48
noia	11	38
difficile	1	44
fiato	1	42
pezzi	5	32
radio	1	35
matto	13	25
paure	2	34
Sesso	6	29
spazio	6	29
problemi	8	26
gesto	1	31
inferno	1	31
muri	1	29
fame	5	25
santi	1	28
pane	8	20
ferita	2	26
malattia	6	22
fatica	7	21
nave	9	18
speciale	3	22
fragili	1	24

sicura	6	19
dubbi	1	23
novità	2	22
pazzia	3	21
brutto	2	21

E ancora, in questo elenco, non si possono trascurare le “brutte parole” (tab. 7.12) con occorrenze di piccola entità che non appaiono nelle tabelle precedenti pur essendo molto significative per la registrazione dei cambiamenti del linguaggio nelle canzoni del Festival: il turpiloquio, così consueto nel linguaggio quotidiano, è ormai legittimato anche nelle canzoni di Sanremo.

Tab. 7.12 – Forme di turpiloquio presenti nel corpus Sanremo2010

Forme	1981-1990	1991-2000	2001-2010
culo	0	2	14
puttana	1	0	4
cazzate	0	2	1
palle	1	2	0
stronza	0	0	3
culetto	0	2	0
cazzata	0	0	1
figata	1	0	0
puttane	0	1	0
coglione	0	0	1
coglioni	0	0	1
sputtanare	1	0	0
stronzi	0	1	0

La presenza delle “parolacce” nei testi della canzoni del Festival di Sanremo è una prova ulteriore di quel fenomeno di indebolimento dell’interdizione verbale cui Nora Galli de’ Paratesi si richiama quando scrive, a proposito del “disfemismo”, che esso “consiste non tanto nella caduta del divieto, ma piuttosto nel suo completo capovolgimento, cioè nell’uso voluto e dissonante dei termini interdetti” (Galli de’ Paratesi, 2009, p. 141). Questo cambiamento nelle strategie eufemistiche

del linguaggio (Galli de' Paratesi, 1964), profondo e significativo in tutti i campi, compreso il linguaggio politico, è da collegare, secondo Galli de' Paratesi, sia ai mutamenti sociali degli ultimi decenni che al diffondersi a tutti i livelli di pratiche di interazione informali che contribuiscono a ridurre ogni precauzione nell'uso di parole che, fino a qualche tempo fa, rientravano nell'ambito del proibito o dell'insulto. In questo, un ruolo importante lo ha giocato anche la retrocessione nell'uso del dialetto che non svolge più quella funzione di espressività esclusiva che gli era propria e che invece è filtrata direttamente nell'italiano parlato contemporaneo.

7.6. CONCLUSIONE

Nel paragrafo precedente abbiamo già individuato alcuni segnali forti di cambiamento nei temi trattati dalle canzoni. Le nuove parole che si affacciano sulla scena negli ultimi anni testimoniano un interesse crescente verso esperienze della vita quotidiana più vicine alla sensibilità del pubblico, specialmente giovanile, già rilevate da altri osservatori nello sviluppo culturale della canzone italiana (Di Franco e Nobile, 1995, p. 49 e sgg.). In particolare sono significativamente rappresentati, dopo il 1964, temi di carattere sociale e inquietudini del mondo contemporaneo che non erano di casa nelle "canzonette" (tab. 7.13).

Dalle parole delle canzoni emerge una certa inquietudine ben rappresentata da *sangue*, <sparare> e *violenza*, ma soprattutto da quelle 260 occorrenze del lemma <paura> nel periodo 1981-2010 che denotano il clima di incertezza e attesa di questi ultimi decenni. Tuttavia il tema dell'*amore* continua a dominare incontrastato (tab. 7.14) anche nel suo rimare più tradizionale con *cuore*.

L'analisi quantitativa del lessico di sessant'anni di canzoni del Festival di Sanremo ha messo in evidenza come i cambiamenti nel linguaggio siano più rilevanti nello stile che nei contenuti veri e propri. Certamente nei testi delle canzoni dal 1965 in poi tendono a scomparire i troncamenti, i diminutivi, i vezzeggiativi e si moltiplicano le contaminazioni linguistiche, soprattutto con l'inglese. Nell'ultimo decennio il registro dell'italiano "parlato" è sempre più presente, non solo per le tematiche affrontate quanto per l'apparire - timido ma significativo, come si è visto - delle "brutte parole".

Tab. 7.13 – Forme grafiche e lemmi presenti nel corpus Sanremo2010 che indicano temi di argomento socialmente rilevante

Forme e lemmi	Occ. 1951-1964	Occ. 1965-1980	Occ.1981-2010
<drogare>	1	20	2
fame	0	5	24
<guerra>	1	20	69
<malattia>	0	6	24
<lavorare>	3	21	15
lavoro	2	6	46
pace	5	21	68
<paura>	10	55	260
sangue	1	4	28
sciopero	0	6	3
solitudine	4	8	42
<sparare>	0	3	34
violenza	0	0	10

Tab. 7.14 – Lemmi presenti nel corpus Sanremo2010 che si riferiscono all'amore

Forme e lemmi	Occ. 1951-1964	Occ. 1965-1980	Occ.1981-2010
<amare>	59	50	275
<amore>	636	856	1.915
<baciare>	67	15	15
<bacio>	162	76	102
<cuore>	369	240	830
<innamorare>	20	45	123

L'italiano del Festival appare, per certi versi, contraddittorio: stabile e rinnovato nello stesso tempo. D'altra parte i testi delle canzoni devono seguire una fonetica che necessariamente deve essere piegata alla metrica e al ritmo, mentre nei contenuti, se vogliono avere successo, devono essere in sintonia con i gusti del pubblico. I testi delle canzoni in parte possono rispecchiare i cambiamenti sociali e il clima culturale da cui prendono vita ma, fondamentalmente, come tutti i "generi", risponde a regole proprie. Se guardiamo allo scorrere del fiume, all'impeto delle parole più frequenti, quello che ci appare è un mondo sempre uguale a se stesso: *Dammi tre parole: sole, cuore e amore*, cantava

Valeria Rossi. E tuttavia, per seguire le tracce del tempo, dobbiamo fare attenzione soprattutto alle parole meno comuni, meno frequenti, ma proprio per questo in grado di suggerire emozioni e immagini nuove. Dal punto di vista metodologico, però, dobbiamo essere consapevoli che così facendo ci stiamo muovendo in un terreno in cui qualità e quantità non si contrappongono e non perseguono il primato dell'una sull'altra, ma si integrano vicendevolmente piegandosi sempre alle nostre domande, alle nostre ipotesi, senza le quali ci muoveremmo nell'oscurità tra impalpabili e indecifrabili presenze.

STRATEGIA DI ANALISI

STEP 1 – ORGANIZZAZIONE, NORMALIZZAZIONE DEL CORPUS E MISURE LESSICOMETRICHE

Il corpus Sanremo2010 è costituito da 1.758 testi di canzoni pubblicati sul sito www.sanremostory.it²¹. Il sito è molto curato, sia nella parte storica che nei commenti. I testi delle canzoni sono trascritti fedelmente. Dal corpus sono state escluse cinque canzoni in lingua inglese presentate negli anni 1981, 1982, 1983 e 1984.

Le canzoni, come di consueto, sono state trattate con normalizzazione in TaLTaC con identificazione di 60 partizioni per gli anni. Il corpus in fase di normalizzazione non è stato lessicalizzato sulla base delle liste delle lucuzioni grammaticali e dei gruppi nominali per lasciare inalterate le forme grafiche ai fini della analisi delle forme tronche. Le misure lessicometriche registrano un'estensione ottimale del vocabolario (type/token ratio = 4,9) e una percentuale di hapax molto contenuta (44%).

STEP 2 – ESPORTAZIONE DELLA MATRICE PAROLE PER TESTI E ANALISI DELLE CORRISPONDENZE BINARIE

Anche in questo caso il trattamento con analisi delle corrispondenze binarie permette di esaminare le caratteristiche comuni dei profili lessicali al fine di

²¹ Il corpus è stato costruito con la collaborazione della dott.ssa Paola Sale per la sua tesi di laurea magistrale in Comunicazione della conoscenza per le imprese e le organizzazioni.

mettere in evidenza omogeneità e differenze tra i testi nel corso degli anni. Dopo aver calcolato le sub-occorrenze della variabile anni dal menu **Analisi – Pre-trattamento** e selezionando la voce *Calcolo delle sub-occorrenze*, con la lista *Vocabolario* della sessione aperta nella finestra di lavoro selezioniamo con il cursore il campo “Occorrenze totali”. A questo punto lo strumento **Text/Data Mining (TDM)** ci permette di selezionare nel vocabolario le voci con occorrenze maggiori o uguali a 12 che è la soglia indicata come ottimale da Taltac nella fase di calcolo delle misure lessicometriche. Con questa soglia la matrice di 2.336 forme che ne risulta, pari a una copertura del testo dell’88,91%, viene esportata dal menu **File – Esporta – In un file di testo** selezionando *Solo i record visibili*. La matrice forme (2.336) per testi (60) è importata in SPAD e trattata con analisi delle corrispondenze binarie.

STEP3 – RICONOSCIMENTO DELLE FORME GRAMMATICALI E IMPRINTING

La procedura di classificazione delle forme grafiche nelle categorie grammaticali e preliminare a qualsiasi altra operazione di classificazione delle parole. Dalla barra degli strumenti cliccando sul bottone **Tagging grammaticale** si accede alla finestra di dialogo. Il tagging è eseguito in forma completa, comprese le opzioni basate su criteri morfologici. Sul *Vocabolario con tagging grammaticale*, dal menu *Calcola*, selezionando *Imprinting*, si ottiene una tabella di tutte le caratteristiche morfologiche del vocabolario: lunghezza delle parole; categorie grammaticali e forme flesse, fino alle voci dei verbi e alla classificazione delle forme pronominali.

RIFERIMENTI BIBLIOGRAFICI

- ANTONELLI G. (2010) *Ma cosa vuoi che sia una canzone. Mezzo secolo di italiano cantato*, Bologna, il Mulino.
- ARCANGELI M. E PIRODDI L. (2010) “Sul palco sessant’anni di parole: a ben guardare, sempre le stesse”, *La Repubblica*, 13 febbraio.
- BORGNA G. (1985) *Storia della canzone italiana*. Roma-Bari, Laterza.
- BORGNA G. (1998) *L’Italia di Sanremo. Cinquant’anni di canzoni, cinquant’anni della nostra storia*, Mondadori, Milano.
- COVERI L. (1996) “Per una storia linguistica della canzone italiana”, in *Parole in Musica. Lingua e poesia nella Canzone d’autore italiana. Saggi critici e antologia di testi*, a cura di L. Coveri, Novara, Interlinea, pp. 13-24.

- DI FRANCO G., NOBILE S. (1995) *L'Italia che si dispera e l'Italia che si innamora. Temi, valori e linguaggi in 25 anni di canzone italiana*, Paper editrice, Roma.
- GALLI DE' PARATESI N. (1964) *Semantica dell'eufemismo. L'eufemismo e l'interdizione verbale nell'italiano contemporaneo*. Torino, Giappichelli.
- GALLI DE' PARATESI N. (2009) "Eufemismo e disfemismo nel linguaggio politico e nell'italiano di oggi", in *Synergie Italie*, numero speciale, pp. 137-144.
- GENNACCARI F., MAFFEI M. (2008) *Sanremo è Sanremo. I retroscena del Festival dal 1951 al 2007*, Curcio Musica, Milano.
- JACHIA P. (1998) *La canzone d'autore italiana 1958-1997. Avventure della parola cantata*, Milano, Feltrinelli.
- VESIGNA, G. (2000) *Sanremo è sempre Sanremo*. Sperling e Kupfer, Milano.

8.

SELF-REMEDIATION: L'IMPORTANZA DI PARTECIPARE AL GRANDE FRATELLO

8.1. INTRODUZIONE

Il processo di costruzione del sé è senz'altro oggetto di numerosi approcci di studio – da quelli di tipo sociologico e psicologico fino a quelli letterari. In questo lavoro si analizza il processo di “rimediazione” del sé che gli aspiranti concorrenti di un *reality show*, il *Grande Fratello* ²², pongono in essere nel momento in cui pensano di poter partecipare alla trasmissione.

La ricerca della partecipazione televisiva come elemento distintivo per la propria biografia ricalca la funzione della *ipseité* delineata da P. Ricoeur (1990). La partecipazione al GF sarebbe quindi un modo per costruire la propria identità come singolarità o individualità. Questi giovani sanno o credono di sapere di avere qualcosa di “particolare”, hanno individuato quale potrebbe essere la loro “stella” – il loro elemento distintivo – ma per poterla fare risplendere o, meglio ancora, accendere è necessario che una telecamera accenda – prima – una “lucetta rossa” su di loro. Ricoeur (1990) definisce questa *ipseité* come l'identità narrativa, in grado di problematizzare il rapporto tra due diversi tipi di sé: da un lato il carattere – che indica un sé che esiste nel tempo e nello spazio –, dall'altro la parola data, che connota un sé che esiste nel tempo e nello spazio di una relazione. L'identità narrativa diventa un ponte tra l'*idem* e l'*ipse*: il carattere e il mantenimento del sé. “Il carattere è quindi il ‘che cosa del chi’, è una sedimentazione, che però il racconto può riaprire” (Besozzi, 2006, p. 153) e anche rivendicare.

²² D'ora in avanti GF.

Rafforza tale asserzione il lavoro sul sé rimediato proposta da J.D. Bolter e R. Grusin (1999), che ci segnalano come si stia manifestando la tendenza a vedere noi stessi all'interno dei media di cui possiamo disporre, finendo per utilizzare questi ultimi anche per definire la nostra identità culturale e farne una necessità biografica.

Il medium che qui viene preso in considerazione è la televisione. Già J. Meyrowitz (1993) ha sostenuto che la televisione è un'arena sociale il cui impatto sul pubblico è di più ampia portata rispetto a quello dei media a stampa, perché la televisione essendo basata sulle immagini è di più facile accesso e comprensione raggiungendo, così, un vasto pubblico. Essendo un medium visivo non richiede particolari abilità per essere interpretata, né ne coltiva di particolari; così ci suggerisce D. Buckingham (2004) quando ripropone la tesi di N. Postman (1983) in merito al rapporto tra infanzia e televisione. Quello sviluppato da Postman non è che uno degli approcci meno ottimisti sull'argomento; oggi le teorie evolutive non vedono più il bambino come spettatore passivo anzi, ne sottolineano le capacità di lettura del testo televisivo e il suo uso per creare interazione con i fratelli o i compagni di gioco (Lemish, 2007).

Prescindendo dalla visione – apocalittica o integrata – che si decida di adottare per guardare a questo fenomeno, quello che qui preme è indicare l'influenza che su una fascia giovanile può avere la realtà televisiva, tanto da indurre questi giovani a ricercare l'esperienza mediata.

8.2. LA RICERCA

Le interviste che qui si analizzano provengono da una ricerca condotta tra giugno e agosto 2009, periodo durante il quale si sono svolte le selezioni per la decima edizione del Grande Fratello. Per analizzare le motivazioni sottese alla partecipazione al GF sono stati sentiti 120 soggetti, ripartiti uniformemente fra Nord, Centro e Sud del nostro Paese. A riprova che a queste selezioni – e di conseguenza a questo tipo di format – è interessato un ventaglio di pubblico con un'età assai diversa si hanno i 48 anni (età massima) di un provinante di sesso maschile e l'età minima di una provinante di 18 anni di Latina. L'età media invece ricade tra i 28 anni.

La traccia d'intervista utilizzata scaturisce, per alcuni suoi punti, dall'osservazione partecipante e dall'ascolto partecipato del “chiacchiericcio” percepito durante una prima fase di osservazione che ha preceduto quella delle interviste. La traccia si compone di quattro aree.

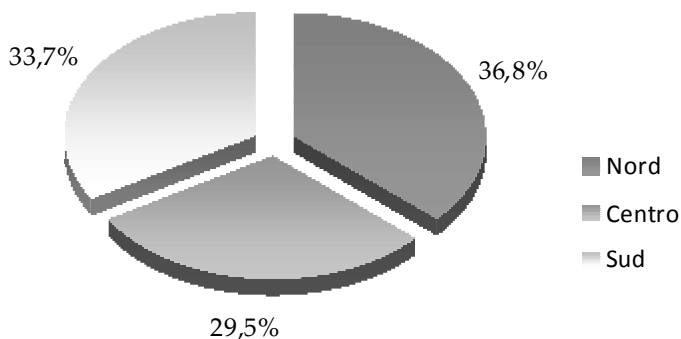


Fig. 8.1 – Distribuzione degli intervistati per zona geografica

- La presentazione di sé: in questa prima parte si chiede all'intervistato di raccontare quello che fa nella vita (attività di studio, lavorative, ecc.); il grado d'istruzione e l'età. Per comprendere quali siano le attitudini dell'intervistato, si procede chiedendogli cosa potrebbe fare di pratico all'interno della casa per il mantenimento della stessa (per esempio: cucinare, pulire, ecc.) e come potrebbe distinguersi dagli altri concorrenti al suo interno. A questo punto, si prosegue facendosi raccontare se ha visto le passate edizioni del programma e se c'è qualcuno dei precedenti concorrenti con cui si identifica per aspetto fisico o comportamento o che gli è piaciuto maggiormente rispetto agli altri.
- La motivazione: quest'area rappresenta il cuore del colloquio²³. All'intervistato è chiesto di spiegare che cosa rappresenti per lui in prima istanza il provino, secondariamente la casa, e poi la possibilità di entrarvi. Quindi gli si chiede se immagina per se stesso un possibile futuro nel mondo dello spettacolo e in che veste. Si indaga, inoltre, se costui abbia già acquisito esperienze nel settore dello spettacolo oppure no. Un rilancio importante in quest'area è la richiesta di cosa l'intervistato voglia trasmettere di sé – mediante la televisione – al resto del pubblico. È qui che l'intervista, a volte, si trasforma in colloquio clinico.
- L'abbigliamento: si chiede una riflessione sul proprio look e su quello degli altri.

²³ Il modo di interrogazione scelto nel porsi nei confronti dei provinanti è l'intervista semi-strutturata condotta con strategia non-direttiva (cfr. Bichi, 2007). Sono state previste una serie di domande aperte all'interno delle quali l'intervistatrice ha introdotto degli interventi, degli stimoli, dei rilanci, pertinenti con la conversazione posta in essere con l'intervistato.

- L'amore: è ormai noto che all'interno della "casa" nascono delle storie sentimentali e che l'entrarvi provochi crisi nelle coppie preesistenti a essa. Partendo dalla domanda sulla propria condizione sentimentale (single, fidanzato/a, coniugato, convivente, ecc.) si chiede a chi ha già un partner se quest'ultimo è a conoscenza della partecipazione al provino e che cosa ne pensi, quindi s'ipotizza al provinante la possibilità di lasciarsi andare a una storia d'amore all'interno della casa (nonostante si abbia già una relazione al di fuori di essa) e cosa accadrebbe alla sua attuale situazione sentimentale. Parimenti ai single è chiesto se un'altra motivazione sottesa alla partecipazione al GF è la ricerca di un'avventura sentimentale e in che modo si lascerebbero andare a essa.

8.3. I SIGNIFICATI DELLA PARTECIPAZIONE AL GRANDE FRATELLO

Obiettivo di questo lavoro è analizzare la motivazione che spinge giovani e adulti a mettersi in fila e aspettare per essere provinati ed entrare nella "casa" più famosa d'Italia.

Dalle codifiche effettuate sui testi appare già che esistano gruppi di motivazioni legate alla possibilità di successo, di racimolare un "gruzzolo", possibilità quest'ultima legata per alcuni direttamente al montepremi in palio per altri a possibilità di lavoro future nel mondo dello spettacolo. C'è chi invece vede nell'entrata nella casa la realizzazione di un'esperienza di vita connessa al "conoscere gente" e al "confrontarsi con loro".

È possibile discriminare a seconda del genere e della zona geografica in cui è stata realizzata l'intervista il prevalere dell'una o dell'altra motivazione.

8.3.1. "FORSE PER I SOLDI"

Sebbene accedere alla casa più famosa e spiata d'Italia appaia non semplice, c'è anche da segnalare che in molti vogliono comunque provarci; perché la "casa" rappresenta una *chances* di vita.

"Le *chances* di vita sono attributi dei singoli. Nella loro esistenza sociale gli individui hanno chances di vita: possono portarle a compimento o farle a pezzi, ma la loro vita è una risposta a queste ciance. Le *chances* di vita sono forme. Possono essere troppo grandi per l'uno e stimolarlo a espandersi e a crescere, possono essere anche troppo limitate e stimolarlo a opporre resistenza. Le *chances* di vita sono possibilità di crescita individuale, di realizzazione di capacità, di desideri, speranze e queste possibilità sono rese disponibili dalle condizioni sociali. Per

ogni determinato individuo e, in certo modo anche per gruppi, ceti, classi sociali, c'è un bilancio di *chances* di vita" (Darendhorf, 1981, pp. 40-41).

È facile a questo punto leggere la partecipazione al GF come una opportunità economica da cogliere. Lo spirito imprenditoriale si conferma presente soprattutto al Nord, dove le interviste che contengono segmenti riconducibili al codice che racchiude i connotati di questa partecipazione al reality sono 12, rispetto alle altre zone geografiche qui sono 9 gli uomini che sostengono tale idea; il triplo rispetto alle donne.

Tab. 8.1 - Il GF inteso come opportunità economica

	Uomo	Donna
Nord	9	3
Centro	5	6
Sud	5	6
Totale	19	15

Questo pubblico, raccolto nei piccoli centri, ha ben chiaro a cosa possa essere utile l'entrata nella casa. Nelle loro motivazioni si rintraccia una dimensione strumentale legata al contesto quotidiano in cui scorre la loro vita.

Emerge uno spirito imprenditoriale degli italiani mozzato dalle spigolosità delle banche e della nostra burocrazia. Già R. Darendhorf (1981) segnala come legate alle *chances* di vita ci siano "opzioni", quindi possibilità di scelta e "legature", intese come vincoli di varia natura cui va soggetto l'individuo.

Come pensa di ovviare il nostro pubblico a tali legature? Pensa bene che la soluzione sia da ricercarsi nella notorietà che si può raggiungere attraverso il mezzo televisivo, quasi senza nessun merito, se non quello di essere stati fortunati, così come nella scelta di un biglietto del *gratta e vinci* vincente.

A intenderlo come tale sono soprattutto gli uomini sentiti nel Nord dell'Italia; mentre del tutto assente è questo codice nelle interviste realizzate nel Centro della nostra penisola.

Tab. 8.2 - Il GF inteso come "svolta fortunata di vita"

	Uomo	Donna
Nord	9	2
Centro	-	-
Sud	1	2
Totale	10	4

C'è anche da segnalare che la partecipazione al GF si ammantava di una dimensione strumentale duplice, da un lato viene identificata come strumento attraverso il quale poter far soldi e dall'altro diventa mezzo per poter realizzare un sogno o sopperire a delle mancanze, certamente per regalarsi una qualità della vita migliore. Vissuta in questo modo, la partecipazione al GF e i soldi a essa legata permettono di risolvere un problema di tipo economico altrimenti non fronteggiabile; permettono infatti: di saldare un debito, di estinguere un mutuo, comprare casa, avviare un'attività imprenditoriale, garantire un futuro sereno ai propri figli. Il Grande Fratello è un modo per ovviare alle difficoltà economiche dovute anche alla congiuntura economica sfavorevole. E allora, bisogna industriarsi sul come farli questi soldi, perché non tutti, anche se venissero selezionati per fare il loro ingresso nella casa, potrebbero accedere al montepremi finale.

La modalità per far soldi è già chiara ai nostri provinciali, è necessario fare l'ospite nelle trasmissioni televisive o nelle serate organizzate in discoteca, perché anche per questo tipo di attività si può essere "uomo senza qualità".

8.3.2. "È UN'ESPERIENZA PER PROVARE QUALCOSA DI NUOVO"

Una prima risposta alla domanda sul perché questi "giovani" si siano messi in fila è ovvia, ovvero sono moderni pionieri del West: sono partiti alla ricerca dell'oro. Un oro rappresentato anche dalla possibilità di trovare espressione e realizzazione per il proprio sé, in termini cognitivi, prima ancora che strumentali. Se pensiamo ai cercatori d'oro così come ce li ha dipinti il cinema western come prima motivazione del loro viaggio c'è quella della ricerca di un guadagno, ma a ben pensare è altrettanto vero che costoro erano avventurieri. Nelle trame degli *Spaghetti western* quanti dei personaggi hanno fatto fortuna e quanti di questi, invece, hanno vissuto un'esperienza costitutiva per la propria vita durante questo viaggio?

Tab. 8.3 - Il GF inteso come "esperienza"

	Uomo	Donna
Nord	13	2
Centro	5	2
Sud	8	7
Totale	26	11

Sono 37 i frammenti di testo in cui si rintraccia l'idea che la partecipazione al GF possa essere una esperienza di vita.

Ci suggerisce A. Elliot (2007) che le persone si vengono a trovare sempre più nella condizione di definire se stessi in relazione a “mega-eventi”. “Come registi delle nostre esistenze, utilizziamo frames emozionali di memoria e desiderio, così come fonti culturali e sociali più generali per dare forma al sé (2007, trad. it. 2010, p. 8)”. Dice ancora Elliot – proponendoci una riflessione di C. Taylor (1990) – che noi siamo dei sé solo in quanto ci stanno a cuore delle questioni.

La partecipazione al GF sarebbe quindi un modo per costruire la propria identità come singolarità o individualità.

D'altro canto la partecipazione al GF è vista come un momento per poter parlare di sé, ovvero poter far conoscere al resto dell'Italia – all'*audience* che resta fuori dalla casa – le proprie esperienze di vita. La casa e la partecipazione al GF possono essere lette anche come necessità per la “narrazione di sé”. Ci dice un ragazzo di 35 anni che all'interno della casa parlerebbe del suo *“cambiamento fisico. Prima ero qualche chilo in più ora in meno. Parlerei di come ci sono riuscito, anche perché mi farebbero tante domande. Non smetti mai di confrontarti”*.

Una ragazza extracomunitaria vorrebbe raccontare di *“quando sono arrivata qui, io sono arrivata dopo quattro anni che i miei genitori stavano qua e in dieci anni a me non sono riusciti a farmi i documenti. Vorrei avere una risposta a questo. Perché a me non sono riusciti a fare i documenti?”*

Tab. 8.4 - Il GF inteso come momento per la “narrazione di sé”

	Uomo	Donna
Nord	6	3
Centro	2	5
Sud	2	4
Totale	10	12

Una sintesi per quest'erma bifronte viene dalle parole di un ragazzo diciannovenne neodiplomato. *“Sarà sia un'esperienza, ma riesci anche a raccontarti e raccontare con questa esperienza. Anche se dopo tutti quanti sanno i tuoi fatti. Non so, forse perché a me piace stare al centro dell'attenzione sempre. Questo peccato di megalomania. Mi piace farmi scoprire davanti alle telecamere da tutti.”*

8.4. DELINEARE I TRATTI DISTINTIVI DELLE DIMENSIONI

In questa prima analisi esplorativa dei contenuti dei codici per le interviste del GF è possibile delineare due gruppi o macro-categorie di significato inerenti il ruolo e la partecipazione a questo *reality show*.

Volendo lavorare sulla costruzione dell'esperienza e gli impatti sul sé sono state segnalate due dimensioni: una cognitiva e l'altra strumentale. I codici che rientrano nella dimensione cognitiva raccolgono i segmenti che si riferiscono ora all'esperienza della partecipazione ma anche alla motivazione. Tali codici possono essere raccolti in un set: "tra esperienza e motivazione".

Tab. 8.5 – Componenti del set "Tra esperienza e motivazione"

	N. Interviste	Segmenti
Esperienza	37	42
Descrizione dell'esperienza	2	2
Sono venuto per un'emozione in più	1	1
Conosci gente	5	5
Stare chiusi in una casa	24	29
Motivazione	9	9
Cosa vorrei trasmettere	10	13
Il GF per comunicare	1	1
Narrazione di sé	23	24
Riscatto sociale	6	8

All'interno di questo set si raccolgono quei codici che fanno capo all'area cognitiva, la tabella 8.5 ci indica il numero di interviste in cui tale codice è stato rinvenuto, mentre la seconda colonna riporta il totale dei segmenti etichettati.

"Sono venuto qui per un'emozione in più" e "Il GF per comunicare" sono degli *NVivo codes*, ovvero sono le parole stesse degli intervistati utilizzate per creare un'etichetta. Entrambi possono essere presi come emblema di questo set cognitivo-emozionale.

Di converso due soggetti distinti dichiarano di essere lì per i soldi ed elevato è il numero delle interviste – ben 43 – in cui si rintraccia l'elemento della partecipazione al GF come strumento per trovare un'occupazione futura nel mondo dello spettacolo.

Tab. 8.6 – Componenti del set “I soldi come farli e a cosa servono”

	N. Interviste	Segmenti
Sono qui per i soldi	2	2
Crisi economica	7	7
Il GF per risolvere i propri problemi economici	16	18
Il GF come colpo di fortuna	21	24
Opportunità economiche future	35	40
Occupazione futura nel mondo dello spettacolo	43	48

Non difficile è a questo punto cogliere il senso di quei sedici soggetti per i quali il GF è un modo per risolvere i propri problemi economici.

STRATEGIA DI ANALISI

STEP 1 – ORGANIZZARE I DATI

Dopo aver proceduto alla sbobinatura dei testi delle interviste è ora necessario:

- 1) organizzare il testo per casi/variabili,
- 2) procedere con la codifica.

Operazione 1

Per ogni intervista siamo in possesso di dati strutturali: sesso, età, zona geografica, occupazione dei soggetti sentiti. Attraverso le procedure presenti in NVivo è possibile collegare ogni intervista alle sue variabili strutturali, creando in questo modo un dataset in cui avremo 120 soggetti con le loro caratteristiche distintive.

Per creare i casi occorre selezionare **Nodes/Cases**, poi dal menu **New/Case in This Folder** si accede alla finestra di dialogo *New Case* e da qui si attribuisce un nome ai casi. Invece per gli attributi occorre selezionare **Classifications/Attribute** dal pannello di navigazione e poi **New/Attribute in This Folder**.

Operazione 2

Caricati i testi è ora possibile procedere all'operazione di codifica. In questo caso si è scelto di operare una codifica dal basso, ovvero i codici non sono già stati individuati a priori bensì emergeranno man mano che si procederà alla lettura del testo.

In questo caso abbiamo 120 documenti da aprire singolarmente, tante quante sono le interviste realizzate. Aperto il primo testo di intervista si procede leggendo riga per riga il documento e selezionando quei segmenti ritenuti rilevanti; a essi si applica una etichetta: il codice, che ha la funzione di chiudere, sintetizzare il contenuto del segmento stesso.

La creazione dei codici è un'operazione puramente concettuale, legata alla sensibilità dell'analista, poiché è lui stesso, in prima persona, che legge ed estrae dal testo quanto ritiene rilevante. In questa operazione il software non si può sostituire al ricercatore.

Dalla stessa porzione di testo possono emergere più codici.

Per esempio, un intervistato ci spiega così il perché della scelta della partecipazione al GF:

“Mi interessa per i soldi e le opportunità future. D'altra parte oggi sempre di più non è il potere che dà la fama ma la fama che dà il potere. Quindi sicuramente un trampolino di lancio senza aspirazioni particolari, però sarebbe una grande cosa per me entrare. Perché può servirmi in un momento poco felice per me familiare e lavorativo. Anche una sorta di stacco un attimo dalla mia vita quotidiana per ritrovarmi a vivere tutt'altra esperienza, quindi una sorta di ristoro e di recupero. Si farebbe il minimo indispensabile lì dentro”.

Questo segmento (*quotation*) ci suggerisce la creazione e applicazione di diversi codici:

- interesse per le future possibilità economiche,
- il GF come trampolino di lancio,
- narrazione e costruzione del sé.
- la casa come momento di incubazione.

Appare evidente come all'interno di questo segmento si affrontino argomenti diversi che possono essere chiusi nei codici testé menzionati.

STEP 2 – INTERROGARE I DATI

Mediante la creazione del casebook è stato possibile organizzare i dati attribuendo ai casi le informazioni strutturali.

A questo punto è possibile interrogare i dati chiedendo loro di estrarre

porzioni di testo collegate alle loro dimensioni.

Come la sociologia economica ci insegna, lo spirito imprenditoriale è maggiormente presente nel Nord dell'Italia piuttosto che a Sud, è una caratteristica degli uomini più che delle donne.

Mediante una *query* di codici si vuole sondare questa ipotesi.

Prima query

Durante la procedura di codifica, i segmenti di testo nei quali era ravvisabile l'idea del GF come opportunità economica sono stati etichettati all'interno dello stesso codice. Mediante la *query* di codici avanzata adesso si chiede di far emergere quei casi in cui è presente tale codice, specificando l'attributo "genere".

A questo punto si ha una selezione di frammenti pertinenti.

Il risultato è esportabile in un foglio formato excel dove è possibile procedere con un'operazione di conteggio per indicare a seconda del genere, in quale zona geografica è presente tale codice.

Seconda query

La tabella 8.3 è il risultato di una seconda *query* eseguita sui dati. Questa volta il codice da rintracciare all'interno delle interviste è "Il GF come svolta fortunata di vita".

STEP 3 – RICHIAMARE LE QUOTATIONS

A volte può impreziosire il testo l'aggiunta di alcune citazioni. A seguito della codifica del testo si avrà che le citazioni con significati simili saranno raccolte all'interno del medesimo codice. Per recuperare un segmento di testo è possibile selezionare da *Free o Tree Nodes* il codice in oggetto a questo punto si può visualizzarne il contenuto nel pannello di navigazione di NVivo oppure esportare in formato word l'intero contenuto del nodo.

Successivamente, un'attenta lettura permetterà di scegliere i segmenti maggiormente esplicativi.

STEP 4 – I SETS

I codici possono essere raggruppati in set. Un set altro non è che un insieme di codici pertinenti dal punto di vista logico. Questa procedura oltre che un lavoro di sintesi rappresenta uno step necessario per riflettere sulle etichette create. Infatti, può accadere che durante la codifica siano stati creati codici simili

che è poi possibile far confluire all'interno di uno solo di questi mediante la procedura *Merge code*.

Prima di creare un set è necessario quindi:

- analizzare i codici creati al fine di unire codici simili e di avere una *overview* del lavoro svolto.

A questo punto mediante l'applicazione **Sets** è possibile raggruppare insieme i codici. Per conoscere i dettagli in merito ai codici che compongono il set da **Tools/Reports** si chiede di estrarre le informazioni in merito alle fonti, ai segmenti, etc. in cui tali codici sono presenti.

RIFERIMENTI BIBLIOGRAFICI

- ANG I. (1991) *Desperately Seeking the Audience*, London, Routledge.
- APPADURAY A. (1996) *Modernity at Large: Cultural Dimension of Globalization*, Minneapolis: University of Minnesota Press.
- BESOZZI E. (2006) *Società, cultura, educazione*, Roma, Carocci.
- BOLTER J.D., GRUSIN R. (2002) *Remediation. Competizione e integrazione tra media vecchi e nuovi*, Milano, Guerini (ed. or. 1999).
- BUCKINGAM D. (2004) *Né con la tv, né senza tv*, Milano, FrancoAngeli (ed. or. 2000).
- DARENDHORF R. (1981) *La libertà che cambia*, Roma-Bari, Laterza.
- ELLIOT A. (2010) *I concetti del sé*, Torino, Einaudi (ed. or. 2007).
- LEMISH D. (2008) *I bambini e la tv*, Milano, Raffaello Cortina Editore (ed. or. 2007).
- MEYROWITZ J. (1995) *Oltre il senso del luogo. L'impatto dei media elettronici sul comportamento sociale*, Bologna, Baskerville (ed. or. 1993).
- POSTMAN N. (1994) *La scomparsa dell'infanzia*, Roma, Armando (ed. or. 1983).
- RIBOLIZI L. (1993) *Sociologia e processi formativi*, Brescia, Editrice La Scuola.
- RICOEUR P. (1990) *Soi-même comme un autre*, Paris, Édition de Seuil.
- SIMMEL G. (1998) *Sociologia*, Torino: Edizioni di Comunità (ed. or. 1983).
- TAYLOR C. (1993) *Radici dell'io. La costruzione dell'identità moderna*, Milano, Feltrinelli (ed. or. 1990).

