

# **Gli sbarchi dei clandestini nei quotidiani: un'analisi testuale esplorativa**

Silvia Bartoletti, Alessandra Garbero, Silvia Montecolle,  
Ferdinando Nisco, Emanuela Recchini, Irene Salerno

Università di Roma "La Sapienza", Facoltà di Economia e Commercio  
silbartoletti@hotmail.com, Alessandra.Garbero@fao.org, fmontecolle@tiscali.it,  
nisco@freemail.it, emanuelarecchini@libero.it, irenesalerno@yahoo.it

## **Abstract**

In this paper we aim to present a method of textual analysis providing a brief description of immigration through the articles published in different newspapers in spring 2003 when Sicily – Italy – was invaded by a great number of clandestines. We have considered five newspapers representing five different Italian geographic areas (north-east, north-west, centre, south and islands). First of all, we aim to analyse forms and verbs used, highlighting specific features and analogies; then, using automatic techniques we aim to have a synthetic knowledge of different points of view picking up the themes of the texts, and the positions that newspapers hold by means of "forms-texts" tables.

## **Riassunto**

Il presente lavoro ha al centro del suo interesse l'analisi del fenomeno migratorio per come esso è stato rappresentato e descritto in alcuni quotidiani italiani nel periodo della primavera dell'anno 2003. La scelta del "momento storico" preso in esame non è stata casuale: l'Italia, infatti, è stata interessata, proprio in tale periodo, da un'ondata migratoria che si è imposta sulle cronache dei principali quotidiani italiani per via della particolare drammaticità con cui si sono svolti gli sbarchi di stranieri clandestini. Analizzando con gli strumenti propri dell'analisi testuale gli articoli estratti da cinque diverse testate giornalistiche italiane scelte secondo un criterio "geografico" ci si è proposti di valutare in che modo i quotidiani scelti hanno affrontato il problema migratorio, cogliendo analogie e specificità attraverso un'analisi multidimensionale dei dati testuali.

**Parole chiave:** analisi delle corrispondenze semplici, analisi testuale, flussi migratori.

## **1. Introduzione**

Nel periodo tra maggio e giugno 2003, la buona stagione e le favorevoli condizioni del mare hanno contribuito all'improvvisa ripresa degli sbarchi sulle coste meridionali dell'Italia di imbarcazioni cariche di clandestini. In particolare, è stata l'isola di Lampedusa, in Sicilia, la principale meta degli sbarchi. I quotidiani italiani hanno discusso per più di un mese dell'argomento nelle pagine dedicate alla cronaca.

Sono stati raccolti articoli omogenei per contenuto – cronaca – e dimensioni, apparsi a cavallo tra i mesi di maggio e di giugno sui seguenti quotidiani: Il Gazzettino Veneto, Brescia Oggi, Il Messaggero, Il Mattino, La Sicilia.

Il criterio con il quale è stata operata la scelta delle testate ha seguito un disegno preciso: si è scelto di adottare un parametro "geografico", per aver così rappresentata, attraverso quotidiani locali, l'Italia tutta intera: il Nord Est (Il Gazzettino Veneto), il Nord Ovest (Brescia Oggi), il

Centro (Il Messaggero), il Sud (Il Mattino); per quanto riguarda, poi, il Meridione, si è ritenuto interessante prendere in considerazione anche il quotidiano “La Sicilia”, per via del peculiare rapporto che questa regione ha intrattenuto con gli sbarchi di stranieri, nella scorsa primavera.

Obiettivo di questa ricerca è stato quello di esplorare, attraverso un’analisi testuale, le parole-chiave caratterizzanti il corpus *in toto*, e le specificità tra i quotidiani oggetto d’analisi, per cogliere similitudini o disuguaglianze dettate da ragioni geografiche – e forse politiche – nel pensiero e nella percezione dei giornali esaminati.

## 2. Il corpus e l’analisi delle forme grafiche

Il corpus risulta composto da circa 37.000 occorrenze totali (N), e da circa 6.380 forme grafiche diverse (V), equidistribuite tra i quotidiani oggetto di studio.

Forme grafiche	Occorrenze	Forme grafiche	Occorrenze	Forme grafiche	Occorrenze
Clandestini	213	isola	85	donne	59
Lampedusa	175	Isola	85	ministro	56
Immigrati	142	bordo	84	costiera	54
Mare	107	persone	84	barcone	52
Italia	100	Libia	83	notte	50
immigrazione	97	Bossi	60	legge	47
guardia	92	porto	60	Lega	38

Tabella 1. Graduatorie delle forme grafiche maggiormente impiegate nel corpus – valori assoluti

In primo luogo si è proceduto con lo studio delle forme grafiche in termini di frequenza d’uso, eliminando gli *hapax* e le parole strumentali quali preposizioni, articoli, ecc. Nella graduatoria delle forme grafiche, costruita rispetto al numero di occorrenze, (tabella 1) la prima parola piena che si incontra è “clandestini”, (che nel *corpus* presenta un numero di occorrenze pari a 213), forma grafica a connotazione prevalentemente negativa, che presenta un numero di occorrenze superiore rispetto ad “immigrati” ed “immigrazione”, a pregnanza semantica prevalentemente neutra. Seguono forme grafiche che individuano geograficamente i luoghi dove sono avvenuti gli sbarchi: “Lampedusa”, “mare”, “Italia” ed “isola”. Anche “persone” e “bordo” sono parole chiave per il *corpus*: la prima indica in maniera ancora più generale i soggetti coinvolti nel fenomeno; la seconda rappresenta il modo in cui queste persone arrivano in Italia: a “bordo” di qualche mezzo, che all’interno del *corpus* è legato a diversi termini (barcone, barca, peschereccio, imbarcazione, ecc.).

### Analisi delle forme verbali

Di particolare interesse si è rilevato lo studio delle forme verbali.

Si è proceduto con la lemmatizzazione (riconducendo tutte le voci verbali al corrispondente tempo all’infinito) riducendo il numero di forme da esaminare di circa il 50%, passando dai precedenti 1.338 differenti tempi verbali a 692 voci all’infinito.

Escludendo i verbi ausiliari e tutti i servili<sup>1</sup>, la graduatoria dei verbi all’infinito col maggior

<sup>1</sup> Tali verbi hanno fatto riscontrare una prevalenza nettamente superiore a tutti i rimanenti. Tale situazione non permetteva di far emergere i reali aspetti caratteristici del testo. Si è pensato, quindi, di escluderli da alcune analisi.

numero di voci coniugate nel *corpus* vede al primo posto “venire” con 60 occorrenze, seguito da “trovare” con 45, da “dire” con 43, da “arrivare” con 40 e da “chiedere” con 37 (tabella 2). Ad un elevato numero di occorrenze non è corrisposto un elevato utilizzo di differenti unità lessicali, ossia di diverse forme verbali coniugate. Nel caso specifico, è risultato che “chiedere”, benché all’ultimo posto della graduatoria delle occorrenze, è al secondo posto in quella delle unità lessicali.

Verbi	Occorrenze	Unità lessicali
Venire	60	13
Trovare	45	7
Dire	43	9
Arrivare	40	10
Chiedere	37	10

Tabella 2. *Graduatoria dei verbi maggiormente impiegati nel corpus – valori assoluti*

Per individuare i verbi maggiormente caratterizzanti il corpus si è effettuato un confronto fra questo ed un modello di riferimento<sup>2</sup>. In particolare, calcolando uno scarto sulle occorrenze si sono individuati i verbi sovra-utilizzati rispetto allo standard scelto: le cosiddette “parole chiave” del *corpus*. Come era logico supporre, i valori più elevati sono stati registrati per verbi strettamente correlati con la descrizione della tematica in esame (tabella 3).

Verbi	Scarto	Verbi	Scarto
Avvistare	98,15	Scortare	27,61
Spiaggiare	58,35	Salpare	23,56
Trasbordare	31,95	Approdare	17,19
Naufragare	29,40	Sbarcare	16,66
Intercettare	27,84	Speronare	16,60

Tabella 3. *Graduatoria dei primi 10 verbi con scarto positivo più elevato*

Considerando la frequenza cumulata associata ai primi 10 verbi maggiormente caratterizzanti il corpus<sup>3</sup> si è osservato che ne La Sicilia e ne Il Gazzettino Veneto è stato registrato il numero di occorrenze più elevate, rispettivamente pari a 44 e 35. Tale situazione sancisce l’impostazione maggiormente indirizzata al mero racconto dei fatti di cronaca da parte di queste due testate rispetto alle rimanenti.

### 3. L’analisi delle specificità nei sub-testi

La tecnica delle specificità permette di comparare gli articoli provenienti dalle differenti testate e di osservare le modificazioni nel profilo lessico-metrico. In sintesi le specificità rappresentano il sovra-impiego o viceversa il sotto-impiego di una forma rispetto ad una soglia di probabilità che nel nostro caso è posta pari al 5%.

Il quotidiano Brescia Oggi rispetto alle altre testate in esame ha presentato un sovra-utilizzo di termini più legati alla una discussione politica del problema. La maggior parte delle forme

<sup>2</sup> Il modello di riferimento scelto è rappresentato dagli articoli apparsi nel corso del 1990 sul giornale La Repubblica, a tiratura nazionale.

<sup>3</sup> Sono tali i verbi con scarto positivo più elevato.



Forme grafiche	Brescia Oggi	Gazzettino Veneto	Il Messaggero	Il Mattino	La Sicilia
Barca			13 +E02	1 -E03	
Barcone				4 -E02	21 +E03
Berlusconi	7 +E03				
bordo				30 +E02	
Bossi	28 +E12		0 -E07		
cadaveri		12 +E03			
Capitaneria					10 +E03
chiede					0 -E02
clandestini				19 -E06	77 +E06
disperati		19 +E03			
donne	22 +E03				
extracomunitari					24 +E07
figli			8 +E04		
guardia	1 -E04				
immigrati	7 -E03				
immigrazione	32 +E09				
intervento				10 +E04	
isola				8 -E03	41 +E07
Italia					14 -E02
Lega	18 +E08		0 -E04		
legge			0 -E05	19 +E03	
maggioranza	6 +E04				0 -E02
ministro	23 +E09				
notte			19 +E03	3 -E03	
organizzazione					12 +E07
Paese				8 +E03	
paesi			1 -E02		
peschereccio			11 +E03		
Pisanu	17 +E07		0 -E05-		
polizia					12 +E03
presidente	10 +E04				
problema		13 +E02			
terra			11 +E03		
vittime			6 +E03		

Tabella 4. Specificità positive e negative nei quotidiani

Come per le forme grafiche, anche per i segmenti è possibile fissare delle soglie di frequenza che ne condizionano la selezione nel testo (frequenza maggiore o uguale a 4); la soglia scelta è abbastanza bassa per ottenere una lista esaustiva. Ma l'elenco ottenuto, così com'è, non è efficiente: risultano selezionate molte sequenze non significative perché costituite soltanto da forme grammaticali, incomplete o semplicemente banali ai fini dell'analisi. Sono stati eliminati questi segmenti e si è proseguita la selezione ricorrendo al criterio di "completezza

Segmento	Occorrenze totali	Lunghezza	Indice IS	Indice IS relativo
centro di accoglienza	54	3	2,85	0,31
navi della marina militare	10	4	2,84	0,17
presunti scafisti	11	2	2,73	0,68
carretta del mare	29	3	2,59	0,28
peschereccio tunisino	8	2	2,25	0,56
acque internazionali	16	2	2,14	0,53
tratto in salvo	7	3	2,02	0,22
motovedetta della guardia	18	3	1,69	0,18
vita migliore	7	2	1,52	0,38
nostro Paese	6	2	1,50	0,37
barca da pesca	7	3	1,47	0,16
bilancio ufficiale	4	2	1,41	0,35
coste siciliane	6	2	1,36	0,34
nostre coste	7	2	1,26	0,31
immigrazione clandestina	8	2	1,09	0,27
cadaveri recuperati	5	2	1,05	0,26
nord africani	6	2	1,02	0,25
verso l'Italia	9	2	0,94	0,23

Tabella 5. Segmenti ripetuti con  $IS > 0,90$

Gli articoli presi in esame, nel sottolineare come il <nostro Paese> rappresenti per alcuni popoli un'attrazione fortissima, non dimenticano di rimarcare la precarietà delle condizioni di viaggio degli extracomunitari (viaggiano trasportati da <presunti scafisti> a bordo di quelle che vengono definite le <carrette del mare>): molto spesso, difatti, la "rotta della speranza" Nordafrica-Sicilia si trasforma in tragedia. Naufragano <barche da pesca> in pessime condizioni che inseguono il sogno di una <vita migliore>, trasformato in incubo di morte (numerosi i <cadaveri recuperati>). Solo l'intervento delle <navi della Marina Militare>, affiancate dalle <motovedette della Guardia Costiera>, evita l'irreparabile. Non sempre la politica riesce a gestire la situazione; accanto ad organismi dello stato, (<Marina Militare>, <Guardia Costiera>, <Guardia di Finanza>), ad affrontare l'emergenza vi sono i <centri di accoglienza> ed organizzazioni di volontariato laiche e confessionali.

## 5. Analisi delle corrispondenze semplici

Per analizzare al meglio le relazioni multidimensionali tra i quotidiani oggetto dello studio e le forme grafiche utilizzate da ciascun giornale, è stata realizzata un'analisi delle corrispondenze semplici con il software statistico SPAD che, sintetizzando le informazioni contenute in ciascun subtesto, evidenzia associazioni e contrasti, al fine di cogliere al meglio analogie e specificità di ciascun quotidiano.

Il *corpus* è stato preventivamente trattato, eliminando le congiunzioni, i pronomi, le preposizioni e le esclamazioni. Si è ritenuto proficuo concentrare l'analisi sui primi due fattori che insieme sintetizzano una discreta quota di inerzia complessiva (tabella 6).

Fattore	Autovalore	Percentuale	Percentuale cumulata
Primo	0,1839	38,18	38,18
Secondo	0,1286	26,71	64,89
Terzo	0,0999	20,74	85,63
Quarto	0,0692	14,37	100,00

Tabella 6. Autovalori ed inerzia spiegata dai primi quattro fattori

L'analisi ha fornito anche le coordinate e i contributi sugli assi fattoriali delle cosiddette frequenze attive che, nel caso specifico, corrispondono ai cinque sub-testi considerati.

Le coordinate permettono di individuare la posizione dei cinque quotidiani sul piano fattoriale (Il Gazzettino Veneto è quello più vicino al baricentro). Proseguendo con l'analisi dei contributi delle frequenze attive sui primi quattro fattori, quello che si evince è che sul primo asse Brescia Oggi è il quotidiano che più apporta un contributo significativo (48%), seguito da Il Mattino (23%). Sul secondo asse, invece, i quotidiani che apportano il contributo maggiore sono Il Messaggero (61.4%) e La Sicilia (35%); quasi nullo il contributo de Il Gazzettino Veneto ai primi due assi (rispettivamente 1,2% e 2,2%).

Analizzando i contributi degli individui attivi sui primi quattro fattori, ad influire di più sul primo asse sono forme grafiche di tipo politico-istituzionale ("Pisanu", "Bossi", "presidente", "Paesi", "uomo", "legge Bossi-Fini", "Berlusconi", "UE", "cooperazione", "maggioranza", "chiede", "Senato", "Camera", "vertice", "voce", "Calderoni", "Frattini", "Carroccio").

Sul secondo asse, invece, hanno peso maggiore le forme grafiche di tipo descrittivo (Lampedusa, mare, Libia, guardia di finanza, motovedette, problema, paesi, qui, peschereccio, terra, ospedale, organizzazione, casa, Somalia, racket) e discorsivo ("racconta", e tutte le forme verbali alla terza persona singolare e plurale).

Il primo asse rappresenta la tipologia di informazione che viene fornita dal quotidiano: in particolare "politica" (semiasse negativo) e "non politica" (di cronaca, pertanto) (semiasse positivo). Il secondo asse, invece, è indicatore della dimensione territoriale dell'informazione: nazionale (semiasse positivo), locale (semiasse negativo).

In figura 1 viene proposto il grafico con la rappresentazione simultanea delle frequenze e degli individui attivi sui primi due assi fattoriali: ciascun giornale ha una caratterizzazione ed una collocazione a sé stante.

Nel primo quadrante (++) è posizionato Il Messaggero. Quella de Il Messaggero è cronaca nazionale "tout court", lontana dalla politica e dalle problematiche locali e regionali, ma vicina all'Italia tutta intera; è cronaca nel senso più vero della parola: descrittiva, innanzitutto (pescherecci, acque, barche, uomini, frontiera, terra, notte, assistenza), anche se talvolta assume tinte patetiche (cadaveri, casa, tragedia, vittime) o polemiche (racket).

Nel secondo quadrante (-+) è posizionato Brescia Oggi; nonostante non sia un giornale di partito, ciò che emerge è la sua natura politica (specie se paragonato agli altri quotidiani oggetto di studio), probabilmente per la realtà geografica di cui è portavoce (Brescia e zone limitrofe) che lo rende inequivocabilmente vicino a Bossi, affrontando il problema in chiave nazional-politico-legislativo (Bossi, senato, maggioranza, presidente, Pisanu, legge Bossi-Fini, immigrazione, Berlusconi, pattugliamento, problema, emergenza).

Nel terzo quadrante (-- ) è posizionato Il Mattino. Analogamente a Brescia Oggi si sofferma sulla dimensione legislativa del problema, anche se in chiave meno politica (legge Bossi-Fini, fenomeno, Italia, acque internazionali); sul secondo asse non è distante da La Sicilia, probabilmente per la sua natura di "quotidiano del Sud" (asilo, accoglienza, forze, ordine, intervento).

Nel quarto quadrante (+-) sono posizionati La Sicilia ed Il Gazzettino. La Sicilia è completamente assorbita dalla dimensione locale del problema. In primo piano tutta una serie di specificità geografiche (Lampedusa, Agrigento, Palermo, Porto Empedocle) ed un resoconto puntuale di tutti gli aspetti organizzativi ed assistenziali (organizzazione, allarme, ospedale, guardia costiera, capitaneria, motovedette, guardia di finanza).

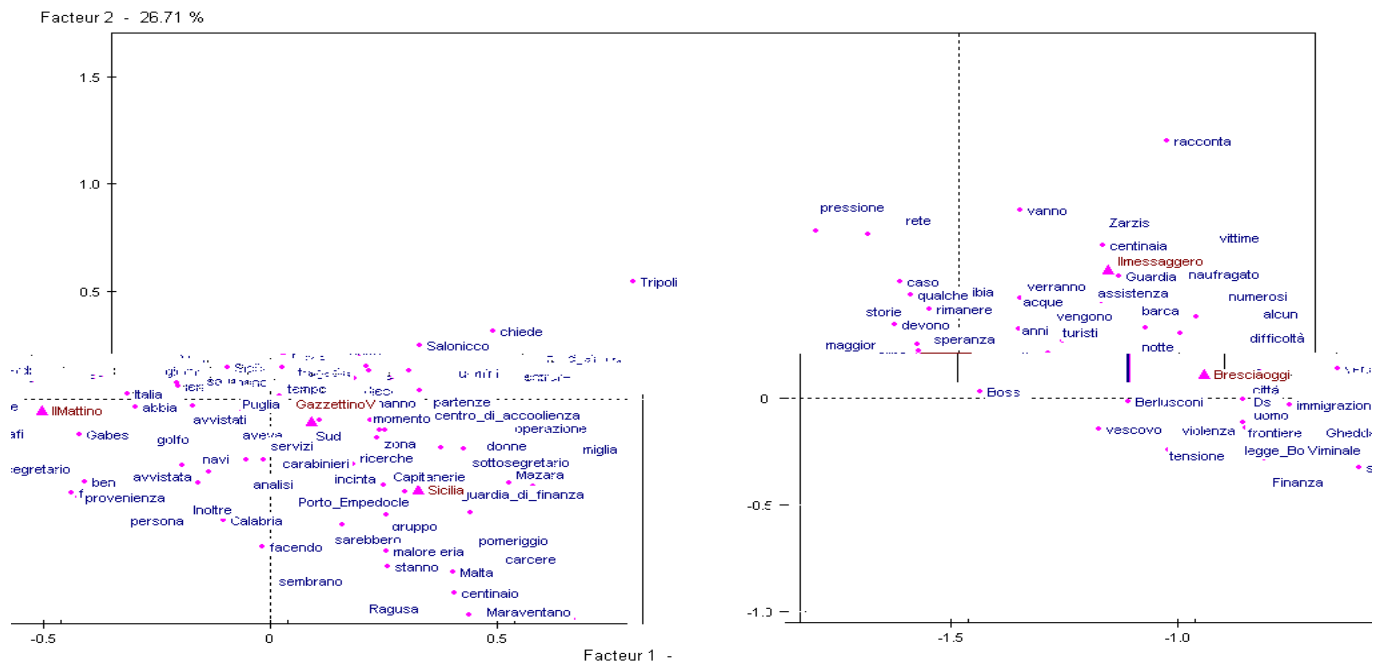


Figura 1. Diagramma quotidiani-forme grafiche

## 6. Conclusioni

L'analisi è stata condotta su un argomento che durante l'ultima primavera ha suscitato particolare interesse, sia per le problematiche organizzative, che un tale flusso di immigrati ha comportato, sia per l'approccio legislativo, che in qualche modo doveva far fronte alla sopraggiunta emergenza, sia per l'aspetto tragico dei numerosi naufragi avvenuti.

L'analisi testuale degli articoli dei cinque differenti giornali raccolti ha portato ad interessanti conclusioni, facendo emergere le differenti impostazioni delle diverse testate.

In particolare, si osserva la netta contrapposizione tra *Brescia Oggi*, che ha affrontato la questione facendo maggiormente riferimento a spunti legislativi e *La Sicilia*, che in quanto giornale dell'isola direttamente coinvolta dagli sbarchi, è stata più attenta a descrivere come si è fatto fronte all'emergenza. Su posizioni intermedie si sono attestate le rimanenti tre testate, pur denotando un'impostazione più politica per il *Mattino* e un'impostazione più di cronaca per *Il Messaggero* e *Il Gazzettino Veneto*.

## Bibliografia

- Bolasco S. (1999). *Analisi Multidimensionale dei dati*. Carocci.
- Bolasco S. e Cipriani R. (1995). *Ricerca qualitativa e computer*. Franco Angeli.
- Lebart L., Salem A. e Berry L. (1998). *Exploring textual data*. Kluwer Academic Publishers.
- Lebart L., Morineau A. e Piron M. (1995). *Statistique Exploratoire Multidimensionnelle*. Dunod.
- Lebart L. e Salem A. (1994). *Statistique textuelle*. Dunod.
- Morrone A. (1993). Alcuni criteri di valutazione della significatività dei segmenti ripetuti. In *JADT* 1993.