

### *La storia dell'analisi del contenuto*

La ricerca empirica sul contenuto della comunicazione risale agli studi teologici del tardo settecento, quando la Chiesa era preoccupata del diffondersi di temi non religiosi nella stampa.

Il primo caso ben documentato di analisi di materiale a stampa si ha in Svezia nel diciottesimo secolo. Si trattava di una raccolta di novanta inni intitolata **Canti di Sion**, di autore sconosciuto<sup>1</sup>.

Fu con la diffusione capillare della stampa che crebbe anche l'interesse nei confronti dei contenuti latenti da essa veicolati. Una delle prime ricerche in questo campo fu compiuta da **Speed**, il quale nel 1893 si chiedeva, mediante un'indagine classificata come analisi quantitativa dei giornali, se questi ultimi dessero effettivamente le notizie oppure no.

La prima applicazione pratica su larga scala dell'analisi del contenuto venne durante la seconda guerra mondiale.

Nel 1927 **Harold Lasswell** analizzò le tecniche di persuasione impiegate su vasta scala durante la prima guerra mondiale (propaganda americana, inglese, francese e tedesca) fra il 1914 ed 1917. È l'inizio della cosiddetta **semantica quantitativa**.

Lo stesso Lasswell nel 1942 fu incaricato dal governo americano di valutare gli eventuali contenuti di propaganda antiamericana da parte del periodico *The Galileian*.

**Lasswell** parte dall'ipotesi che il linguaggio costituisca un pilastro del potere. Secondo Lasswell, l'elaborazione del mito politico non può prescindere dalla costruzione di **simboli-chiave**.

1949 (*Language of politics*): Studio degli slogan adottati dalla propaganda ufficiale sovietica in occasione della Festa del Primo Maggio (studio quantitativo: calcolo dei coefficienti di correlazione fra la ricorrenza di dati simboli e l'accento posto su di essi in diversi periodi storici, a partire dalla istituzione dell'Unione delle Repubbliche Socialiste Sovietiche; creazioni di undici diverse categorie relative a particolari tipi di "simboli-chiave").

In occasione della **Hallerton House Conference** (1955), l'analisi del contenuto diventa oggetto di una serrata discussione. Si accusavano le tecniche di analisi del contenuto di quantofrenia e di inadeguatezza concettuale.

Lo sviluppo dei **programmi informatici** ha dato nuova linfa alle tecniche di analisi del contenuto a partire dalla fine degli anni '60.

### *Definizioni*

Stando al noto modello di **Harold Lasswell** sui mezzi di comunicazioni di massa, l'analisi del contenuto avrebbe dovuto riguardare il "che cosa" (*what?*) del modello a 5 W.

**Franco Rositi**<sup>2</sup> dà la seguente definizione di analisi del contenuto: "un insieme di metodi che sono orientati al controllo di determinate ipotesi su fatti di comunicazione e che a tale scopo utilizzano procedure di scomposizione analitica e di classificazione, normalmente a destinazione statistica, di testi e di altri insiemi simbolici".

**Stone** definisce l'analisi del contenuto come una "tecnica di ricerca che identifica in modo sistematico e oggettivo determinate caratteristiche presenti in un testo"

**Cartwright** dà questa definizione: *analisi del contenuto e codifica sono sinonimi attraverso i quali ci riferiamo alla descrizione oggettiva, sistematica e quantitativa di qualsiasi comportamento simbolico*.

---

<sup>1</sup> Dowring K., 1955, «Quantitative semantics in 18<sup>th</sup> century Sweden», in *Public opinion quarterly*, XVIII, 4, pp.389-394

<sup>2</sup> Rositi F., 1988, «L'analisi del contenuto», in *La ricerca sull'industria culturale*, a cura di F. Rositi e M. Livolsi, La Nuova Italia Scientifica, Roma

La definizione di **Bernard Berelson** (1952): "La *content analysis* è una tecnica di ricerca per la descrizione obiettiva, sistematica e quantitativa del contenuto manifesto della comunicazione" è una delle più note e citate. Essa è stata sottoposta a diverse critiche: innanzitutto esso di obiettivo non ha che l'oggetto e le procedure di ricerca che possono essere ripetibili e standardizzabili; in secondo luogo, assai più di frequente l'a.d.c. viene utilizzata non per rilevare il contenuto manifesto di un messaggio, bensì quello latente; **Rositi** fa notare come si tratti in realtà non di una tecnica, bensì di un insieme di tecniche; anche il concetto di descrizione è suscettibile di qualche perplessità: gli atti comunicativi non sono mai univoci, ma piuttosto plurivoci ed equivoci e quindi di difficile descrizione.

Oggi, con la locuzione "analisi del contenuto" intendiamo tanto l'approccio "classico", legato all'analisi interpretativa del testo sulla base di categorie (ex-ante o ex-post), quanto l'analisi strettamente **testuale** di tipo lessicometrico<sup>3</sup>.



Numerose controversie si aprirono in merito all'antinomia tra ricerche *qualitative* e *quantitative*. A sostegno delle prime si schierarono **Albert George** e **Kracauer** (autore di un'importante ricerca sull'evoluzione della cinematografia tedesca e sulla propaganda nazista), mentre le seconde, largamente utilizzate negli Stati Uniti nel corso degli anni Cinquanta, furono il cavallo di battaglia di **Lasswell**.

**Krippendorff**<sup>4</sup> distingue tre tipi di disegno della ricerca:

- quelli che tendono a stimare alcuni fenomeni in un certo contesto di dati rappresentano il tipo fondamentale;
- quelli che si propongono di saggiare la sostituibilità di un metodo per mezzo dell'analisi del contenuto;
- quelli per la verifica di ipotesi, che mettono a confronto i risultati ottenuti mediante l'analisi del contenuto con i dati ottenuti indipendentemente e riguardanti fenomeni non inferiti per mezzo della stessa ricerca;

Secondo **Irving Janis**, esistono diverse strade che portano alla *content analysis*:

- Analisi del contenuto pragmatica (classificazione dei segni a seconda delle loro probabili cause o dei loro probabili effetti)
- Analisi del contenuto semantica
- Analisi delle designazioni
- Analisi delle attribuzioni
- Analisi delle asserzioni
- Analisi dei veicoli segnici (classificazione del contenuto a seconda delle proprietà psicofisiche dei segni)

I passi fondamentali dell'analisi del contenuto sono i seguenti:

- Produrre i dati
- Estrarre il campione dai documenti
  - I campionamenti possibili sono i seguenti:
    - Campione casuale
    - Campione stratificato: il campione riflette distinzioni note a priori nell'ambito della popolazione
    - Campione sistematico
    - Campione a grappolo: i grappoli hanno confini naturali; si campiona prendendo interi "chicchi" all'interno dei quali sono contenute le unità
    - Campione a due o più stadi, che è una modificazione del campionamento a grappolo. Se per esempio si volesse estrarre un campione di famiglie italiane con un campionamento a due stadi, si può

<sup>3</sup> Tuzzi A., 2003, *L'analisi del contenuto. Introduzione ai metodi e alle tecniche di ricerca*, Carocci, Roma

<sup>4</sup> Krippendorff K., 1980, *Analisi del contenuto. Introduzione metodologica*, Eri, Torino

considerare che ciascuna famiglia é residente in un comune. Dapprima vengono estratti i comuni e quindi, da ciascun comune estratto, si sorteggiano le famiglie

- Definire il contenuto delle categorie  
Le categorie devono essere mutamente esclusive, basate su un unico fondamento logico e non sovrapponibili
- Definire l'unità di analisi
- Definire le unità di campionamento
- Definire le unità di registrazione: sono le parti separatamente analizzabili di un'unità di campionamento; queste si possono suddividere in:
  - Unità fisiche: per es., i singoli fotogrammi di un film
  - Unità sintattiche: le parole (nel caso di mezzi di comunicazione verbali) oppure le azioni i spettacoli teatrali, gli avvenimenti di un dramma, le riprese di un film (nel caso di mezzi di comunicazione nonverbali)
  - Unità di riferimento: definite in base a particolari oggetti, avvenimenti, persone, azioni, Paesi o idee a cui un'espressione si riferisce
  - Unità di proposizione (e nuclei): **Osgood** propone di studiare tutte le proposizioni che rientrano in uno dei due modelli che seguono:
    - § oggetto di atteggiamento / connettore verbale / termine valutativo
    - § oggetto di atteggiamento / connettore verbale / oggetto di atteggiamento
  - Unità tematiche: **Katz** fa l'esempio di unità tematiche da lui utilizzate che riguardavano le richieste alle autorità per ottenere favori o esenzioni in Israele; queste unità sono, secondo **Krippendorff**, le più idonee all'analisi del contenuto, ma vengono spesso evitate a causa dello sforzo cognitivo che richiedono nella loro identificazione
  - Unità di contesto: circoscrivono l'informazione contestuale che può entrare nella descrizione dell'unità di registrazioneLe unità di registrazione nel loro complesso convogliano l'informazione presente nell'unità di campionamento e forniscono la base per l'analisi. Le unità di contesto infine interessano il processo di descrizione delle unità di registrazione
- Definire il sistema di enumerazione
- Registrare i dati
- Adattare i dati
- Trarre inferenze (secondo Krippendorff)
- Analizzare i dati

### ***Le tecniche di analisi del contenuto***

#### L'analisi del contenuto quantitativa

##### *Analisi delle frequenze*

Nel 1949, Yakobson e **Lasswell** si proposero di analizzare gli *slogans* conati in occasione della festa del Primo Maggio in Unione Sovietica, nel periodo compreso tra il 1918 e il 1943.

I due autori partivano dalla convinzione che gli eventuali mutamenti intervenuti nel contenuto degli *slogans* riflettessero i più generali mutamenti della linea politica ufficiale del Partito Comunista in quegli stessi anni.

Utilizzando come unità di registrazione il simbolo-chiave, essi individuarono 11 categorie di classificazione:

- Simboli rivoluzionari (dittatura del proletariato, rivoluzione d'ottobre, rivoluzione mondiale)
- Simboli controrivoluzionari (fascismo, capitalismo, borghesia)
- Simboli nazionali (la nostra terra, sicurezza interna, paria)
- Simboli universali (umanità, internazionale comunista)
- Simboli di politica interna (piano quinquennale, fattorie di stato, disoccupazione)
- Simboli di politica estera (atti ufficiali dell'URSS nelle relazioni internazionali)
- Simboli di gruppo sociale (Armata Rossa, intellettuali, proletari)

- Nomi propri di persone (Lenin, Marx, Hitler)
- Simboli liberali tradizionali (cittadino, libertà, democrazia)
- Simboli morali (lealtà, onestà, responsabilità)
- Simboli d'azione (viva! abbasso! vittoria!)

### Analisi delle contingenze

Nel 1959 Osgood propose una procedura che presupponeva la classificazione delle unità di registrazione e che consisteva nel determinare la presenza contemporanea di una o più unità di categorie diverse in uno stesso testo o segmento di testo assunto come unità di contesto.

L'analisi delle contingenze può risultare utile per rilevare tratti di personalità, atteggiamenti, rappresentazioni sociali, intenzioni "latenti" del produttore del testo.

**Tabella 1 - Matrice delle presenze dei simboli chiave nelle unità di analisi**

	A1	A2	.....	An
U1	+			+
U2		+		
.				
.				
.				
.				
Un		+		+

### Indici verbali

- **Indici lessicali** (Parole non classificate in categorie);  
es.: TTR (*Type Token-ratio*) viene usato nelle ricerche in campo psichiatrico e come indicatore di competenza linguistica o di livello culturale
- **indici grammaticali** (Parole classificate in categorie grammaticali)  
es.: VAQ (*verbe/adjective quotient*), quoziente verbi-aggettivi. Nella ricerca psichiatrica viene usato come indicatore di disadattamento e di instabilità affettiva<sup>5</sup>  
Osgood & Walker<sup>6</sup> (*Motivation and language behaviour*) usano il rapporto  $(N+V)/(A+Av)$  come indicatore di stereotipia lessicale
- **indici semantici** (Parole classificate in categorie di significato)  
es.: DRQ (*Discomfort relief quotient*), usato da Dollard e Mowrer<sup>7</sup> come indicatore di tensione emotiva  $M/M+B$  (M = parole esprimenti malessere e il totale complessivo di parole esprimenti benessere e malessere)

<sup>5</sup> Amerio P., 1966, «Studio sul comportamento verbale mediante analisi quantitativa del rapporto verbi-aggettivi», in *Rendiconti*, n. 13, pp.3-42

<sup>6</sup> Osgood C.E. e Walker E.G., 1959, «Motivation and language behaviour: content analysis of suicide notes», in *Journal of abnormal social psychology*, 59, pp.58-67

<sup>7</sup> Dollard J. E Mowrer O.H., 1947, «A method of measuring tension in written documents», in *Journal of abnormal social psychology*, 42, pp.3-32

Classe di indici	Esempio	A cosa servono
Indici grammaticali	<ul style="list-style-type: none"> <li>• VAQ quoziente verbi/aggettivi</li> <li>• N+V/A+Av (somma dei nomi e dei verbi rapportata alla somma degli aggettivi e degli avverbi)</li> </ul>	<ul style="list-style-type: none"> <li>• Come indicatore di disadattamento e di instabilità affettiva. Più sono alti i valori, più sono disadattati i soggetti.</li> <li>• A individuare eventuali differenze tra gli autori o i testi considerati all'interno di ciascun genere.</li> <li>• A individuare eventuali differenze tra testi di diverso genere di uno stesso autore.</li> <li>• A individuare eventuali differenze tra generi diversi indipendentemente dal livello "letterario" dei singoli testi o dei singoli autori</li> <li>• Stereotipia lessicale.</li> </ul>
Indici lessicali	TTR (rapporto tra numero di parole diverse e numero di parole complessivamente utilizzate)	Misura la varietà del linguaggio, la stereotipia lessicale, è usato in campo clinico come indicatore di disagio psicologico
Indici semantici	DRQ (Discomfort relief quotient) dato da M/M+B (somma delle parole esprimenti Malessere e Benessere)	Indicatore di tensione emotiva

#### Analisi delle valutazioni

- Indice di squilibrio (Contenuto rilevante = Neutro + Favorevole + Sfavorevole); a questo indice possono essere attribuiti punteggi che variano o da 1 a +1, oppure da 3 a +3; si hanno due indici, uno di squilibrio favorevole, l'altro di squilibrio sfavorevole:

$$IF = (F2 Fs)/rt=(fs)/rf/t$$

$$IS = (SFS2)/rt=(fs)/r. S/t$$

- Analisi degli asserti valutativi (Osgood)

#### Analisi del contenuto computerizzata

Si basa sull'applicazione di programmi informatici dotati di vocabolari gestibili dall'utente o automatici.

Tra questi:

- il **General Inquirer** (Harvard, anni '60);
- lo **Stanford Political Dictionary**, basato sul differenziale semantico. Due fasi: 1) I giudici decidono a quale dimensione ricondurre la parola; 2) i giudici stabiliscono l'intensità (il punteggio)

#### Analisi delle corrispondenze lessicali

L'analisi delle corrispondenze lessicali (ACL) fa riferimento ad alcune tecniche per l'analisi quantitativa dei dati testuali che sono state messe a punto nell'ambito della cosiddetta "scuola" francese dell'*Analyse des données*<sup>8</sup>.

L'ACL consente operazioni di 3 tipi:

1. sintesi dell'informazione contenuta nei dati
2. visualizzazione delle associazioni multiple tra parole
3. connessione tra dati testuali e dati di contesto

<sup>8</sup> Amato E., 1988, «L'analisi delle corrispondenze lessicali: una proposta per il trattamento automatico di dati testuali», in *La ricerca sull'industria culturale*, a cura di F. Rositi e M. Livolsi, La Nuova Italia Scientifica, Roma, p.96

	<b>Aggettivi</b>	<b>Pronomi</b>	<b>Verbi ausiliari</b>	<b>V. non ausiliari</b>	<b>Totale</b>
<b>Autore 1</b>	2	10	15	5	32
<b>Autore 2</b>	18	2	6	6	32
<b>Autore 3</b>	8	6	8	16	38
<b>Autore 4</b>	6	12	20	3	41
<b>Autore 5</b>	7	9	6	18	40
<b>Totale</b>	41	39	55	48	183

L'ACL consente inoltre di ricavare dal corpus testuale le frasi "tipiche" dell'intero corpus", seguendo due criteri tra loro alternativi:

- quello della distanza dal chi quadrato e
- quello della frequenza delle parole

L'ADC deve essere considerata come l'applicazione della tecnica delle analisi delle corrispondenze multiple (ACM) a variabili lessicali (parole). Essa si configura pertanto come tecnica di analisi dei dati piuttosto che come tecnica di rilevazione. Tuttavia, essa condivide con gli approcci lessicometrici alcuni aspetti, come per esempio la distinzione preliminare tra **word token** e **word type**.

