

Il lessico della guerra nei newsgroups della categoria it.politica durante la guerra in Iraq

Luca Giuliano

Dipartimento di Contabilità nazionale e analisi dei processi sociali –
Università degli studi di Roma “La Sapienza”
luca.giuliano@uniroma1.it

Abstract

How have reacted the Italian newsgroups to the war in Iraq? During the 28 days of “declared” war, what have been the main political and ideological guidelines?

The monitoring of the 8 newsgroups of the it.politica category has permitted to explore this topic. The more frequent non empty word in the corpus is <guerra> (15.605 occurrences on 5.220.932 occurrences in total). This paper intends to analyze the use of this form with a minimal loss of information and a meaningful gain of thematic deepening in a very “noisy” and “dirty” source. For this purpose, the analysis is performed on a sub-text extracted by concordances with keywords <guerra>.

In complex, there is a climate of general opposition to the war, although with various ways and reasonings. It is not always possible to give a description of the newsgroups based on the political choice of the users. They search the argument and the challenge with the political opponents.

Riassunto

Come hanno reagito i newsgroups italiani alla guerra in Iraq? Durante i 28 giorni di guerra “dichiarata” quali sono stati gli orientamenti politici ed ideologici prevalenti?

Il monitoraggio degli 8 newsgroups della categoria it.politica ha permesso di esplorare questo tema a partire dalla forma grafica <guerra>, la più frequente nel corpus (escluse le parole vuote) con 15.605 occorrenze su 5.220.932 occorrenze in totale. L’obiettivo è di minimizzare la perdita di informazione ma, nel contempo, ottenere un guadagno significativo di approfondimento tematico in una fonte di informazione nota per essere molto “rumorosa” e “sporca”. A questo scopo, l’analisi viene condotta su un sub-testo estratto dal corpus e composto dalle concordanze della forma <guerra>.

In complesso, ne emerge un clima di opposizione generalizzata alla guerra, sebbene con modalità e argomentazioni diverse. Non sempre è possibile dare una caratterizzazione precisa del newsgroup in base all’orientamento politico che lo identifica nominalmente. Gli utenti dei newsgroups cercano la discussione e il confronto polemico con gli avversari politici.

Keywords : newsgroups, textual analysis, correspondence analysis, politics, war.

1. Introduzione

La guerra in Iraq è iniziata ufficialmente alle 4:14 (ora italiana) del 20 marzo 2003, con l’annuncio del presidente degli Stati Uniti, George Bush: “E’ cominciata la guerra di liberazione dell’Iraq”. In effetti già poco meno di un’ora prima, alle 3:35 (5:35 ora locale), i primi missili americani cadevano su Bagdad, mentre i marines e i soldati britannici entravano in territorio iracheno. Il 2 aprile le truppe americane arrivavano alla periferia della capitale. Il 9 aprile, alle 16:49, le televisioni di tutto il mondo mostravano l’abbattimento della statua di Saddam Hussein in piazza al Ferdous. Da questo momento in poi l’Iraq passava progressivamente sotto il controllo anglo-americano. Il 16 aprile le operazioni di guerra potevano considerarsi finite con la richiesta, inoltrata dalla Casa Bianca all’ONU, di togliere l’embargo sull’Iraq.

I testi di cui trattiamo in questa comunicazione provengono da newsgroups di argomento politico. I newsgroups rappresentano una parte consistente della messaggeria elettronica asincrona. Sono più di 25.000 e generano circa 1.000.000 di messaggi al giorno. La loro rilevanza nei processi di comunicazione, non solo all'interno della comunicazione mediata dal computer, è stata ampiamente documentata (Smith, 1999; Choi e Danowski, 2003). Il sistema Usenet è facilmente accessibile ed è conosciuto in tutto il mondo dal 1979. In questi anni si è sviluppato fino a coprire un arco completo di argomenti e con diramazioni in moltissime aree linguistiche e culturali. Per la loro struttura sociale fondamentalmente "anarchica", i newsgroups sono particolarmente adatti per l'analisi delle correnti endogene nei conflitti sociali emergenti (Stubbs, 1998; Giuliano, 2002; Smith, 2002). L'archivio dei messaggi con interfaccia web, che è stato per diverso tempo gestito da Deja News, recentemente è stato acquistato da Google.

I newsgroups coinvolgono milioni di persone con una complessità di temi e di interazioni sociali che non si possono analizzare con gli strumenti attuali di navigazione e di "information retrieval" (Whittaker *et al.*, 1998).

L'analisi automatica dei dati testuali può offrire un contributo rilevante per l'individuazione dei nuclei di significato in una massa gigantesca di informazioni. In questa comunicazione si prendono in esame i newsgroups italiani durante la guerra in Iraq, la prima vera e propria guerra nell'era di Internet.

2. Corpus dei messaggi analizzati ed estrazione delle concordanze

La seconda guerra del Golfo ha avuto una durata di 28 giorni. Il periodo di riferimento per il prelevamento dei messaggi dai newsgroups della categoria it.politica va dal 18 marzo al 18 aprile.

Newsgroups	N. messaggi	Media msg al giorno	Identificato nel testo con:
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
<i>it.politica.ulivo</i>	<i>349</i>	<i>11,40</i>	<i>scartato</i>
<i>it.politica.libertaria</i>	<i>245</i>	<i>7,65</i>	<i>scartato</i>

*Tavola 1. Newsgroups della categoria it.politica secondo i messaggi nel periodo considerato:
18 marzo – 18 aprile 2003*

I newsgroups della categoria it.politica sono 10, ma it.politica.libertaria e it.politica.ulivo 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.

Il corpus è costituito così da un file di 33Mb denominato IRAQ28 sottoposto a normalizzazione completa in TALTAC 1.6 con le seguenti caratteristiche lessicometriche:

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

Tavola 2. Misure lessicometriche del corpus IRAQ28

Sebbene già “ripulito” dell’header dei messaggi (l’intestazione che contiene le indicazioni di mittente, destinatario, data di invio, soggetto, server di posta, ecc.), il corpus si presenta con un’alta componente di “rumore” dovuta a varie fonti: marcatori HTML, indirizzi internet e di email, caratteri non riconosciuti dal server, errori ortografici, “firme” dei mittenti che utilizzano disegni in caratteri ASCII, emoticon, parole e interiezioni gergali, quoting, copia-e-incolla da pagine web, sovrapposizione di idiomi diversi, ecc.).

L’alta quota di rumore contenuta nel corpus è misurabile approssimativamente attraverso le forme grafiche non riconosciute dal tagging grammaticale di TALTAC. Le forme grafiche non riconosciute sono 76.984 (pari al 42,98% delle forme grafiche distinte contro una media del 5% nei testi letterari e giornalistici). Nonostante questo, per l’analisi del lessico e per l’analisi multidimensionale del testo, il rumore può essere trascurato. Infatti la quota maggiore di forme non riconosciute (26,67% delle forme grafiche) si trova tra gli hapax (le forme grafiche di occorrenza 1). Assumendo come riferimento per l’analisi la soglia di frequenza 17 consigliata da TALTAC, la percentuale di rumore relativa al testo da analizzare scende drasticamente al 5,04% delle occorrenze, con una copertura del testo del 90,98%.

La fascia di alta frequenza contiene 133 forme grafiche, 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 newsgroups agli eventi in corso.

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

Tavola 3. Forme grafiche principali appartenenti alla fascia di alta frequenza nel corpus IRAQ28

Tra le forme più significative, “guerra” è la più frequente. Seguono i principali riferimenti al contesto (Iraq, Saddam, Bush, USA, americani, pace). Pertanto, avendo come obiettivo di analizzare il tema principale del corpus con il minimo di trattamento del testo e con procedure il più possibile standardizzate e automatiche si è proceduto nel seguente modo:

- Estrazione delle concordanze della forma grafica “guerra/guerre” (Lexico 3).
- Estrazione dei segmenti con soglia 7 (SPAD_5.0).

- Costruzione della tabella lessicale a soglia 20 ed eliminazione dei segmenti ripetuti vuoti, banali e/o ridondanti.
- Analisi delle corrispondenze sulla tabella lessicale dei segmenti ripetuti per newsgroups (SPAD 5.0).
- Analisi delle corrispondenze sulla tabella lessicale a soglia 50 delle forme per concordanze (*mots-réponses*) con SPAD-T 1.5.
- Classificazione delle concordanze sugli assi fattoriali individuati con SPAD-T 1.5.

L'estrazione delle concordanze (70 caratteri prima e 70 caratteri dopo la forma pivot <guerra>) ha portato alla individuazione di 18.698 stringhe che costituiscono il sub-corpus 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

Tavola 4. Misure lessicometriche del sub-corpus delle concordanze GUERRA-IRAQ

Il rumore contenuto nel sub-corpus è notevolmente ridotto. Le forme grafiche non riconosciute dal tagging grammaticale di TALTAC sono 8.659 (pari al 26,28% delle forme grafiche distinte, ivi incluso il taglio di parole dovuto a +/- 70 caratteri intorno al pivot); 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.

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

Nei newsgroups analizzati il clima di opposizione verso la guerra è prevalente. Il segmento ripetuto più frequente, a parte i segmenti di carattere grammaticale, è <contro la guerra> con 780 occorrenze.

I newsgroups si differenziano per il riferimento politico, salvo i due newsgroups generali ("politica" e "internazionale") che dovrebbero rispettivamente essere dedicati alla politica italiana e alla politica internazionale. Gli altri indicano chiaramente una scelta ideologica: it.politica.destra accoglie le discussioni intorno alla cultura di destra estrema; it.politica.leganord tratta argomenti relativi al federalismo; it.politica.polo è dedicato ai partiti di maggioranza della "Casa delle Libertà"; it.politica.pds è dedicato al partito dei Democratici di Sinistra; it.politica.rifondazione tratta argomenti che riguardano la sinistra di Rifondazione Comunista. Tuttavia gli utenti dei newsgroups non si differenziano nettamente secondo questi schieramenti. E' piuttosto diffusa la consuetudine di inviare una parte dei messaggi in tutti i newsgroups contemporaneamente (*crosspost*) sebbene sia vietato dalla *netiquette*. Gli utenti amano dibattere dei loro argomenti preferiti con gli "avversari". Accade piuttosto spesso che alcuni utenti si lascino andare a vere e proprie provocazioni, fingendosi appartenenti alla parte politica avversa o inondando il newsgroups di insulti.

3. Analisi dei segmenti ripetuti

L'analisi delle corrispondenze binarie segmenti/testi con soglia 20 effettuata con il programma SPAD 5.0 permette di individuare con qualche difficoltà i gruppi tematici caratteristici dei diversi newsgroups.

Numero	Autovalori	% di inerzia	% Cumulata di inerzia
1	0.0685	22.37	22.37
2	0.0587	19.17	41.54
3	0.0522	17.05	58.59
4	0.0415	13.57	72.16
5	0.0338	11.03	83.19
6	0.0287	9.36	92.55
7	0.0228	7.45	100.00

Tavola 5. Analisi delle corrispondenze dei segmenti ripetuti per testi: estrazione degli autovalori

Newsgroups	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

Tavola 6. Analisi delle corrispondenze dei segmenti ripetuti per newsgroups: coordinate e contributi assoluti delle frequenze attive

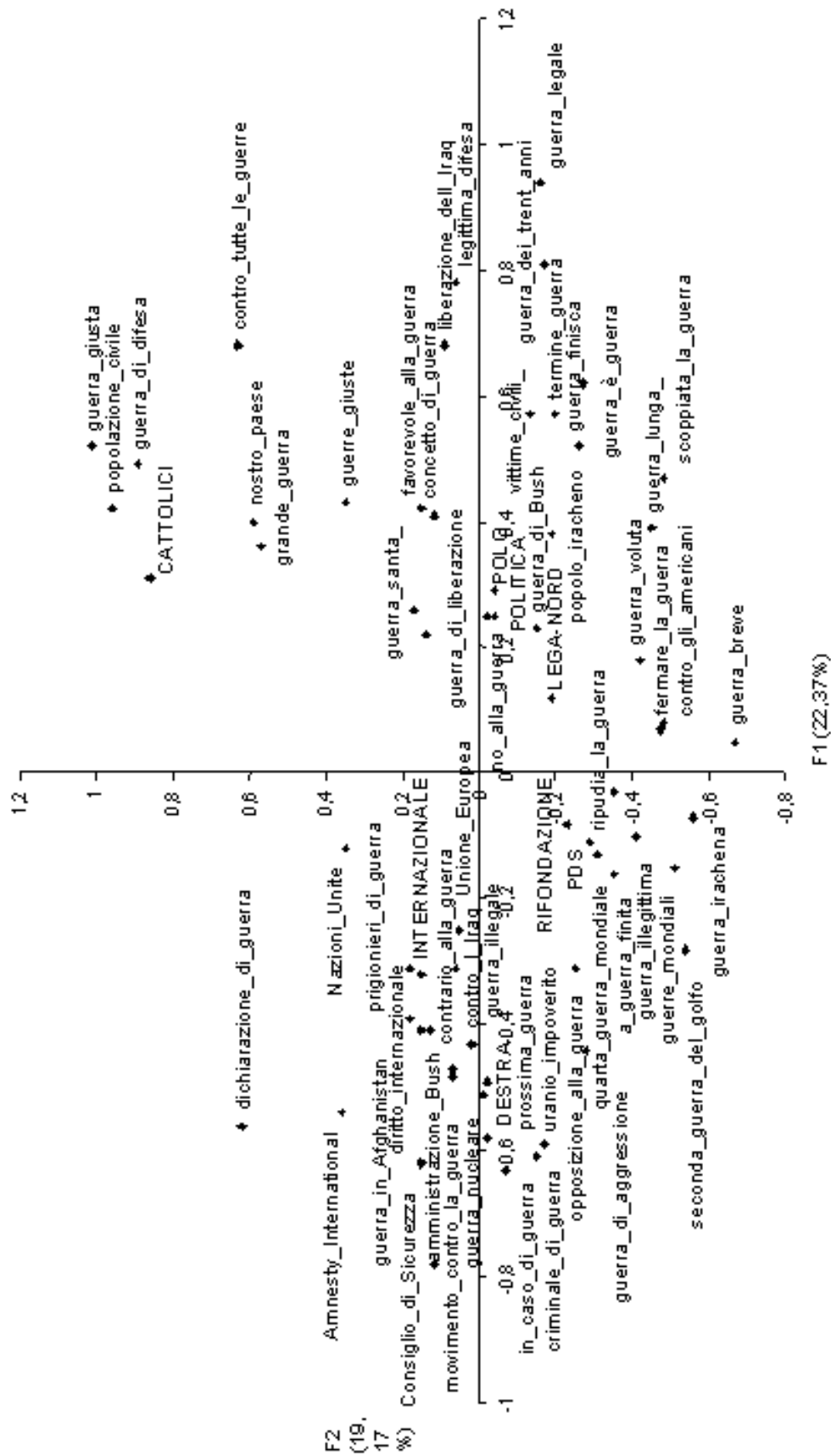
In questa comunicazione vengono presi in esame i primi due assi fattoriali che spiegano il 41,54% della variabilità complessiva. Le coordinate e i contributi assoluti delle frequenze attive permettono di individuare una diversificazione attesa sul primo fattore tra i due newsgroups principali, it.politica 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 tab. 7 sono riportati solo i segmenti ripetuti che presentano un contributo significativo per la formazione di almeno uno dei due fattori estratti.

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	0,9
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	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	0,3
guerra giusta	0,52	1,01	4,4	19,5
guerra illegale	-0,43	0,02	1	0
guerra illegittima	-0,1	-0,41	0,1	1,1
guerra in Afghanistan	-0,39	0,18	1	0,3
guerra irachena	-0,07	-0,56	0	3,2
guerra legale	0,94	-0,16	3	0,1
guerra lunga	0,39	-0,45	1,7	2,8
guerra nucleare	-0,51	-0,01	1	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
legittima difesa	0,78	0,06	2,2	0
liberazione dell Iraq	0,68	0,09	1,6	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
nostro paese	0,4	0,59	1,1	2,8
nuova guerra	-0,48	0,07	1,6	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
quarta guerra mondiale	-0,13	-0,31	0,1	0,8
ripudia la guerra	-0,03	-0,35	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	-0,2	1,1	0,2
Unione Europea	-0,25	0,05	0,7	0
uranio impoverito	-0,61	-0,15	1,7	0,1
vittime civili	0,57	-0,13	2,8	0,2

Tavola 7. Analisi delle corrispondenze dei segmenti ripetuti a soglia 20 per newsgroups: coordinate e contributi assoluti dei segmenti

Graf. 1 - Proiezione dei segmenti sul piano fattoriale degli assi 1 e



I contenuti emergono più chiaramente esaminando la proiezione dei segmenti ripetuti sul piano fattoriale formato dagli assi 1 e 2 (Graf. 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>.

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 newsgroups 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 alcune 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>.

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.

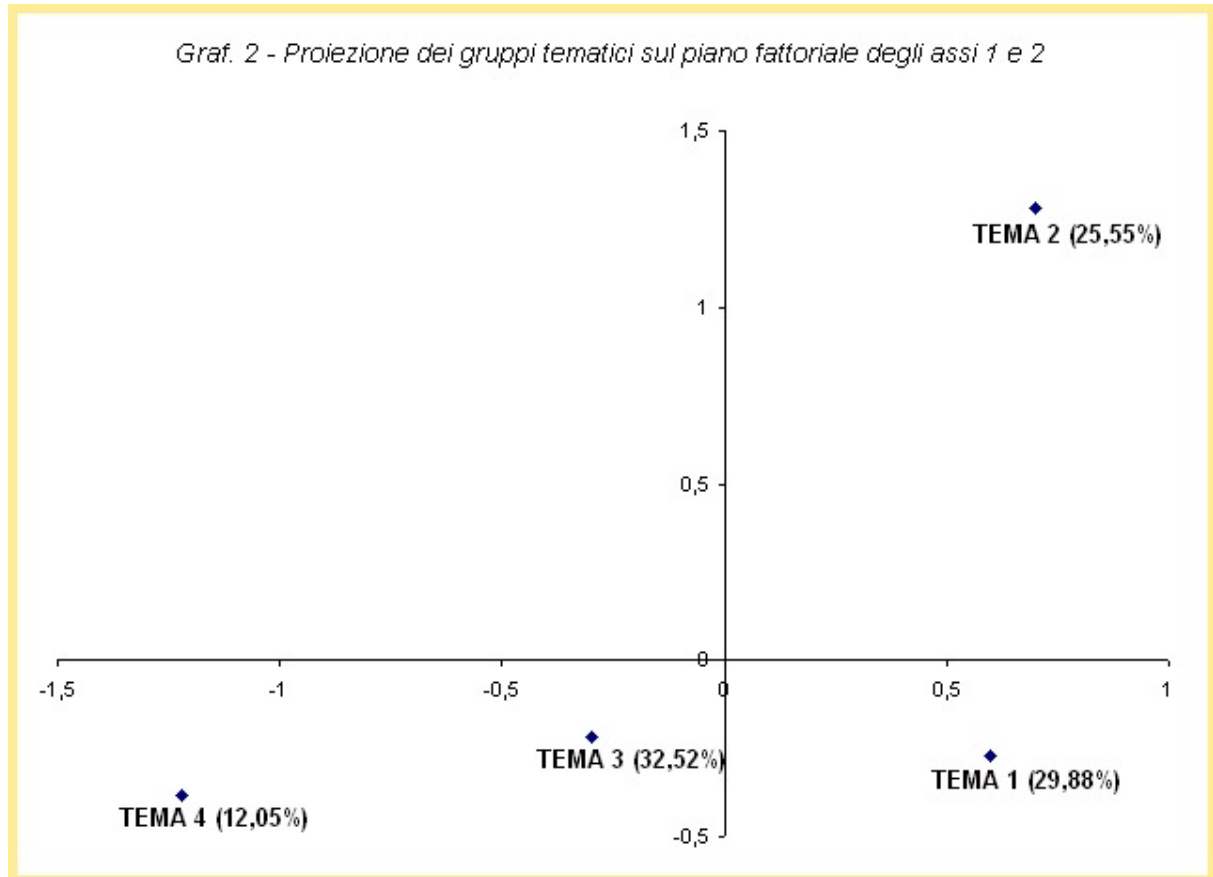
4. Analisi delle corrispondenze e classificazione delle concordanze.

L’individuazione dei temi dibattuti nei newsgroups può essere effettuata anche attraverso l’analisi delle co-occorrenze all’interno delle concordanze della forma <guerra>. L’analisi delle corrispondenze sulla tabella lessicale delle forme per concordanze (mots-réponses) effettuata alla soglia di frequenza 50 con la fase ASPAR del programma SPAD-T 1.5 per DOS ha permesso di individuare due fattori principali che, anche in questo caso, non permettono di evidenziare temi caratteristici di ciascun newsgroup ma permettono di individuare con maggiore dettaglio i nuclei di significato che sono stati dibattuti nel periodo preso in esame.

Per una sintesi più efficace di questo risultato si è ritenuto opportuno procedere ad una classificazione delle concordanze. L’ipotesi-guida di questa strategia di analisi è che dalla individuazione dei temi trattati sia possibile risalire ad una quantificazione, sebbene molto approssimata, del peso di ciascuno di essi nei messaggi del periodo preso in esame. Il peso dei gruppi tematici è rappresentato dalla frequenza relativa delle concordanze classificate all’interno dei rispettivi gruppi.

Nella *cluster analysis* effettuata sulle concordanze sono stati individuati 4 gruppi tematici.

Si riportano qui di seguito le descrizioni dei gruppi tematici sul primo piano fattoriale. Le forme caratteristiche che descrivono ciascun gruppo di concordanze sono elencate in ordine di specificità con valori del test sempre inferiori a 0,01. Le specificità sono state ottenute con la seguente catena di procedure: RECIP – PARTI – MOTEX – MOCAR.



Gruppo tematico 1. Riferimenti alla politica estera italiana e alla polemica tra maggioranza e opposizione (29,88% delle concordanze)

Onu, giusta, Papa, esistere, sinistra, Unione_Europea, posizione, Francia, risoluzione, dire, Italia, opposizione, Germania, preventiva, partecipare, Dio, legittima, dottrina, internazionale, morale, giuste, parole, favorevole, politico, dire_che, legale, politica, dovere, D'Alema, contrari, parola, nome, Fassino, illegale, chiesa, Berlusconi, impedire, porre, potere, costituzione, problema, unilaterale, Cofferati, concetto, umanità, umanitario, illegittimo, pace.

Gruppo tematico 2. Riferimenti alla polemica sul pacifismo e sulle sue motivazioni (25,55% delle concordanze)

in piazza, contro, scendere, manifestazione, popolo, pace, liberazione, manifestare, Saddam, Iraq, credere, andare, Paese, Roma, bandiera, pacifisti, protesta, persona, post, corteo, questa, mondo, motivazione, pacifista, obiettivo, aprile, usare, gente, unità, sicuro, iracheno, ragione, a_favore_di, giovani, governo, intendere, prova, opinione, scrivere, continuare, liberare, campagna, dittatore, volere, santa, nazionale, esempio, portare, appoggiare, discutere, lotta.

Gruppo tematico 3. Riferimenti alla politica internazionale americana e alle motivazioni dell'intervento militare (32,52% delle concordanze)

mondiale, prima_di, seconda_guerra_mondiale, perdere, guerra_fredda, crimine, militare, americano, alleati, vincere, quarta, Stati_Uniti, durare, prima_guerra_mondiale, uomo, nemici, venire, sangue, anno, potenza, alla_fine_di, effetto, ogni, sperare, terrore, terza, amministrazione, scoppiare, storia, mese, secolo, Bush, terribile, stare, passare, minaccia, tre, statunitense, petrolio, politici, due, nemico, paura, criminale, rispondere, sapere, fronte, Siria.

Gruppo tematico 4. Riferimenti alle cronache della guerra e al regime di Saddam (12,05% delle concordanze)

civili, morti, di massa, guerra del golfo, vittima, distruzione, Afghanistan, armi, durante la guerra, bambino, città, pentagono, soldato, bomba, bombardamento, centinaio, tornare, iracheni, missili, militari, Iran, notizia, golfo, milione, migliaio, decina, occidentali, danni, colpire, Bagdad, arrivare, propaganda, popolazione, giorno, donna, giornalista, prigioniero, morire, morte, truppa, Vietnam, vita, capo, tv, affare, embargo, raccontare, sporca, totale.

Il tema prevalente è il n. 3 (32,52% delle concordanze). Le forme grafiche classificate in questo gruppo individuano l'argomento che ha interessato la maggior parte degli utenti dei newsgroups: il dibattito intorno alle motivazioni dell'intervento militare e alle sue giustificazioni sul piano del diritto internazionale. Emergono diversi riferimenti ai crimini compiuti da Saddam Hussein, alle scelte dell'amministrazione Bush, alla storia delle guerre mondiali che hanno caratterizzato il XX secolo, allo "scontro tra civiltà". Le motivazioni vengono ricondotte alla paura del terrorismo, alla liberazione dell'Iraq dalla dittatura, quanto alla presenza dei pozzi petroliferi.

5. Conclusioni

La strategia di analisi adottata si è rivelata adeguata alla individuazione dei temi principali del dibattito nei newsgroups presi in esame. La presenza di alte componenti di "rumore" nei messaggi è stata neutralizzata con l'analisi delle concordanze della forma <guerra> senza comportare perdite significative per l'analisi del contenuto. Le scelte compiute con l'utilizzazione di TALTAC come strumento di riconoscimento delle forme grafiche, normalizzazione, disambiguazione e tagging ha dato dei buoni risultati anche sul piano della "pulitura" automatica del testo.

Nei messaggi non è stato possibile individuare tendenze ideologico-politiche distinte in sintonia con le scelte tematiche dei newsgroups. 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 newsgroups 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. Saranno necessarie ulteriori analisi sul corpus complessivo dei newsgroups, e non soltanto sul sub-corpus delle concordanze, per poter confermare o meno l'esistenza di un lessico della guerra caratteristico delle diverse posizioni ideologiche.

Bibliografia

- Beaudouin V., Fleury S. e Velkowska J. (2000). Études des échanges électroniques sur internet et intranet : forums et couriers électroniques. In *Actes des JADT 2000*, vol. (1) : 17-24.
- Bolasco S. (1999). *Analisi multidimensionale dei dati*. Carocci.
- Bolasco S., Baiocchi F. e Morrone A. (2000-2003). *TALTAC 1.6. Trattamento Automatico Lessico Testuale del Contenuto*. CISU.
- Choi J.H. e Danowski J. (2003). Making a Global Community on the Net – Global Village or Global Metropolis ? : A Network Analysis of Usenet Newsgroups. *Journal of Computer Mediated Communication*, vol. (7/3). <http://www.ascusc.org/jcmc/vol7/issue3/choi.html>.

- Giuliano L. (2002). G8-2001 : la rivolta nel monitor. Analisi testuale dei messaggi nel newsgroup <it.eventi.g8.genova> durante gli scontri di piazza. In *Actes des JADT 2002* : 301-311.
- Lebart L., Morineau A., Bécue M. e Hausler L. (1993). *SPAD-T 1.5 (DOS)*. CISIA.
- Lamalle C., Martinez W., Fleury S. e Salem A. (2002). *Lexico3*. Université de la Sorbonne Nouvelle – Paris 3.
- Rosen D., Woelfel J., Krikorian D. e Barnett G.A. (2003). Procedures for Analyses of Online Communities. *Journal of Computer Mediated Communication*, vol. (8/4). <http://www.ascusc.org/jcmc/vol8/issue4/rosen.html>.
- Smith M.A. (1999). Invisible crowds in cyberspace : Mapping the social structure of the Usenet. In Smith M.A. e Kollock P. (Eds), *Communities in cyberspace*. Routledge : 195-219.
- Smith M.A. (2002). Mapping Social Cyberpaces : Measures and Maps of Usenet, a Computer Mediated Space. *Dissertation Abstracts International, A : The Humanities and Social Sciences*, vol. (63/1) : 382-A.
- Stubbs P. (1998). Conflict and Co-Operation in the Virtual Community : eMail ed the Wars of the Yugoslav Succession. *Sociological Research Online*, vol. (3). <http://www.soc.surrey.ac.uk/socresoline/>.
- Tuzi A. (2003). *L'analisi del contenuto*. Carocci.
- Whittaker S., Terveen L., Hill W. e Chemy L. (1998). The dynamics of mass interaction. In *Proceedings of Conference on Computer Supported Cooperative Work* : 257-264. <http://www.acm.org/pubs/citations/proceedings/cscw/289444/p257-whittaker/>.