

Metodi misti di analisi dei testi come terreno di integrazione tra spiegazione e interpretazione

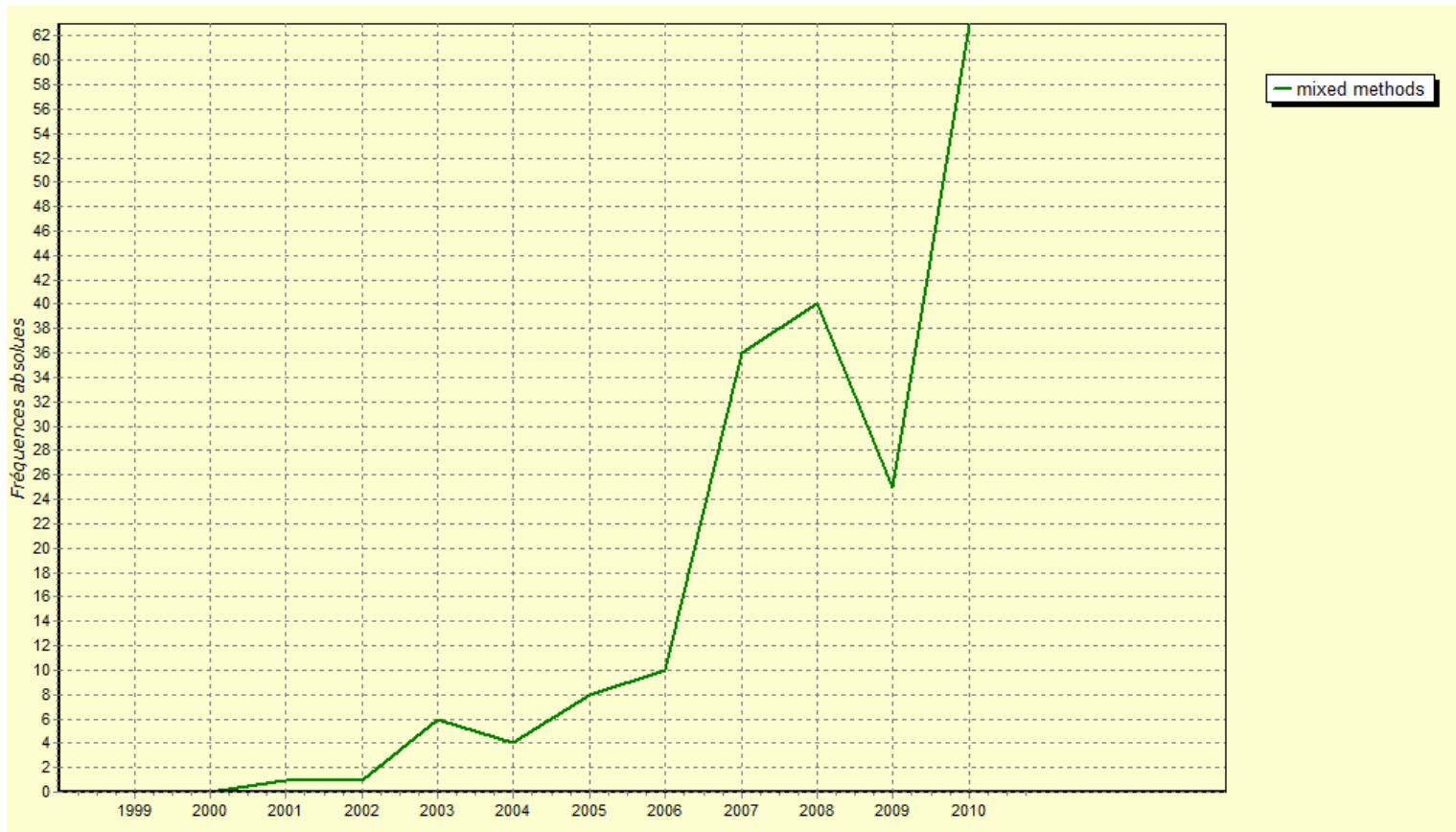
Luca Giuliano



SAPIENZA
UNIVERSITÀ DI ROMA

Dipartimento di Scienze statistiche

Mixed methods: una serie temporale di dieci anni



*Fig. 1 – Mixed methods negli abstracts del Sociological Abstract (CSA) dal 2000 al 2010
(197 Frequenze assolute per 17.615 schede dal 1952 al 2010)*

La nascita “ufficiale” dei metodi qualitativi

Of course, qualitative research is inherently multimethod in focus. However the use of multiple methods, or triangulation, reflects an attempt to secure an in-depth understanding of phenomenon in question. **Objective reality can never be captured. We know a thing only through its representation.**

(Denzin, N. K., & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In N. K. Denzin and Y. S. Lincoln (Eds.), *Handbook of qualitative research*. Thousand Oaks, CA: Sage, pag. 5)

Triangolazione

Triangulation is not a tool or a strategy of validation but an alternative to validation.

(Denzin, N. K., & Lincoln, Y. S. (1994). Introduction: Entering the field of qualitative research. In N. K. Denzin and Y. S. Lincoln (Eds.), *Handbook of qualitative research*. Thousand Oaks, CA: Sage, pag. 5)

I have repeatedly suggested that the sociologist should examine his problem from as many different methodological perspective as possible.

(Dezin, N. (1970), Strategies of multiple triangulation. In Denzin, N. (ed.), *The Research Act in Sociology: a Theoretical Introduction to Sociological Method*, New York, McGraw-Hill, p. 297)

Una tassonomia della triangolazione

Triangolazione dei dati	Riguarda le fonti di dati, survey, serie temporali, ecc. E' tipico dell'analisi secondaria.
Triangolazione del ricercatore	Riguarda la collaborazione tra diversi ricercatori con competenze e punti di vista diversi che replicano l'uno il lavoro dell'altro su ricerche specifiche.
Triangolazione teorica	Consiste nell'assumere quadri teorici differenti, anche dal punto di vista disciplinare.
Triangolazione metodologica	E' quella in cui si combinano metodi diversi sia nel momento della rilevazione-osservazione che dell'analisi. Tipico è il cross-over qualitativo-quantitativo. Non è il solo, ma rappresenta una anticipazione dei <i>mixed methods</i> .

Downward , P., Mearman, A. (2007), Retroduction as mixed-methods triangulation in economic research: reorienting economics into social science, *Cambridge Journal of Economics* , 31, p. 81.

L'integrazione dei metodi dal punto di vista delle “pratiche di ricerca”

1 - Ipotesi/teorie definite, strutturate, generalizzabili, per le quali ci attendiamo che vi siano delle proprietà classificabili finite, chiuse e ripetibili che vogliamo mettere alla prova (**approccio ipotetico-deduttivo**)

2 – Abbiamo solo alcuni indizi e tentiamo di “inventare” delle ipotesi plausibili perché non conosciamo abbastanza il nostro oggetto di studio e **siamo portati a considerare ogni “oggetto” del nostro interesse come unico**. La nostra ricerca sarà allora per “classi aperte” in cui ogni classe, definite certe proprietà, contiene provvisoriamente (e potenzialmente) un solo oggetto ritenuto temporaneamente irripetibile (**approccio ipotetico-induttivo**).

Le due “culture”

Fin da piccoli veniamo abituati alla separazione di 2 culture:

- quella (più nobile?) rappresentata dal mondo del *quaderno a righe*, dove vivono le parole, i discorsi, le narrazioni e le descrizioni, prevalentemente “qualitative”;
- quella (più pratica?) rappresentata dal mondo del *quaderno a quadretti*, dove vivono i numeri, i calcoli, i diagrammi, gli schemi, le figure geometriche: tutti oggetti più o meno “quantitativi”.

(La Torre, M. (2005), *Le parole che contano. Proposte di analisi testuale automatizzata*, Franco Angeli, Milano, p. 13)

Gestire le informazioni contenute nei testi



- Privilegiare le parole rispetto ai numeri (ermeneutica)
- Sottoporre le parole al dominio dei numeri (dati testuali)

Testi e dati testuali



“Chiamiamo **testo** ogni espressione o insieme di espressioni fissate mediante la scrittura” (Ricoeur, 206, p. 213).

Dati espressi in linguaggio naturale provenienti dalle fonti più diverse : indagini sul campo (domande aperte o interviste), analisi di frammenti o testi corti (abstract, bibliografie, manifesti , messaggi) raccolti in una collezione di documenti costituenti un **corpus di dati testuali** (Bolasco, 2004, pag. 10).

Cos'è un testo?

“Chiamiamo testo ogni espressione o insieme di espressioni fissate mediante la scrittura” (Ricoeur, P. 2006, p. 213).

“Ma cosa viene fissato nella scrittura?”

Dietro ogni scrittura vi è un discorso (*speech*).

Priorità psicologica del discorso rispetto alla scrittura: l'esperienza umana del parlante.
Priorità sociologica del discorso: contesto economico, politico e culturale del linguaggio.

Il testo prende il posto del parlare.
Un lettore prende il posto dell'ascoltatore.

DISCORSO = TESTO + CONDIZIONI DI PRODUZIONE
TESTO = DISCORSO – CONDIZIONI DI PRODUZIONE

“Les territoires respectifs apparaissent dans un rapport d'inclusion, la linguistique textuelle est alors un sous-ensemble du vaste domaine de l'analyse du discours” (Jaubert, 2002, p. 77).

Dalla singolarità del testo alla pluralità del corpus

- “In questo ambito di ricerca si preferisce parlare di *corpus* di testi, indicando con questo termine una qualsivoglia **raccolta di frammenti testuali**, fra loro **confrontabili** sotto qualche **punto di interesse**” (Bolasco, 1997, p. 167).

Il passaggio dal singolare al plurale mette in crisi l'opposizione concreto/astratto apparentemente risolta dalla opposizione discorso / testo.

Le scienze sociali (compresa la linguistica) non si occupano della singolarità ma della pluralità.

- “Content analysis is a research technique for making **replicable** and **valide** inferences from **textes** (or other meaningful matter) to the context of their use” (Krippendorff, 2004, p. 19).

Teoria \Rightarrow corpus \Rightarrow dati testuali

Un corpus è sempre il risultato di un intervento operativo. Un corpus è “costruito” per mezzo di una selezione di proprietà definite da un punto di vista teorico. Un corpus è inscindibile da una teoria.

- Corpus dei dati testuali grezzi
- Corpus dei dati testuali pertinenti
- Corpus dei dati testuali costruiti

Spiegazione

- “Il compito principale delle scienze sociali è spiegare i fenomeni sociali. Non è il loro solo mandato, ma è il più importante, e gli altri gli sono subordinati e ne dipendono” (Elster, 31).
- “Spiegare un evento vuol dire rendere conto del perché sia accaduto citando un evento *precedente* come causa” (Elster, pag. 31).

Interpretazione

- “Interpretare un’azione significa spiegarla nei termini delle credenze e dei desideri (motivazioni) antecedenti dell’agente.” (Elster, p. 82).

Universi dell'esperienza

- L'universo di ciò che può essere pienamente rappresentabile.
- L'universo della bruta effettività dei fatti.
- L'universo dei segni che stabiliscono relazioni e connessioni tra gli oggetti che appartengono agli altri due universi.

Metodologia della interpretazione e della spiegazione

	Interpretazione	Spiegazione
Domande di ricerca	Come si comportano gli attori sociali nei processi in cui sono coinvolti?	Quali regolarità di comportamento caratterizzano una popolazione di attori sociali?
Relazioni/conessioni	Sostantive, empatiche	Formali, basate sulla similarità/differenza
Descrizione	Genetica, narrativa, storica	Tassonomica
Obiettivo	Capire i nessi dal punto vista del senso che vi attribuiscono gli attori (motivazioni)	Determinare un modello causale (probabile) rappresentativo e generalizzabile
Metodi	Ricercatori sul campo, focus sul contesto, etnografie, analisi qualitative	Ricerca su larga scala, campioni rappresentativi, analisi statistiche
Punti deboli	Impossibilità di generalizzare e replicare	Generalizzazione condizionata dal rispetto dei vincoli probabilistici
Punti forti	Rappresentatività delle proprietà specifiche dei fatti	Replica, validazione e falsificazione delle ipotesi
Risultati	Conclusioni singolari	Conclusioni esaustive e necessarie

Ermeneutica quantitativa

- Cosa **non** è l'ermeneutica quantitativa: non è sottrazione della soggettività all'ermeneutica ed espropriazione del soggetto da parte del ricercatore. Non è l'"ermeneutica oggettiva" di Ulrich Oevermann (1979) secondo il quale colui che produce il testo non ne può afferrare il senso profondo perché la passione o l'ideologia che lo pervade lo nascondono alla sua vista e pertanto deve essere il ricercatore, come soggetto estraneo, a far emergere il significato oggettivo o latente che è portatore di senso (inteso come "inconscio collettivo").
- Per ermeneutica quantitativa intendo invece un disegno della ricerca che si pone l'obiettivo di spiegare il testo (e il comportamento linguistico che ne è l'origine) in presenza di una informazione di massa che lo sottrae alla sua relazione privilegiata con un lettore (testi senza lettori).
- Per ermeneutica quantitativa intendo la restituzione di un momento interpretativo (quello del lettore) in un percorso fondamentalmente esplicativo. E' la qualità che getta una luce sulla quantità in vista di una ricomposizione.

Festival di Sanremo: 1951-2010

- Corpus: 1.758 testi di canzoni (1.792 Kb) per un totale di 342.000 occorrenze e 17.000 forme grafiche.

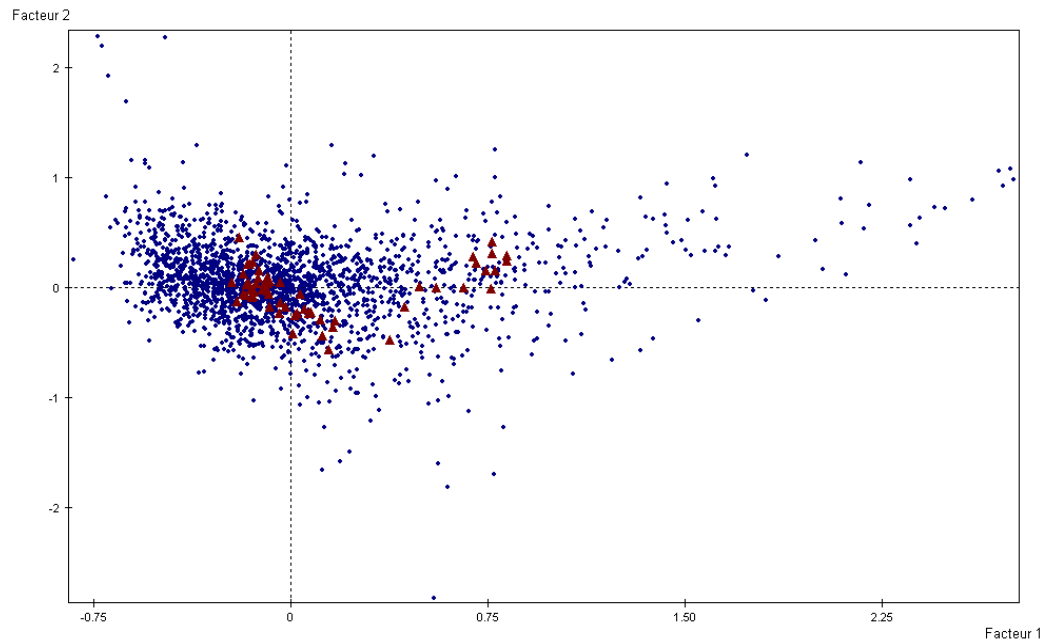


Fig. 2 - Analisi delle corrispondenze della tabella lessicale Sanremo2010 forme per testi: rappresentazione delle forme (soglia $N \geq 12$) sul piano fattoriale F1-F2

Festival di Sanremo: evoluzione cronologica

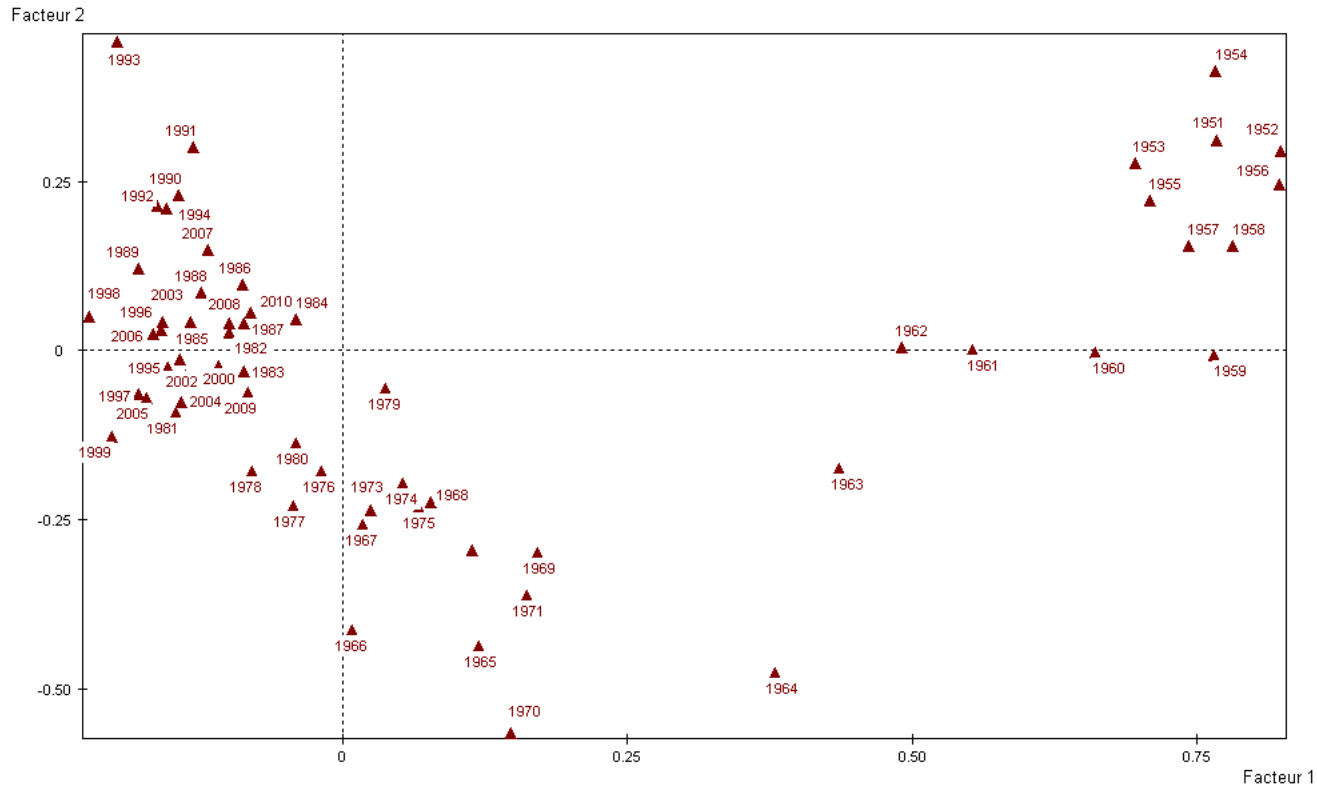


Fig. 3 - Analisi delle corrispondenze della tabella lessicale Sanremo 2010 forme per testi: rappresentazione delle frequenze attive (anno di edizione) sul piano fattoriale F1-F2

Amore: evoluzione cronologica

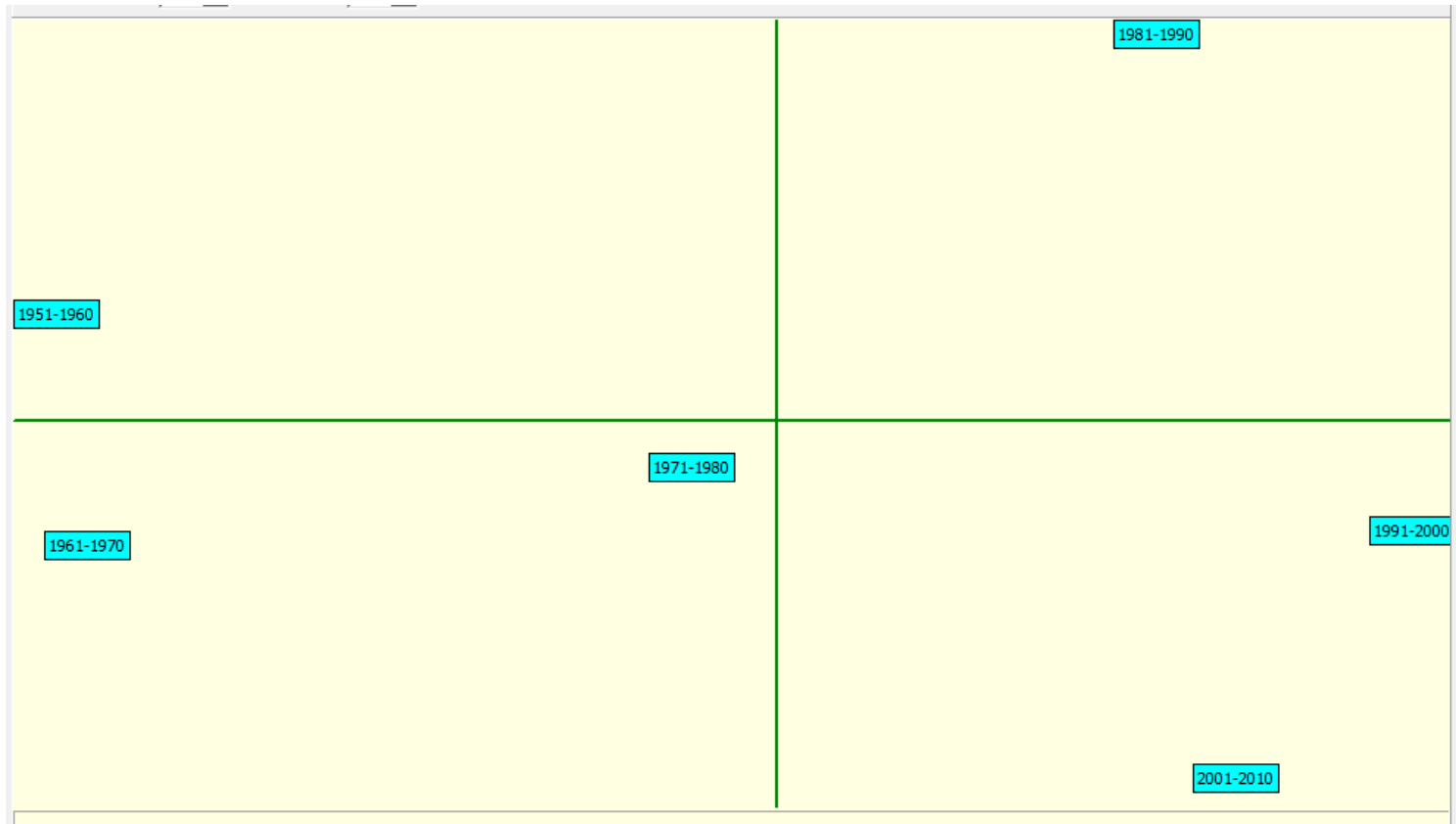


Fig. 4 - Analisi delle corrispondenze della tabella lessicale Concordanze - Amore -Sanremo 2010 forme per testi: rappresentazione delle frequenze attive (periodi decennali) sul piano fattoriale F1-F2

Amore: specificità per periodi decennali

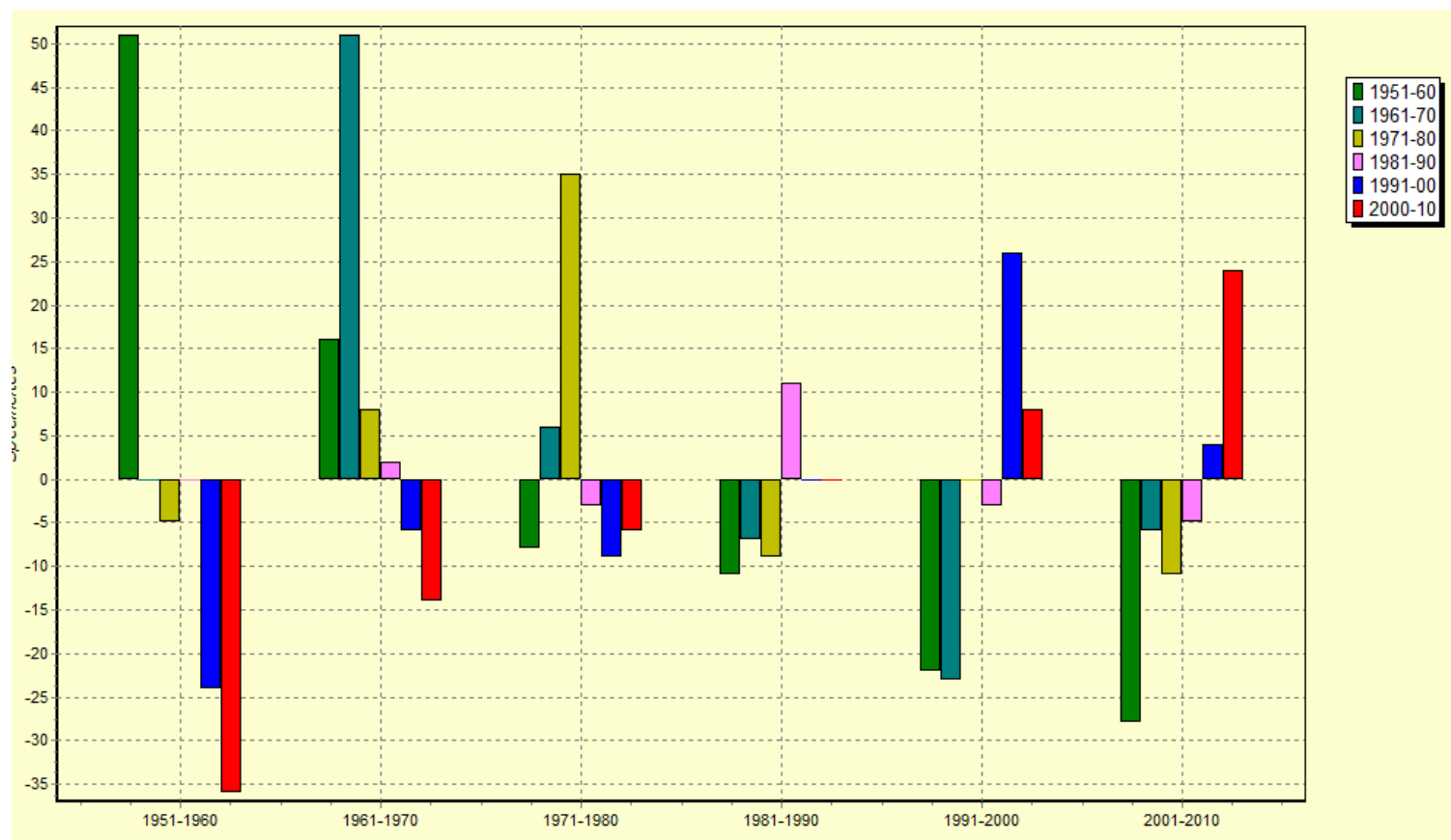
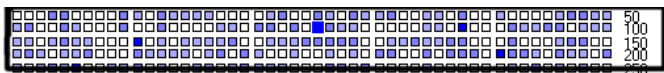
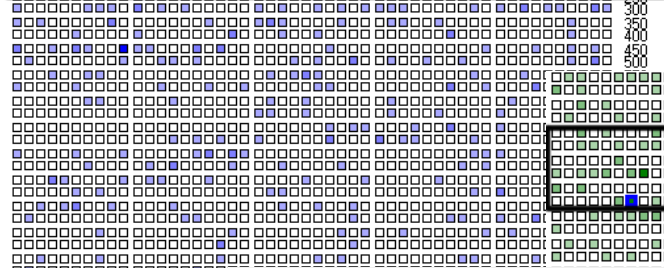


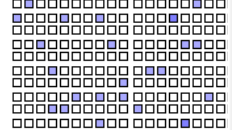
Fig. 5 – Gruppi di forme costituiti dalle prime 10 parole specifiche per ciascun decennio sulla base del sub-corpus Sanremo-Amore e riposizionati sul corpus totale.



1951-60



1961-70



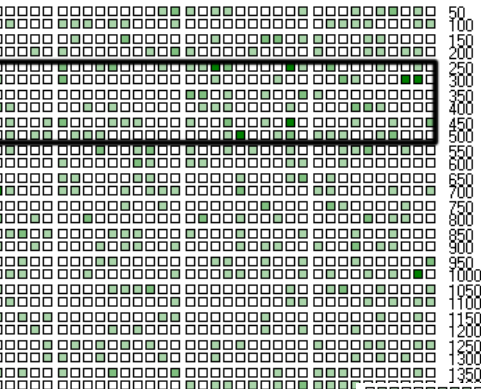
Nom du groupe :

Le motif :

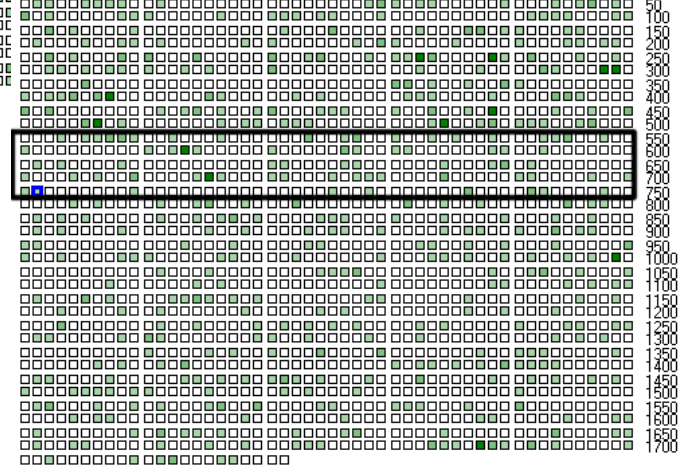
Ajouter

Rechercher Enregistrer

Supprimer Charger



1971-80



Forme	Fréquence
ciao	136
pizza	22
sole	673
taxi	24
piangere	111
caro	59
brucio	9
caldo	49
gondola	36
chiama	53

Analisi di Blog sull'anoressia (Pro-Ana)

Costruzione del corpus:

- ✓ Selezione di 11 blog (410 post) per un corpus di 57.132 occorrenze

Analisi lessicometrica:

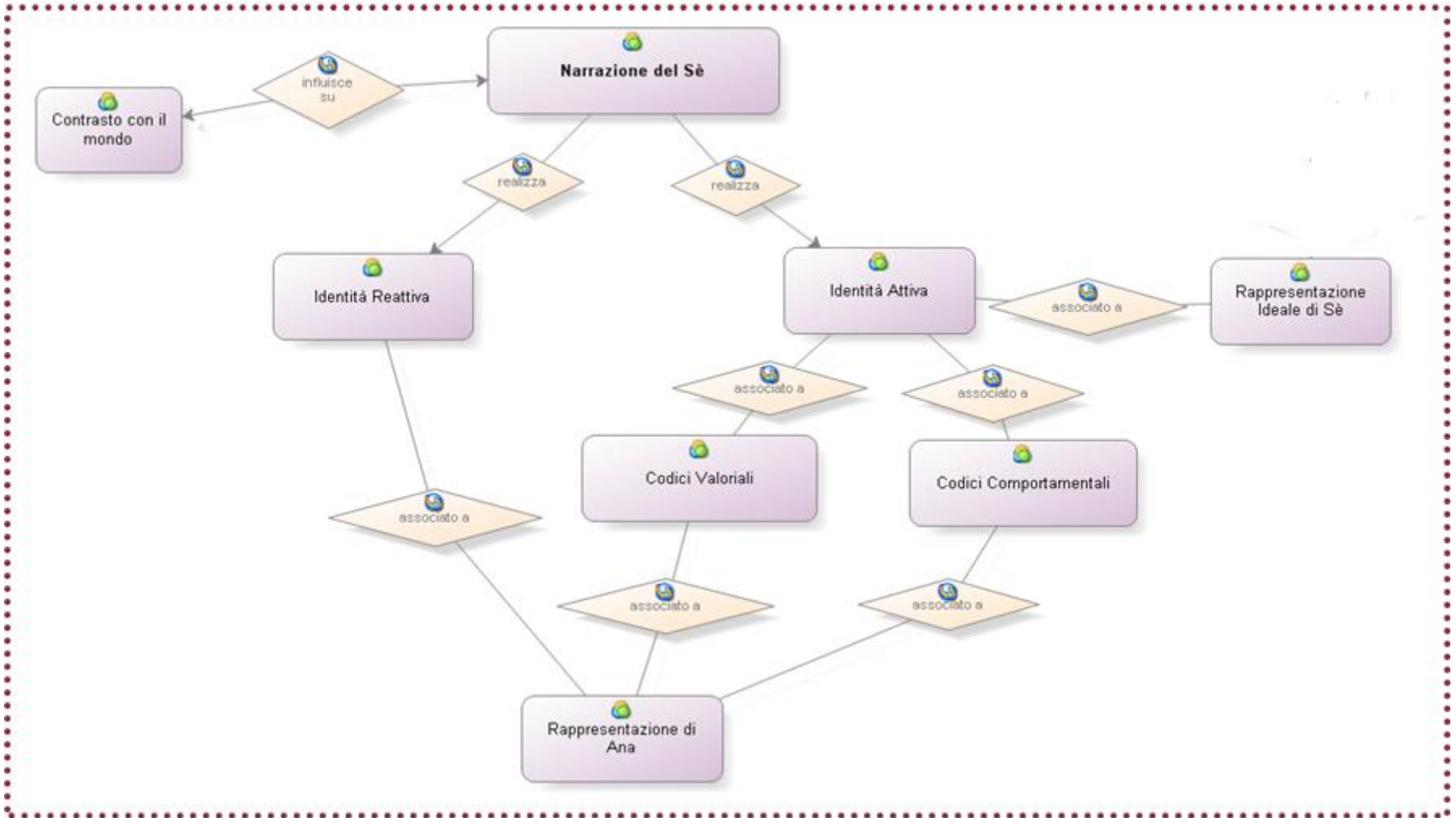
- ✓ Elaborazione del corpus in TaLTac2: pre-trattamento, parsing, normalizzazione, lessicalizzazione ed estrazione del Linguaggio Peculiare (LIPE)

Analisi ermeneutica:

- ✓ Esportazione del corpus in N_Vivo e codifica delle porzioni di testo contenenti il LIPE
- ✓ Raggruppamento dei codici in Set (macrocodici) ed individuazione delle relazioni tra di essi
- ✓ Rappresentazione sintetica delle suddette relazioni in una mappa concettuale

D'Amen, B. – Pontecorvo, M.E. (2011) . L'anoressia che fa rete: un'integrazione degli approcci lessicometrico ed ermeneutico nell'analisi del fenomeno Pro-Ana,

Mappa concettuale (N-Vivo)



Per concludere...

(...) comprensione e spiegazione non sono due metodi tra loro opposti. A rigore, solo la spiegazione è metodica. La comprensione è il momento non metodico che precede, accompagna e ingloba la spiegazione. Per contro, la spiegazione sviluppa analiticamente la comprensione (P. Ricoeur, 1987, p. 90).

Le analisi dei dati testuali centrate sulle forme grafiche **sono indipendenti dalla lingua**. Si tratta di un approccio **puramente formale** che privilegia i **segni** (significanti) per arrivare al **senso** (in quanto insieme di significati) come rappresentazione del contenuto/discorso (S. Bolasco, 2004, p. 10).