

# I processi cognitivi nell'analisi concettuale dei documenti: una postilla tra biblioteconomia e linguistica

di Alberto Cheti

## Premessa

La nozione di 'soggetto', il modello della 'relazione di soggetto' tra un'opera, i temi di cui tratta e i concetti che ne fanno parte, la natura e la valenza dei processi logici che sottostanno alle operazioni di analisi concettuale dei documenti, questi argomenti sono stati oggetto di indagine del GRIS (Gruppo di ricerca sull'indicizzazione per soggetto dell'Associazione italiana biblioteche), attraverso un percorso di cui si richiama qui le principali tappe<sup>1</sup>.

*Testo e contesto nell'analisi concettuale dei documenti* uscì nel 1994 come contributo per gli *Scritti in onore di Diego Maltese*, ma le proposte che vi sono contenute erano già state presentate qualche anno prima, nel maggio del 1991, durante un corso di aggiornamento organizzato dalla Sezione toscana dell'AIB presso la Biblioteca comunale di Bagno a Ripoli. Fu quella la prima di una lunga serie di esperienze formative dedicate a questo tema, compreso un seminario che tenni nell'anno accademico 1999-2000 al Corso di laurea in conservazione dei beni culturali presso l'Università di Pisa. Da queste esperienze è nato anche il *Manuale ipertestuale di analisi concettuale*, frutto di un'intensa attività di aggiornamento professionale sui temi dell'analisi concettuale dei documenti e dell'indicizzazione per soggetto, promossa dall'Università degli studi di Bologna tra il 1992 e il 1996<sup>2</sup>.

ALBERTO CHETI, 50053 Empoli (FI), e-mail chetialberto@infinito.it.

**1** La riflessione del GRIS è a sua volta debitrice dei seminari di aggiornamento sull'analisi concettuale dei documenti tenuti da Daniele Danesi, Luigi Crocetti e Diego Maltese negli anni Ottanta, a partire da quello organizzato dalla Sezione Lazio dell'Associazione italiana biblioteche (9-11 febbraio 1983).

**2** Alberto Cheti, *Testo e contesto nell'analisi concettuale dei documenti*. In: *Il linguaggio della biblioteca: scritti in onore di Diego Maltese*, raccolti da Mauro Guerrini. Firenze: Regione Toscana, Giunta regionale, p. 237-239 (pubblicato successivamente in: *Il linguaggio della biblioteca: scritti in onore di Diego Maltese*, a cura di Mauro Guerrini. Milano: Bibliografica, 1996, p. 833-855). *Manuale ipertestuale di analisi concettuale*, realizzazione e strutturazione ipertestuale di Serena [Serafina] Spinelli, <<http://www.sba.unifi.it/CMpro-v-p-46.html>> (ultima consultazione: 01/10/2015). Allo stesso URL è consultabile una versione dei materiali didattici impiegati nei corsi sull'indicizzazione per soggetto secondo il metodo GRIS, compresi quelli relativi all'analisi concettuale dei documenti.

In termini generali, riferibili a qualsiasi campo dell'agire pubblico, istituzionale, *Testo e contesto* cercava di rispondere all'esigenza di analizzare il processo, le fasi di cui si compone, gli elementi in gioco, le criticità rilevate, le variabili del contesto ecc., allo scopo di migliorarne la qualità dei risultati. Nello specifico, avanzava un'ipotesi di lavoro basata sull'utilizzo di alcune nozioni e metodi tratti, in modo semplificato e scevro da formalismi, dalla linguistica testuale. In questa prospettiva, proponeva anche una rilettura di alcuni passi essenziali della norma UNI ISO 5963 sull'analisi concettuale. Il testo e le sue proprietà, i temi e le loro relazioni, il tema di base, la nozione di macrostruttura, gli elementi del contesto, il processo di analisi, la natura e la valenza degli atti cognitivi sottostanti, le macro-operazioni di sommarizzazione: queste le principali nozioni presenti in *Testo e contesto*, illustrate anche attraverso alcune esemplificazioni.

Nella miscellanea Maltese, *Testo e contesto* è seguito da un contributo di Stefano Tartaglia dal titolo *Per una definizione di 'soggetto'*<sup>3</sup>, dove, a conclusione di un'ampia e puntuale rassegna, si propone una definizione dell'entità 'soggetto' incentrata sulla nozione di 'tema di base', che sarà sostanzialmente ripresa nella Guida GRIS<sup>4</sup>. Il 'tema di base' costituirà, infatti, una parola chiave nell'elaborazione del GRIS, tanto che, in anni più recenti, sarà al centro di una proposta di riformulazione della 'relazione di soggetto' in FRBR<sup>5</sup>. Ma il quadro non sarebbe completo senza considerare i contributi relativi alle fasi a valle del processo di indicizzazione, in cui le nozioni di 'concetto' e di 'soggetto' sono viste dal versante della ricerca, la quale in funzione delle esigenze dell'utente prende la forma di ricerca a due stadi, ossia per 'termine e stringa'<sup>6</sup>.

Questi contributi, sparsi nell'arco di un ventennio e nati in occasioni e con destinazioni diverse, forniscono un quadro sostanzialmente omogeneo, al quale ci possiamo riferire, per comodità, con l'espressione utilizzata da Leda Bultrini nella sua presentazione al 10<sup>th</sup> International ISKO Conference (Montréal, 2008): «un modello di *aboutness* elaborato dal GRIS». A questo quadro le note che seguono aggiungono solo qualche chiarimento, sollecitato anche dalle osservazioni di Piero Cavalieri in

**3** Stefano Tartaglia, *Per una definizione di 'soggetto'*. In: *Il linguaggio della biblioteca* cit., p. 789-801 (p. 856-867 nell'edizione dell'Editrice Bibliografica cit.).

**4** Associazione italiana biblioteche. GRIS – Gruppo di ricerca sull'indicizzazione per soggetto, *Guida all'indicizzazione per soggetto*. Roma: AIB, 1996 (ristampa con correzioni 2001).

**5** Raffaella Gaddoni; Federica Paradisi, *GRIS e modello FRBR* (dattiloscritto della relazione tenuta al convegno "L'indicizzazione per soggetto in Italia nell'epoca del rifacimento del Soggettario", 50° Congresso nazionale AIB, Roma, 30 ottobre 2003). Leda Bultrini; Alberto Cheti, *The Italian model* (dattiloscritto della relazione tenuta al convegno "International ISKO Conference Conceptual models of aboutness, Montréal, 5-8 agosto 2008). Alberto Cheti, *Il punto di vista del GRIS sulla 'relazione di soggetto'*, in: *Principi di catalogazione internazionali: una piattaforma europea? Considerazioni sull'IME ICC di Francoforte e Buenos Aires: atti del Convegno internazionale, Bibliocom, 51° Congresso AIB (Roma, 27 ottobre 2004)*, a cura di Mauro Guerrini. Roma: Associazione italiana biblioteche, 2008, p. 91-100.

**6** Andrea Fabbrizzi, *Indici semantici e ricerca nel catalogo elettronico* (testo dell'intervento al seminario "L'indicizzazione semantica nei cataloghi italiani", organizzato dalla Sezione Toscana dell'AIB e dalla Biblioteca nazionale centrale di Firenze, Firenze, Biblioteca nazionale centrale, 13 settembre 2004), <<http://www.iskoi.org/doc/catalogo.htm>> (ultima consultazione: 01/10/2015). Emanuela Casson; Andrea Fabbrizzi; Aida Slavic, *Subject search in italian OPACs: an opportunity in waiting?* In: *Subject access: preparing for the future*, edited by Patrice Landry [et al.], Berlin: De Gruyter Saur, 2011, p. 37-50.

*La biblioteca crea significato*<sup>7</sup>. Muovendo da un brano tratto dal *Manuale ipertestuale*, Cavalieri formula alcuni giudizi intorno alle nozioni ‘processi cognitivi’, ‘contenuto concettuale’, ‘aboutness’, ‘tema’, ‘temi particolari’, le quali presupporrebbero l’adesione a una gnoseologia di tipo mentalista, e conclude:

La possibilità che i temi possano essere individuati direttamente nel ‘contenuto concettuale’ richiede l’adesione forte ad una semantica e, soprattutto, ad una gnoseologia di tipo mentalista o idealistica. L’estrazione immediata di argomenti e temi da un documento, se si vuole sfuggire alla prospettiva solipsistica berkeleiana, richiederebbe, perché fosse possibile ottenere risultati condivisibili, l’accettazione di una prospettiva ontologica da ‘terzo mondo’ platonico, fregeiano o popperiano. Accettando queste visioni i temi risulterebbero riconoscibili direttamente e oggettivamente perché esistono, così come sono, sotto forma di concetti indipendenti dalla mente che li conosce, perché elementi di un mondo terzo rispetto alla realtà materiale e alle menti da cui chi parla o scrive li assume per esprimere ciò che è necessario per rappresentare il mondo<sup>8</sup>.

Sul mentalismo esprime preoccupazione anche Roberto Ventura in *La biblioteca tra linguaggio e rappresentazione*<sup>9</sup>, invitando a chiarire la valenza attribuita ai processi cognitivi di sommarizzazione del testo:

Una questione alla quale vale la pena accennare è quella relativa alla valenza che vogliamo attribuire ai processi cognitivi di sommarizzazione del testo: quando andiamo alla ricerca delle regole tramite le quali tale processo di riduzione è praticato, dobbiamo stare attenti a non cadere in una visione di carattere mentalistico, visione dalla quale l’approccio cognitivo non sarebbe del tutto esente: la rappresentazione del testo effettuata dall’indicizzatore presupporrebbe, in tale prospettiva, un’equazione con la rappresentazione della conoscenza; essendo il problema fondato sulla chiarificazione di come la mente funzioni, di quali regole la mente segua nel processo di rappresentazione indicale, è evidente che sommarizzare un testo significa, in pratica, sommarizzare concetti e conoscenze, entità immateriali che possono assumere forma testuale e che il linguaggio può pertanto adeguatamente rappresentare. Gli indici ci darebbero allora una rappresentazione contratta della conoscenza, la cui formazione seguirebbe regole mentali<sup>10</sup>.

Raccogliendo l’invito di Ventura, ci si chiede: che cosa s’intende per ‘processi cognitivi’ nell’ambito dell’analisi concettuale dei documenti? Quale valenza si è voluto attribuire ad essi, quando sono stati spiegati ed esemplificati nella formazione e aggiornamento professionale? Quali implicazioni operative ne sono state tratte? Dalle regole di sommarizzazione ne consegue una ‘rappresentazione del testo’ equivalente a una ‘rappresentazione contratta della conoscenza’?

<sup>7</sup> Piero Cavalieri, *La biblioteca crea significato: thesaurus, termini e concetti*. Milano: Bibliografica, 2013.

<sup>8</sup> *Ivi*, p. 77-78.

<sup>9</sup> Roberto Ventura, *La biblioteca tra linguaggio e rappresentazione: per un’indagine filosofica della biblioteconomia*, «JLIS.it», 1 (2010), n. 1, p. 41-61.

<sup>10</sup> *Ivi*, p. 52-53.

Dunque, il compito affidato a queste note non è quello di fornire una descrizione sistematica e compiuta delle operazioni di analisi concettuale dei documenti, bensì di tentare una sorta di negoziazione del significato di ‘processi cognitivi’, che consenta di coglierne, in quest’ambito, la valenza e l’efficacia, come pure di verificarne la coerenza con un certo modello di analisi e indicizzazione per soggetto, quale quello che informa la Guida GRIS e il *Nuovo soggettario*.

I processi cognitivi sono identificati qui in alcuni processi inferenziali, che, con una certa approssimazione e senza alcuna pretesa di esaustività, si possono suddividere nelle seguenti tipologie, che nella pratica tendono a sovrapporsi:

1. processi semantici e sintattici, basati sull’interazione testo-contesto e finalizzati all’interpretazione delle parole e delle loro relazioni;
2. processi di sommarizzazione, basati sulle relazioni logiche dei temi all’interno del testo e finalizzati all’individuazione della loro rilevanza nella struttura del testo;
3. processi di tipo valutativo, basati sulla relazione documento-bisogni informativi e finalizzati alla valutazione della pertinenza dei temi individuati, ai fini del recupero.

I chiarimenti forniti qui attingono prevalentemente a materiali didattici: esempi, schemi, esercizi. Se all’inevitabile fallacia degli esempi si aggiunge il carattere non sistematico, interlocutorio di queste note e l’assenza di un apparato di riferimenti bibliografici adeguato alla natura dei temi trattati, ben s’addice ad essi la qualificazione di ‘postilla’. Ma il motivo principale sta nel modo in cui i processi sono descritti, diverso da quello che ci si aspetterebbe da uno scritto di biblioteconomia, che dovrebbe dar conto degli scopi, dei principi, delle norme, delle procedure che presiedono a una determinata operazione; mentre qui lo schema descrittivo è prevalentemente quello inferenziale (se, allora) e il tipo di inferenza consiste in un ragionamento nel quale conoscenze testuali e conoscenze contestuali interagiscono per approdare alla fine a una conoscenza metatestuale (tema, *topic*, *aboutness*), quale «ipotesi sulla validità che certe voci dell’indice, o parimenti alcuni soggetti, possano avere per diventare occasioni di ritrovamento del documento, e quindi del suo contenuto semantico»<sup>11</sup>. I ragionamenti che seguono possono contribuire a gettare qualche luce sulle condizioni di questa ‘validità’.

### Un modello interattivo

Se il tipo di inferenza dipende dalla «natura dell’oggetto di partenza»<sup>12</sup>, allora i processi inferenziali coinvolti nell’analisi concettuale dei documenti si possono considerare, in termini generali, come processi di tipo abduttivo. Infatti, l’oggetto di partenza è sempre un ‘testo’ (il documento che si sta analizzando e, al suo interno, un’occorrenza lessicale, un enunciato, una serie di enunciati). Il punto di arrivo è la traduzione di un’espressione (l’espressione del ‘soggetto’ del documento) in un’altra espressione (la rappresentazione del ‘soggetto’ nel linguaggio di indicizzazione): si tratta, dunque, di un processo di interpretazione. Il percorso dall’oggetto di partenza (il testo) a quello di arrivo (la stringa di soggetto) si sviluppa attraverso una serie di passaggi intermedi, che si avvalgono della mediazione di conoscenze

<sup>11</sup> Alfredo Serrai, *Indici, logica e linguaggio: problemi di catalogazione semantica*. Roma: Consiglio nazionale delle ricerche. Laboratorio di studi sulla ricerca e sulla documentazione, 1974, p. 52.

<sup>12</sup> Salvatore Zingale, *Il ciclo inferenziale: deduzione, induzione, abduzione: dall’interpretazione alla progettualità: working paper, 2009*, <<http://europa.uniroma3.it/cipriani/files/59a778e0-31a6-4f8b-ad2d-fcb7aacof21.pdf>>, p. 1 (ultima consultazione: 01/10/2015).

linguistiche, extralinguistiche, intertestuali, rappresentazioni schematiche ecc., selezionate dal contesto. Dell’abduzione il processo di analisi condivide anche la natura ipotetica, probabilistica dell’oggetto di arrivo: la pertinenza del soggetto rispetto ai bisogni conoscitivi degli utenti, la sua rappresentazione come mezzo efficace di reperimento del documento. L’abduzione interviene quando si debbano interpretare sintomi, indizi, parole. «Ma diversamente non accade quando voglia interpretare il valore che un dato enunciato, una parola chiave, una intera vicenda, assumono in un testo»<sup>13</sup>.

Da questo punto di vista, l’analisi concettuale dei documenti può essere descritta secondo un modello interattivo, schematicamente rappresentato nella Fig 1. Nella colonna di sinistra sono descritte sommariamente le operazioni di analisi e di indicizzazione del documento; in quella di destra compaiono i principali elementi del contesto, ossia l’insieme delle conoscenze possedute; in quella centrale sono indicati schematicamente i processi di elaborazione, che consistono nell’attivazione di conoscenze dal contesto in funzione dell’analisi e dell’indicizzazione del documento.

Analisi del documento	↔	Processi di elaborazione	↔	Contesto
<ul style="list-style-type: none"> <li>➤ Esame del documento (testo e paratesto)</li> <li>➤ Individuazione dei temi e dei loro nessi logici (congiunzione / implicazione)</li> <li>➤ Individuazione del tema di base/temi particolari e selezione dei concetti</li> <li>➤ Rappresentazione del/dei soggetto/i in uno o più enunciati</li> <li>➤ Traduzione del/dei soggetto/i nel linguaggio di indicizzazione</li> </ul>		<ul style="list-style-type: none"> <li>➤ Attivazione di conoscenze linguistiche, extralinguistiche e intertestuali</li> <li>➤ Attivazione di strutture semantiche, schemi, scene</li> <li>➤ Applicazione di strategie: tecniche di esame del documento, macro-operazioni ecc.</li> </ul>		<ul style="list-style-type: none"> <li>➤ Conoscenze linguistiche: linguaggio naturale e linguaggio di indicizzazione</li> <li>➤ Conoscenze extralinguistiche:                             <ul style="list-style-type: none"> <li>• conoscenza del mondo</li> <li>• conoscenza del contesto documentario</li> <li>• conoscenza dei bisogni informativi degli utenti</li> </ul> </li> <li>➤ Conoscenze intertestuali</li> <li>➤ Strategie di analisi: tecniche di esame del documento, macro-operazioni ecc.</li> </ul>

Figura 1 – Modello interattivo di analisi concettuale dei documenti

13 Umberto Eco, *Semiotica e filosofia del linguaggio*. Torino: Einaudi, 1984, p. 42.

Dunque, ciascun passaggio intermedio del percorso di analisi concettuale dei documenti implica due momenti, che interagiscono tra loro:

- un processo dal basso (*bottom-up*) o percettivo;
- un processo dall'alto (*top-down*) o cognitivo.

Il processo percettivo consiste nell'esame dell'informazione contenuta nel documento, mentre quello cognitivo, con cui interagisce, consiste nell'attivazione di vari tipi di conoscenze possedute, che svolgono una funzione euristica nell'interpretazione delle parole e degli enunciati, nella selezione dei concetti e nella scelta dei termini di indicizzazione. Tra queste vi sono le conoscenze relative a concetti singoli e modelli concettuali (strutture semantiche e schemi) impiegati nel linguaggio di indicizzazione.

Un modello interattivo, con una componente *bottom-up* (induttiva) e una *top-down* (deduttiva) che si combinano, è generalmente considerato dai linguisti il più appropriato a tutti i livelli di analisi (delle parole, delle proposizioni e dei discorsi), poiché «at all levels, textual information and contextual information are closely intertwined»<sup>14</sup>.

Scrivete Beghtol:

To analyse the cognitive processes of both generating and comprehending actual language systems (i.e. texts or discourses) text linguists borrow from cognitive psychology the concepts of two kinds of mental information processing: 'top-down', concept-driven or deductive processing; and 'bottom-up', data-driven or inductive processing. Empirical research shows that these two kinds of mental processing appear to occur continuously during a reader's comprehension of a text<sup>15</sup>.

L'esempio seguente (Fig. 2), relativo all'opera di Rita Valentino Merletti, *Leggere ad alta voce*, mostra un'applicazione parziale di questo modello, nella quale le conoscenze selezionate si riferiscono soprattutto a informazioni registrate nel linguaggio di indicizzazione, che, come si è visto nello schema precedente, rappresentano solo una porzione delle conoscenze possedute. Man mano che si procede nell'esame del documento, si attivano conoscenze che il testo stesso richiama: ora sono le conoscenze linguistiche ed extralinguistiche sulla lettura ad alta voce, ora le informazioni che su di essa sono registrate nel linguaggio di indicizzazione, ora i modelli concettuali ivi incorporati che ad essa si possono applicare. È specificato nel documento un agente, un oggetto, un beneficiario della lettura ad alta voce? Ne sono trattati aspetti particolari? Il documento è destinato a una determinata classe di persone? E così via.

<sup>14</sup> Teun A. van Dijk; Walter Kintsch, *Strategies of discourse comprehension*. New York: Academic press, 1983, p. 221.

<sup>15</sup> Clare Beghtol, *Bibliographic classification theory and text linguistics: aboutness analysis, intertextuality and the cognitive act of classifying documents*, «Journal of documentation», 42 (1986), n. 2, p. 84-113: p. 88.

(Rita Valentino Merletti, *Leggere ad alta voce*. Milano: Mondadori, 1996)

Esame		Conoscenze possedute
<p>Un libro sulla lettura ad alta voce? Un intero libro? Cosa mai si potrà scrivere dopo la seconda pagina? [...] Queste domande mi sono state rivolte ogni volta che cercavo di illustrare il mio progetto. E non si può negare che gli interrogativi abbiano una loro giustificazione. La lettura ad alta voce è uno di quegli argomenti su cui ci sentiamo tutti competenti. (p. 9)</p>	<p><i>bottom up</i> →</p> <p><i>top down</i> ←</p>	<p>Competenze linguistiche ed extralinguistiche sulla lettura ad alta voce.</p> <p>Informazioni sulla lettura ad alta voce registrate nel sistema di indicizzazione:</p> <ul style="list-style-type: none"> <li>voce del thesaurus (il termine e il suo corredo di relazioni): elementi descrittivi, relazioni semantiche, tipi di concetti, rinvio a dizionari ed enciclopedie ecc.;</li> <li>stringhe di soggetto contenenti il termine.</li> </ul> <p>Modelli concettuali propri del linguaggio di indicizzazione, per esempio lo schema dei ruoli, che:</p> <ol style="list-style-type: none"> <li>attiva la seguente cornice della lettura ad alta voce:</li> </ol>  <pre> graph LR     A([oggetto]) --- B[Letture ad alta voce]     B --- C([agente])     B --- D([beneficiario])     </pre> <ol style="list-style-type: none"> <li>fornisce alcune informazioni procedurali: SE in un documento è presente un concetto di attività transitiva a tre argomenti, ALLORA verifica se sono presenti anche l'oggetto, l'agente, il beneficiario.</li> </ol>
<p>Questo testo non intende essere un manuale tecnico della lettura ad alta voce, né intende rivolgersi a chi ad essa si dedica a livello professionale. Scuole, biblioteche, laboratori di lettura, librerie specializzate sono impegnate sul fronte della diffusione della lettura [...]. Si è dunque sviluppata, tra gli addetti ai lavori, una grammatica di riferimento la cui discussione non rientra negli scopi di questo libro. (p. 11)</p>	<p><i>bottom up</i> →</p> <p><i>top down</i> ←</p>	<p><b>Manuali</b> Tipi di concetti: Forme. Informazioni specifiche: testi che avviano allo studio di una disciplina, di una tecnica, di un'arte, un'attività ecc.; sono utilizzati per scopi didattici o divulgativi. Relazioni: BT [Documenti secondo la presentazione di contenuti e testi, la forma bibliografica]. Ruoli: Forma. Ruolo extranucleare o complementare. I concetti formali indicano la forma (fisica o intellettuale) di un documento. Informazioni procedurali: è opportuno assegnare un'indicazione di forma, SE può modificare la valutazione da parte dell'utente della rilevanza del documento.</p> <p><b>Destinazione</b> Ruoli: Destinazione. È un ruolo extranucleare o complementare. La destinazione indica la classe di persone a cui il documento è rivolto.</p>
<p>Il modo di leggere ad alta voce che proponiamo in queste pagine ha a che fare con una dimensione più intima e familiare della lettura [...]. Non un manuale tecnico, ma una riflessione su quanto della lettura ci sembra importante trasmettere ai nostri bambini. [...]. Basterebbe questo dunque a spiegare perché è importante leggere ad alta voce ai nostri bambini. (p. 11)</p>	<p><i>bottom up</i> →</p> <p><i>top down</i> ←</p>	<p><b>Aspetti</b> Ruoli: Aspetti. E' un ruolo extranucleare o complementare, da assegnare a termini che precisano il tipo di approccio intellettuale al tema, per esempio, delimitandone gli aspetti trattati nel documento. Informazioni procedurali: è opportuno assegnare un'indicazione di forma intellettuale, solo SE il tema è studiato al di fuori del suo contesto proprio o tipico.</p> <p><b>Beneficiario</b> Ruoli: beneficiario. È un ruolo nucleare. È il ruolo da assegnare al termine che rappresenta il beneficiario dell'azione. Da non confondere con la destinazione del documento.</p>

Figura 2 – Esempio di attivazione di processi cognitivi

Il modo dell'inferenza può essere esemplificato in questi termini: se un documento parla della lettura ad alta voce, poiché un'azione transitiva a tre argomenti, come la lettura ad alta voce, può stabilire la relazione con l'oggetto, l'agente e il beneficiario, allora è probabile che nel documento siano specificate una o più di queste funzioni come figure chiave nell'espressione del tema del documento.

L'indicizzatore elabora questa inferenza a partire da due premesse: la prima è il dato ricavato dal documento; la seconda, invece, va a cercarla nel contesto, in questo caso nello schema dei ruoli nella sua funzione euristica. Questo esempio può essere ricondotto a un tipo di inferenza definibile come 'inferenza da schemi'. Tuttavia, un analogo modo inferenziale – interazione tra testo e contesto – caratterizza anche le inferenze lessicali e i processi di sommarizzazione. Questi tre tipi di inferenze sono trattati nei paragrafi successivi.

Chiamare questo 'andirivieni' tra testo e contesto una «passeggiata inferenziale»<sup>16</sup> può apparire qui una metafora troppo disinvolta; tuttavia, ne mantiene la sceneggiatura generale, che prevede la fuoriuscita dal testo per rientrarvi con le conoscenze, selezionate dal contesto, utili ai fini dell'analisi.

Se la prima delle due premesse è un dato ricavato dal documento, allora essa presuppone una qualche tecnica di esame del documento. Volendo usare l'immagine della lettura come un «percorso a sorvolo»<sup>17</sup>, si potrebbe dire che la premessa si forma quando l'esame del testo trova un punto di appoggio, un aggancio, un indizio che ci faccia capire di cosa si tratta. Illustrate nei corsi di formazione e aggiornamento, di queste tecniche (di cui fanno parte l'indagine sulle fonti, per esempio, la funzione e le caratteristiche dei titoli, i tipi di lettura ecc.) non si parlerà in queste note.

Il contesto è inteso qui come l'insieme delle conoscenze possedute o, in analogia con la metafora elaborata da Umberto Eco, come l'enciclopedia disponibile<sup>18</sup>. Tuttavia, nel paragrafo successivo le espressioni 'contesto' e 'contestuale' sono utilizzate anche per indicare un co-testo, ossia l'enunciato o gli enunciati in cui compare o è applicata un'occorrenza lessicale.

### Inferenze lessicali

Se, come scrive Valentino Merletti, «la lettura ad alta voce è uno di quegli argomenti su cui ci sentiamo tutti competenti», allora si può supporre che la sua interpretazione non presenti problemi. Supponiamo, invece, che in un documento ricorra l'espressione 'gatti di biblioteca'. Qual è il significato di quest'occorrenza? Le pre-conoscenze linguistiche ed extralinguistiche ci inducono a formulare ipotesi, di cui solo l'esame del documento può verificare la pertinenza. Sono, quelli di cui si parla nel documento, i gatti che vivono nelle biblioteche, dove, adottati da solerti bibliotecari e circondati dall'affetto dei lettori, svolgono molteplici funzioni, da acchiappatopi a beniamini dei frequentatori, mascotte ecc.? Supponiamo che sia così e che, sul calco di «Library cats» (da Library of Congress subject headings – LCSH), si adotti il termine di indicizzazione 'Gatti di biblioteca'. Possiamo, tuttavia, associare lo stesso contenuto concettuale a un'altra occorrenza della medesima espressione, quella che ricorre, per esempio, nel titolo *Gatti di biblioteca* del volume di Michèle Sacquin (Milano: Officina libraria, 2010)? Ragionevolmente no, trattandosi in quest'ultimo caso

16 Umberto Eco, *Lector in fabula: la cooperazione interpretativa nei testi narrativi*. Milano: Bompiani, 1983, p. 117-119 (1. ed. 1979).

17 Paul Cornea, *Introduzione alla teoria della lettura*, a cura di Gheorghe Carageani. Firenze: Sansoni, 1993, p. 156.

18 U. Eco, *Semiotica e filosofia del linguaggio* cit, p. 106 e seguenti.

non di gatti ‘in carne e ossa’, ma di gatti ‘di carta’, ossia di gatti rappresentati in poesie, racconti, proverbi, aneddoti, immagini ecc., tratti dalla collezione di manoscritti e stampe della Bibliothèque nationale de France. Anche questi sono gatti e, come i primi, si aggirano tra scaffali e libri, ma abitano nelle biblioteche in un altro senso, al quale presumibilmente associare una diversa forma espressiva nel linguaggio di indicizzazione. E che dire del *Felis bibliothecarum* di Luigi Crocetti<sup>19</sup>?

Dovendo indicizzare un documento su ‘esercizio provvisorio dei bilanci pubblici’, l’indicizzatore verifica nel vocabolario del sistema di indicizzazione la presenza del termine ‘esercizio provvisorio’:

(dal *Thesaurus* del *Nuovo soggettario*)

**Esercizio provvisorio**

Categoria/Faccetta: Azioni: Attività

Nota d’ambito: La continuazione temporanea dell’esercizio dell’impresa del fallito, che il tribunale può disporre dopo la dichiarazione di fallimento, quando dall’interruzione improvvisa possa derivare un danno grave e irreparabile.

TT                    Attività

BT                    [Attività economico-gestionali]

RT                    Fallimento

Falliti

(stringhe di soggetto)

**Imprese – Esercizio provvisorio**

Il significato di ‘esercizio provvisorio’ riferito all’impresa fallita è compatibile con l’associazione della medesima espressione all’esercizio finanziario degli enti pubblici nel periodo che intercorre tra la scadenza dell’anno finanziario e l’approvazione del nuovo bilancio? Prevalgono i tratti comuni (la continuazione di un’attività, la temporaneità, la prevenzione del danno), oppure le differenze (impresa vs. ente pubblico, diritto privato vs. diritto pubblico ecc.)? Il significato lessicale, così com’è registrato nel *thesaurus*, consente entrambe le contestualizzazioni? È l’istruzione da modificare, per ampliare i possibili contesti d’uso del termine, oppure, confermandone le restrizioni e i vincoli, si preferisce adottare un nuovo termine?

Analogamente, come rappresentare nel linguaggio di indicizzazione l’espressione ‘imposta di capitazione’, quando è usata nei documenti per riferirsi ora alla moderna Poll tax, ora all’imposizione fiscale per testa di epoca medievale (testatico), ora al *tributum pro capite* di epoca romana (*capitatio*)?

Per l’espressione ‘tessere’ si possono prevedere almeno due contesti: il contesto dell’arte e della tecnica del mosaico e il contesto del tesseramento presso enti, partiti, organizzazioni: evidentemente, quest’espressione significa una cosa diversa in ciascuno dei due contesti.

L’espressione ‘sviluppo’, registrata come processo di applicabilità generale, si può riferire a innumerevoli fenomeni (economia, memoria, personalità, corpo umano, fotografie ecc.): in tutte queste occorrenze, è sufficiente l’esplicitazione del contesto a precisarne il significato? E così via.

Queste domande sottendono una specie di negoziazione del significato, che ha lo scopo di tradurre un’espressione (la sua occorrenza nel documento) in un’altra espressione (il termine di indicizzazione). In questo senso, il processo di ‘negoziato’ è un processo di ‘interpretazione’. In caso di conflitto, tale da richiedere un’eventuale

<sup>19</sup> Luigi Crocetti, *Lat. scient. Felis bibliothecarum*, «Biblioteche oggi», 17 (1999), n. 6, p. 79.

rinegoziazione del significato, si possono dare due opzioni: mantenere il significato lessicale del termine così com'è fissato nel *thesaurus*, creando un nuovo termine di indicizzazione per l'occorrenza dell'espressione nel documento in esame; oppure, modificare il significato lessicale del termine di indicizzazione.

Di queste negoziazioni rimane traccia nel *thesaurus* (nelle note d'ambito, di orientamento, nelle note storiche), quando ne è derivato un cambiamento del significato lessicale del termine. Esempi (dal *Thesaurus del Nuovo soggettario*):

#### **Anagrafe commerciale**

Categoria/Faccetta: Agenti: Organizzazioni

Nota d'ambito: Uffici, istituiti presso le Camere di commercio, che si occupano del servizio di registro delle imprese e di tutte le attività economiche.

TT Organizzazioni  
BT Camere di commercio  
RT Registro delle imprese

Nota storica: Precedentemente anche con il significato di Registro delle imprese (BNI 1999-2005).

#### **Calzolerie**

Categoria/Faccetta: Agenti: Organizzazioni

TT Organizzazioni  
BT Aziende artigiane  
RT Calzature  
Calzaturifici  
Calzolai  
Negozio di calzature

Nota d'ambito: Laboratori artigiani in cui si producono e riparano calzature. Per il significato di negozi usare Negozi di calzature. Per il significato di aziende industriali di calzature usare Calzaturifici.

Nota storica: Precedentemente anche con il significato di Negozi di calzature (Soggettario).

#### **Locomozione**

Categoria/Faccetta: Azioni: Attività

Nota d'ambito: Attività di spostamento di organismi viventi (animali e umani) da un luogo a un altro, dovuto all'attivazione del loro apparato locomotore (la deambulazione dell'uomo e di alcuni vertebrati superiori, il volo degli uccelli, la reptazione dei rettili, il nuoto dei pesci). Per il movimento degli organismi in generale usare Movimenti. Per la locomozione nell'ambito della meccanica pura e applicata usare Moto.

TT Attività  
BT [Attività relative agli spostamenti e ai trasporti]  
RT Animali  
Apparato locomotore  
Moto  
Movimenti  
Psicomotricità  
Spostamento  
Uomo

Nota storica: Precedentemente con il significato di Moto.

**Ruoli**

Categoria/Faccetta: Cose: Strumenti

Nota d'ambito: Nelle amministrazioni pubbliche, private o in altri enti, la composizione e l'ordinamento del personale. Per il significato di prospetti in cui sono registrati nomi di persone che fanno parte di enti usare Elenchi. Per il significato degli elenchi dei contribuenti, a fianco del cui nome compare il credito accertato o dichiarato e l'ammontare di imposta che ciascuno di loro è tenuto a pagare, usare Ruoli dei contribuenti.

TT Strumenti

BT [Strumenti del mondo del lavoro]

RT Personale

Nota storica: Precedentemente anche con il significato di Ruoli dei contribuenti.

Queste voci registrano sia l'analisi diacronica del significato (il mutamento), sia quella sincronica (la stabilità), in un processo in cui ogni occorrenza di un'espressione in un nuovo documento mette in questione potenzialmente il significato di quell'espressione registrato nel sistema di indicizzazione; al tempo stesso, il significato di un'espressione registrato nel sistema di indicizzazione fornisce le istruzioni che ne regolano l'applicazione in contesti determinati. Dunque, la negoziazione del significato è da intendersi come una comparazione dinamica tra 'significato lessicale' e 'significato contestuale', ossia tra il significato dell'espressione come convenzionalmente registrato nel *thesaurus* e il suo significato in un co-testo determinato, che, a seconda dello stadio del processo di indicizzazione in cui ci troviamo, è il testo del documento che si sta analizzando o l'enunciato che ne esprime il *topic* oppure la stringa di soggetto che lo rappresenta nel catalogo.

Se questa comparazione, che si verifica nel corso dell'analisi concettuale del documento, può comportare un cambiamento nel significato di un termine, al tempo stesso il significato stabilizzato del termine di indicizzazione consente di fare inferenze valide e cogenti, basate su regole condivise e convenzioni. Il modo dell'inferenza può essere schematizzato così: se  $x$  (termine di indicizzazione) implica  $y$  (significato lessicale di  $x$ ), allora  $x$  può/non può essere applicato nel contesto  $p$  (enunciato nel quale un'occorrenza di  $x$  implica/non implica  $y$ ). Se 'Gatti di biblioteca' significa 'gatti che abitano nelle biblioteche', allora non si può impiegare questo termine nell'indicizzazione di un documento il cui *topic* è 'gatti come tema letterario e artistico'. Se si presuppone convenzionalmente che 'Tessere' siano oggetti costituiti da piccoli pezzi di marmo, vetro ecc., allora il termine verrebbe usato impropriamente per indicizzare un documento dove ricorra l'espressione 'tessere' nell'ambito del 'tesseramento' a un partito o a un'associazione. Se 'Anagrafe commerciale' è un'organizzazione, allora non può essere impiegata per significare un documento come il 'registro delle imprese'. Se 'Gatti' sono 'Felini', nulla impedisce che possano essere considerati, nei diversi contesti, animali sacri, da pelliccia, di laboratorio, da compagnia, animali con determinati fabbisogni nutritivi, che si ammalano, che possono causare malattie nell'uomo o in altri animali, gatti di strada, gatti di biblioteca, gatti di bordo, acchiappatopi, temi letterari e artistici, persino sapienti lettori, poiché in nessuno di questi contesti è messo in crisi il rapporto di implicazione tra gatti e felini.

Dunque, il significato lessicale registrato nel *thesaurus* può essere interpretato anche come una «istruzione per l'inserzione contestuale»<sup>20</sup> o come un sistema di «istruzioni

<sup>20</sup> U. Eco, *Semiotica e filosofia del linguaggio* cit., p. 72.

orientate al testo»<sup>21</sup>. Tutti gli elementi che compongono le voci del *thesaurus* (categoria, relazioni semantiche, note, link ecc.) concorrono a stabilire se un termine di indicizzazione possa/non possa essere inserito in un determinato contesto (enunciato/stringa di soggetto): per esempio, se 'Locomozione' si possa impiegare in un contesto zoologico oppure meccanico o in entrambi. Al tempo stesso, l'occorrenza di un'espressione in un contesto (il testo del documento) seleziona un senso che trova/non trova corrispondenza in un determinato termine d'indicizzazione, in virtù del contenuto di istruzioni di quest'ultimo. Così, l'occorrenza 'locomozione' in un testo di meccanica della locomozione non trova corrispondenza nel termine di indicizzazione 'Locomozione', poiché l'istruzione contenuta nel suo significato lessicale ne restringe l'ambito all'attività di spostamento di organismi viventi (animali e umani) da un luogo a un altro, dovuto all'attivazione del loro apparato locomotore; trova corrispondenza, invece, nel termine 'Moto', assunto a significare il movimento nell'ambito della fisica.

Se in relazione ai possibili contesti il significato di un termine può essere interpretato come istruzione, in relazione agli usi contestuali concretamente realizzati può essere interpretato come un «potenziale semantico», che si attualizza nell'enunciazione<sup>22</sup>; nel nostro caso, nelle tre situazioni di enunciazione: il testo del documento, l'enunciato che ne esprime il *topic*, la stringa di soggetto. Il significato del termine registrato nel *thesaurus* fa da ponte tra l'una e l'altra. Come si manifesta il potenziale semantico nel sistema di indicizzazione? Principalmente, nel legame tra il termine e le stringhe di soggetto in cui compare, ciascuna delle quali ne prende in carico il 'potenziale semantico' e lo attualizza in una particolare configurazione di contenuto. Nella valenza di 'istruzione', è associato a un termine un insieme di condizioni di applicazione, in quella di 'potenziale semantico' un insieme di applicazioni in contesti reali, che nel linguaggio di indicizzazione è assimilabile all'elenco delle stringhe di soggetto collegate al termine.

Contesti possibili e contesti reali sono, poi, riunificabili nella nozione di '*contextual relations*': il significato di una parola si riflette pienamente nelle sue relazioni contestuali; o anche: il significato di una parola è costituito dalle sue relazioni contestuali, intendendo sia le relazioni paradigmatiche che quelle sintagmatiche<sup>23</sup>.

Queste valenze del significato sottendono approcci almeno in parte differenti, oggetto di discussione in campo linguistico: per esempio, la nozione di 'potenziale semantico' delle parole oscilla da una posizione che lo concepisce come un insieme di sensi possibili, come un contenuto semantico immagazzinato in concetti lessicali, a cui le parole sono associate attraverso processi di astrazione e a cui sono vincolati i sensi determinati che le parole acquisiscono nei diversi contesti d'uso, a una posizione in cui il potenziale semantico è definito come «the collection of past uses of a word w on the basis of which similarities can be established between source situations (i.e., the circumstances in which a speaker has used w) and target situations (i.e., candidate occasions of application of w)»<sup>24</sup>, senza bisogno di presupporre alcun significato convenzionale delle parole.

21 U. Eco, *Lector in fabula* cit., p. 15.

22 Patrizia Violi, *Significati lessicali e pratiche comunicative: una prospettiva semiotica*, «Rivista di Linguistica», 15 (2003), n. 2, p. 321-342: p. 331.

23 Alan D. Cruse, *Lexical semantics*. Cambridge: Cambridge university press, 1986, p. 15-20.

24 Per una rassegna delle diverse visioni del significato lessicale cfr. Luca Gasparri; Diego Marconi, *Word meaning*. In: *The Stanford encyclopedia of philosophy*, edited by Edward N. Zalta, 2015, <<http://plato.stanford.edu/archives/sum2015/entries/word-meaning/>> (ultima consultazione: 01/10/2015), da cui è tratta la citazione.

Qui, tuttavia, queste valenze sono intese come aspetti complementari del complesso rapporto delle parole con il contesto nelle sue varie forme, in quanto aspetti verificabili nell'ambito dell'indicizzazione per soggetto: al significato di un termine come istruzione corrisponde l'insieme degli elementi prescrittivi insiti nella relativa voce del *thesaurus*, al significato come potenziale semantico l'insieme dei sensi che il termine assume nelle stringhe di soggetto di cui è componente semantico, al significato che si riflette nelle relazioni contestuali l'insieme di relazioni (sinonimica, generica, associativa, attributiva, agentiva, oggettiva ecc.) che il termine contrae con altri termini nelle voci del *thesaurus* e nelle stringhe di soggetto, ognuno dei quali partecipa in diverso modo all'espressione del suo significato.

Resta un'ultima valenza da analizzare: il significato lessicale come insieme di proprietà. Prima torniamo per un momento ai gatti, da cui è iniziata questa discussione. Se l'esempio della 'locomozione' mostra che il significato lessicale deve valere (essere implicato) in tutti i contesti, l'esempio dei 'gatti' mostra l'altra faccia della medaglia, ossia che il «significato contestuale va molto al di là dei significati lessicali»<sup>25</sup>.

Che un gatto sia non solo un felino domestico, ma anche l'animale che le classificazioni zoologiche definiscono come *felis catus*, l'animale adorato dagli Egiziani, l'animale che appare nell'*Olympia* di Manet, l'animale mangiare il quale era una leccornia nella Parigi assediata dai prussiani, l'animale cantato da Baudelaire, l'animale che Collodi associa per astuzia e malvagità alla volpe, l'animale che in una certa favola è al servizio del marchese di Carabas, un infingardo amante della casa che non muore di inedia sulla tomba del padrone, l'animale prediletto delle streghe e così via, sono tutte interpretazioni dell'espressione 'gatto'. Tutte sono registrate, poste intersoggettivamente in qualche testo di quella immensa e ideale biblioteca il cui modello teorico è l'enciclopedia. Ciascuna di queste interpretazioni definisce sotto qualche aspetto cosa sia un gatto, e tuttavia fa conoscere sempre qualcosa di più circa un gatto. Ciascuna di queste interpretazioni vale ed è atualizzabile in un determinato contesto, ma l'enciclopedia dovrebbe idealmente provvedere istruzioni onde interpretare nel modo più fruttuoso l'espressione 'gatto' in numerosi contesti possibili<sup>26</sup>.

La natura enciclopedica del significato, così com'è rappresentata nel linguaggio di indicizzazione, veniva illustrata in *Un thesaurus per le biblioteche generali* mediante l'esempio dei 'pesticidi', su cui si costruivano diversi tipi di enunciati, afferenti a molteplici ambiti disciplinari e aree di attività, mostrando come alcuni di essi si strutturino in forma di dizionario, altri in forma di relazioni sintattiche con altri termini, secondo certe strategie di ottimizzazione della divisione di compiti e dell'interazione tra le due dimensioni del linguaggio di indicizzazione (il vocabolario e la sintassi), entro le quali prende corpo il significato di un termine. E si annotava:

Del resto, i linguisti e gli studiosi della semantica sono propensi a considerare il significato non come un'entità statica e isolata, ma come un fenomeno dinamico e relazionale: il significato di un termine, si dice, è sempre in relazione ai suoi contesti d'uso. C'è, dunque, un significato di base delle parole, al quale sono associati in modo relativamente stabile alcuni contenuti concettuali e per il quale un concetto

<sup>25</sup> U. Eco, *Semiotica e filosofia del linguaggio* cit., p. 72.

<sup>26</sup> *Ivi*, p. 108-109.

è inserito in un sistema strutturato; c'è, però, anche un aspetto mobile, flessibile del significato, che si manifesta quando un termine, associato a determinati contesti, riceve particolari valenze, che arricchiscono ed estendono il suo significato di base, o assume addirittura significati diversi, che da esso si irradiano<sup>27</sup>.

Il contenuto concettuale associato a un termine è un'altra forma di interpretazione del significato lessicale. Se nella sua valenza di 'istruzione', è associato a un termine un insieme di condizioni applicative, in quella di 'potenziale semantico' un insieme di applicazioni reali e in quella di 'relazioni contestuali' l'insieme di tutti i contesti concepibili con cui si accorda, nella valenza di 'contenuto concettuale' è associato al termine un insieme di proprietà. È questa rappresentazione che conferisce stabilità al significato, impone vincoli ai processi interpretativi, dà luogo a inferenze in forma di implicazione (se gatto allora felino vs. se gatto allora probabilmente (in qualche contesto) animale sacro o da pelliccia o di laboratorio o tema letterario ecc.) e limita la rinegoziazione solo ai casi in cui è messo in questione quest'insieme di proprietà. Senza questo elemento di stabilità non sarebbe possibile neppure la negoziazione.

Quali tipi di proprietà e in che modo si selezionano? Sono stati elaborati vari modelli per stabilire il tipo, il numero e il peso delle proprietà nella definizione del significato di un termine: la teoria classica, i modelli 'a tratti differenziali', i modelli basati sul 'prototipo' ecc. Non si entra nel merito di questi diversi approcci. Nella Guida GRIS e nella Guida al *Nuovo soggettario*<sup>28</sup>, queste proprietà sono caratterizzate nei seguenti modi:

- si tratta di attributi essenziali, che esprimono il significato di base di un termine;
- seppure formalizzato, questo significato è sempre il più vicino possibile al significato tipico del termine nel linguaggio naturale, a garanzia dell'espressività del termine di indicizzazione;
- questo significato, inoltre, è costituito spesso da un fascio di significati sovrapposti o parzialmente coincidenti, seppure tra loro coerenti;
- gli attributi che esprimono il significato di base hanno prevalentemente la funzione di stabilire l'appartenenza categoriale di un termine, riferendolo a una categoria indipendente dal contesto, ossia valida in ciascuno dei differenti ambiti nei quali può comparire;
- sia il significato di base che l'appartenenza categoriale sono generalmente riconosciuti e condivisi e possono essere stabiliti facendo riferimento a opere standard come dizionari o enciclopedie.

Come esplicitamente richiamato nei due strumenti normativi sopra citati, il nucleo essenziale di queste affermazioni deriva dal 'principio della definizione unica'. Non potendo approfondirne qui il significato e le implicazioni, mi limito a coglierne sinteticamente tre aspetti, quali risultano dalle spiegazioni date dal suo ideatore, Jason Farradane, messe in luce da Elisabetta Viti nella sua tesi di dottorato<sup>29</sup>, allo

<sup>27</sup> Alberto Cheti, *Un thesaurus per le biblioteche generali*, «Bollettino AIB», 32 (1992), n. 1, p. 29-41: p. 33.

<sup>28</sup> AIB. GRIS, *Guida all'indicizzazione per soggetto* cit., p. 60-63. Cfr. anche Biblioteca nazionale centrale di Firenze, *Nuovo soggettario: guida al sistema italiano di indicizzazione per soggetto. Prototipo del thesaurus*. Milano: Bibliografica, 2006, p. 51-52.

<sup>29</sup> Elisabetta Viti, *Interoperabilità fra thesauri generali e thesauri specialistici in ambito economico-finanziario: il caso del Nuovo soggettario* [tesi di dottorato di ricerca]. Udine: Università degli studi di Udine, 2012, p. 52-62. Cfr. anche Claudio Gnoli, *Progress in synthetic classification: towards unique definition of concepts*, «Extensions & corrections to the UDC», 29 (2007), p. 167-182.

scopo di trovare alcune analogie con quanto elaborato in ambito linguistico: il contenuto (quali caratteristiche del concetto include la definizione unica), lo scopo (perché c'è bisogno di una definizione unica), il modo in cui si selezionano le proprietà in essa contenute. È – spiega Farradane – una definizione libera da associazioni non necessarie («it must be free from unnecessary associations»<sup>30</sup>), ossia è una definizione che contiene soltanto le proprietà essenziali del concetto. Il suo scopo è dare a un concetto una collocazione unica (un posto unico) in una struttura classificatoria: «one term, one place»<sup>31</sup>. Questa collocazione è stabilita mediante la cancellazione delle proprietà non necessarie: «the fundamental class placing of an item [...] should be decided by stripping off extraneous characteristics to leave common (intrinsic, essential) characteristics by which the item was uniquely definable»<sup>32</sup>.

Che il significato lessicale sia un insieme composito e strutturato di proprietà è ampiamente riconosciuto in ambito linguistico: si distinguono, per esempio, 'attributi definienti' e 'attributi caratteristici'<sup>33</sup>, 'proprietà essenziali' e 'proprietà tipiche'<sup>34</sup>, oppure *crierial, expected, possible, unexpected e excluded traits*<sup>35</sup>. Anche qui, come per la nozione di 'potenziale semantico', le rappresentazioni della struttura delle proprietà, riconducibili a modelli linguistici differenti, non sono del tutto sovrapponibili. Tuttavia, quello che caratterizza questi diversi tipi di proprietà è:

- il grado di resistenza (le proprietà tipiche sono cancellabili, quelle essenziali no), basato su ciò che una comunità di parlanti ritiene essenziale in un dato momento per il significato di un termine: «Potremmo dire che è essenziale ciò che non può essere negato senza ripattuire il significato dei termini, senza cioè modificare la convenzione su cui era basato il precedente significato per introdurne una nuova, diversa»<sup>36</sup>; «Le proprietà essenziali diventano dunque quelle che non bisogna disconoscere se, in un certo contesto, si vuole mantenere aperto il discorso, e che possono essere negate solo a prezzo di ripattuire il significato dei termini che stiamo usando»<sup>37</sup>;
- il grado di necessità, che si manifesta come grado di coerenza dei processi inferenziali: «We shall do this initially by defining a number of statuses (degrees of necessity) of semantic traits»; per esempio, «crierial and excluded traits can be diagnosed by means of entailment relations between sentences: for instance, 'animal' is a crierial trait of 'dog' because 'It's a dog' entails 'It's an animal'; 'fish' is an excluded trait of 'dog' because 'It's a dog' entails 'It's not a fish'»<sup>38</sup>;
- il grado di 'applicabilità', ossia il grado di validità in una pluralità di contesti: «alcune proprietà sembrano infatti essere più indipendenti dal contesto, e quindi più costitutive

**30** Jason Farradane, *Fundamental fallacies and new needs in classification*. In: *Theory of subject analysis: a sourcebook*, edited by Lois Mai Chan, Phyllis A. Richmond, Elaine Svenonius. Littleton (Colo.): Libraries unlimited, 1985, p. 199-209: p. 204.

**31** *Classification Research Group bulletin n. 10*, «Journal of documentation», 29 (1973), n. 1, p. 51-71: p. 56.

**32** *Classification Research Group bulletin n. 6*, «Journal of documentation», 17 (1961), n. 3, p. 157-172: p. 166.

**33** Remo Job; Rino Rumiati, *Linguaggio e pensiero*. Bologna: Il Mulino, 1984, p. 69-71.

**34** Patrizia Violi, *Significato ed esperienza*. Milano: Bompiani, 1997, p. 215 e seguenti.

**35** A. D. Cruse, *Lexical semantics* cit., p. 15-20.

**36** P. Violi, *Significato ed esperienza* cit., p. 219.

**37** Umberto Eco, *Kant e l'ornitorinco*. Milano: Bompiani, 2008, p. 208 (1. ed. 1997).

**38** A. D. Cruse, *Lexical semantics* cit., p. 16-17.

di altre, nel senso che paiono venire genericamente attivate in tutti i contesti» (Violi, cit. da Eco<sup>39</sup>);

- il livello di 'rilevanza' come criterio per l'appartenenza categoriale: «le proprietà non sono tutte allo stesso livello di importanza per quanto concerne l'appartenenza categoriale»<sup>40</sup>; «Lo schema dizionario è uno strumento di classificazione, non uno strumento di definizione»<sup>41</sup>; «Siccome non si può combattere contro le inerzie del linguaggio, mi adatterò anch'io a chiamare categorie queste voci classificatorie, ma sia chiaro che esse non contribuiscono immediatamente a dirci che cosa una cosa sia bensì come essa viene gerarchicamente ordinata in un sistema di concetti di base e concetti superordinati e subordinati»<sup>42</sup>.

Nell'ambito dell'indicizzazione per soggetto, il significato come insieme di proprietà (che non sono tutte allo stesso livello di importanza), si mostra con particolare evidenza nell'analisi a faccette. Per esempio, i termini che denotano 'oggetti' presentano vari tipi di proprietà: il materiale, la forma, il tipo di lavorazione, la funzione. Queste sono anche le 'caratteristiche di divisione' registrate nel *Thesaurus del Nuovo soggettario* sotto forma di 'etichette di nodo', che servono a organizzare i termini appartenenti alla classe degli 'oggetti'. Ci sono, tuttavia, termini che riflettono più di una caratteristica: per esempio, nel termine 'contenitori di metallo' si combinano due proprietà, il materiale e la funzione; lo stesso vale per 'soprammobili di porcellana'; mentre, nel caso di 'funi metalliche' si tratta di oggetti che presentano una specifica forma e uno specifico materiale. Per decidere la rispettiva 'sottofaccetta' di appartenenza, si stabilisce un 'ordine di precedenza' tra queste caratteristiche, considerando, per esempio, la funzione più rilevante del materiale e della forma, la forma più rilevante del materiale. Così, quando si applica l'analisi ai termini che denotano oggetti con specifiche funzioni, tutte le altre proprietà, diverse dalla funzione, sono rimosse come irrilevanti. Ma questo non è molto diverso dall'affermare che, nel caso degli oggetti con una specifica funzione, le proprietà funzionali «sembrano intuitivamente possedere uno statuto più centrale», «giocare un ruolo più essenziale delle proprietà percettive nella definizione del significato», cosicché le proprietà funzionali sono incancellabili, in quanto essenziali, le altre invece sono cancellabili, tant'è che «variazioni di forma che lasciano inalterata la funzione non hanno le stesse conseguenze semantiche che hanno alterazioni della funzione con forma inalterata»<sup>43</sup>.

Tuttavia, «le proprietà essenziali non sono tali in virtù di un'ontologia forte», ma di «una convenzione storicamente e culturalmente definita»<sup>44</sup>; in un linguaggio di indicizzazione, anche di una convenzione definita dalla comunità bibliotecaria, allo scopo di costruire un sistema di classificazione onnicomprensivo, al cui interno trovino un posto preciso l'uovo e la colonna vertebrale, il profumo di violetta e l'arrampicarsi<sup>45</sup>, compito che non sembra rientrare in nessun programma semantico della linguistica. Comune è però il fatto che, alla base delle decisioni circa le categorie attraverso le quali

39 U. Eco, *Semiotica e filosofia del linguaggio* cit., p. 133.

40 P. Violi, *Significato ed esperienza* cit., p. 221.

41 U. Eco, *Kant e l'ornitorinco* cit., p. 196.

42 *Ivi*, p. 126.

43 Cfr. P. Violi, *Significato ed esperienza* cit., p. 227-230, da cui sono tratte le citazioni.

44 *Ivi*, p. 222.

45 U. Eco, *Kant e l'ornitorinco* cit., p. 196.

organizzare i concetti, ci sia una relazione con le proprietà contenute nel significato dei termini. Ciò vale anche per categorie molto generali, come quelle che definiscono la macrostruttura di massimo livello del *Thesaurus* del *Nuovo soggettario*, sebbene la distanza tra la loro generalità e il significato dei singoli termini abbia bisogno di essere colmata attraverso livelli intermedi di specificazione delle proprietà che determinano l'organizzazione concettuale. Le 'sedie' sono 'oggetti' e come tali appartengono alla categoria delle 'cose', ma questa appartenenza, rappresentata nella 'catena gerarchica' del termine, si manifesta come un *continuum* di proprietà, da quelle più specifiche a quelle più generali e astratte: così, le 'sedie' sono un tipo di 'sedili' (oggetto fatto apposta per potersi sedere), che a loro volta sono un tipo di 'arredi' (complesso degli oggetti d'uso, delle suppellettili, dei mobili e simili che servono a rendere una abitazione o un altro ambiente adatto alla sua utilizzazione), i quali appartengono alla classe degli 'oggetti secondo la funzione e l'ambito d'uso' e così via. Ad ogni livello è la proprietà individuata come essenziale che determina l'appartenenza, mentre le altre proprietà, per esempio quelle che distinguono le 'sedie' da altri tipi di 'sedili', come 'divani', 'poltrone' ecc., sono cancellate<sup>46</sup>.

Un altro indizio di questa gerarchia di proprietà si ha con i cosiddetti termini sincategorematici, per esempio 'fiori artificiali', 'uova di cioccolato': in questi termini la 'differenza' o 'modificatore' ('artificiali', 'di cioccolato') non specifica una sottoclasse del *focus* o testa ('fiori', 'uova'), anzi ne contraddice l'appartenenza, cosicché il *focus* non identifica, come avviene normalmente, la classe cui il termine nel suo insieme appartiene e per questo è detto sincategorematico. Si può spiegare questo fenomeno linguistico supponendo che la 'differenza' blocchi una proprietà ritenuta essenziale per questa classe di concetti, ossia quella di 'organismo vivente': per questo, i fiori artificiali non sono fiori, anche se conservano, dei fiori, altre proprietà, come l'aspetto esteriore e la funzione ornamentale.

Come si vede, vi sono molte analogie tra le spiegazioni elaborate in ambito biblioteconomico e quelle in ambito linguistico, pur nella diversità dei programmi semantici cui si riferiscono. Soprattutto, ne consegue che parlare di proprietà essenziali non implica necessariamente il riferimento a qualità astratte, permanenti e imm modificabili, anche se, una volta associate a un termine, hanno una validità generale, che non può essere individualmente modificata. Parallelamente, le strutture concettuali (e i processi cognitivi sottostanti) che si basano sulle proprietà dei termini, come le strutture di un *thesaurus*, non sono avulse dalla componente linguistica e semantica: sono convenzionali, ma non arbitrarie, e come tali condivisibili.

Questa condizione si rispecchia anche nel processo materiale di costruzione cooperativa del *Thesaurus* del *Nuovo soggettario*. Il modello interattivo, mostrato nelle pagine precedenti, in generale e specificamente riguardo al significato dei termini, configura processi cognitivi che appaiono tutt'altra cosa da processi mentali intesi come un'attività psichica individuale o un'attività algoritmica identica in tutti gli individui o addirittura come «un'attività svolta da un'entità unica esistente al di là delle menti dei singoli uomini»<sup>47</sup>. Il modello si attualizza piuttosto in un percorso fatto di: a) proposte (di creazione e strutturazione di nuovi termini o di modificazione di termini esistenti), che nascono dall'attività concreta di indicizzazione dei documenti, b) regole condivise (principi, regole e procedure del sistema), c) convenzioni (accordi

<sup>46</sup> Sul ruolo delle proprietà nella definizione e organizzazione delle classi cfr. anche Claudio Gnoli, *Per una demarcazione ontologica dei concetti nell'organizzazione della conoscenza*. In: *Prospettive ontologiche: realismi a confronto*, a cura di Roberto Poli. Brescia: Morcelliana, 2014, p. 303-317.

<sup>47</sup> P. Cavalieri, *La biblioteca crea significato* cit., p. 68.

sul significato associato ai termini nuovi o revisionati). Questi tre elementi (proposte, regole, convenzioni) rispecchiano materialmente i tre momenti del processo inferenziale (abduittivo, interpretativo), descritto all'inizio: le due premesse (proposta e regola), la conclusione (convenzione). Così è del tutto condivisibile quanto afferma Cavaleri nella *pars construens* del suo libro:

È l'interazione degli indicizzatori con dei documenti concreti, o meglio gli elementi linguistici che compongono il loro paratesto e il loro testo, a determinare la necessità di produrre degli enunciati che rappresentino degli argomenti e di conseguenza di inserire nuovi termini nel thesaurus. È questa attività che determina il senso di un termine in un contesto specifico. Il senso dei termini deriva perciò dall'interazione che persone concrete – i bibliografi, i catalogatori – realizzano con dei documenti che vengono indicizzati per essere inseriti all'interno di uno specifico insieme, reale o virtuale, di documenti<sup>48</sup>.

E ancora:

I termini il cui significato è quello rintracciabile nei documenti in cui vengono usati sono termini i cui significati non sono oggettivi e certi in senso forte, ma sicuramente sono termini i cui significati possono essere condivisi, discussi, contrattati<sup>49</sup>.

Resta da descrivere, con la medesima modalità esemplificativa, un'altra forma di inferenza che si verifica nell'analisi concettuale dei documenti, le inferenze da schemi, per poi passare ai processi di sommarizzazione, non prima però di aver detto qualcosa sulla nozione di concetto, che tanto peso ha nella *pars destruens* del libro di Cavaleri.

### Concetto

Cavaleri lamenta il fatto che teorie, standard, sistemi di indicizzazione ci dicano ben poco del concetto: le definizioni fornite non ne chiariscono sufficientemente la natura e le funzioni. È vero. Eppure, l'espressione 'concetto' vi è ampiamente utilizzata. Dunque, per essere coerenti con un certo approccio al significato delle parole, è nell'uso che se ne fa che si dovrebbe cercare di precisare o ampliare il significato di 'concetto' contenuto nella sua definizione, evitando il paradosso di un ragionamento come il seguente: se «la teoria dei thesauri e gli standard che la riassumono ci dicono che il significato di ogni termine è un concetto», poiché «dei concetti che si adattino al ruolo che viene loro assegnato nelle strutture thesaurali non possiamo dire praticamente nulla: né che cosa sono, né come li conosciamo, né quale sia la loro eventuale struttura interna», allora ne consegue che anche «i termini rimarrebbero senza significato e non sapremmo quando e perché definire due termini come equivalenti e su che cosa basare l'assegnazione di relazioni gerarchiche o associative»<sup>50</sup>.

I concetti di cui, in ambito biblioteconomico, non possiamo dire praticamente nulla sono i concetti intesi come rappresentazioni mentali degli individui, come strutture di conoscenza 'interne', come entità astratte e prelinguistiche: concezione attribuita al punto di vista cognitivo, come in questo giudizio critico di Talja:

<sup>48</sup> *Ivi*, p. 206.

<sup>49</sup> *Ivi*, p. 195.

<sup>50</sup> *Ivi*, p. 187.

any processing of information is mediated by concepts and categories which, for the individual, are the model of his world. In the cognitive viewpoint, concepts and categories are defined as “mental representations” [...] Speech is understood as the unmediate expression of the individual’s original thoughts and experiences. [...] The individual’s thinking is understood in Platonic terms: as pre-linguistic, immaterial ideas which can for communication purposes, as if in afterthought, be attached onto the signs of language. But there are no immaterial concepts, categories or ideas. [...] No concepts, thoughts or meanings can exist outside language<sup>51</sup>.

Quando si parla del ‘concetto di pioggia’, del ‘concetto di verità’ o del ‘concetto di gatto’, si accede a una rappresentazione mentale di individui, il cui contenuto è un’idea immateriale e prelinguistica? Generalmente, queste espressioni si usano quando ci si riferisce a un determinato contenuto associato al termine, o quando si distingue un tipo di fenomeni da un altro o si organizza l’albero dei *directories*, per usare la metafora di Eco<sup>52</sup>: quando si dice, per esempio, che la pioggia è un fenomeno atmosferico e, più precisamente, un tipo di precipitazione atmosferica, distinto dalla neve e dalla grandine, cosicché possiamo usare l’espressione ‘pioggia’ in connessione con questo tipo di fenomeni in ogni contesto in cui se ne parla. Uso che non è dissimile da quello che se ne fa nell’indicizzazione per soggetto.

Stabilire quale sia la natura dei concetti nella nostra mente esula dai compiti della biblioteconomia. Tantomeno è in questione il contenuto variabile di un determinato concetto nella mente di singoli individui. Nel linguaggio di indicizzazione, il formato della rappresentazione di un concetto è una struttura semantica costituita dal corrispondente termine di indicizzazione e da un insieme di relazioni con altri termini, così com’è registrata nella voce del *thesaurus*; mentre, le informazioni specifiche sul concetto a cui il termine dà accesso sono espresse dalle stringhe di soggetto di cui il termine è componente semantico. Questo è quello che possiamo dire dei concetti nel linguaggio di indicizzazione. Allora, il punto non è se i concetti esistano o quale rappresentazione mentale abbiano, ma se il formato adottato sia funzionale ai processi di comprensione, categorizzazione, comunicazione ecc., che caratterizzano i programmi biblioteconomici di recupero dei documenti.

Potremmo piuttosto vedere un approccio ‘mentalista’ ai concetti nella supposizione di una relazione diretta e immediata tra concetti e termini, come sembrano adombrare le parole di Talja e com’è implicito nell’enunciazione ripetutamente stigmatizzata da Cavaleri che «il significato di un termine è un concetto». Si è visto però, nel paragrafo precedente, che questa relazione diretta non c’è: gli esempi portati sono anche una narrazione di ‘scarti’ tra termini e concetti. In effetti, la relazione è mediata dai significati, cosicché termine, significato e concetto si possono disporre ai lati del cosiddetto triangolo semiotico, ognuno dei quali rappresenta una delle tre dimensioni del linguaggio di indicizzazione: la dimensione linguistica, la dimensione del significato e quella concettuale. Alla dimensione linguistica fanno carico gli aspetti culturali, sociali, intertestuali, convenzionali del significato; mentre alla dimensione concettuale gli aspetti dell’organizzazione, intesa sia come struttura del lessico (organizzazione delle

51 Sanna Talja, *Constituting ‘information’ and ‘user’ as research objects: a theory of knowledge formations as an alternative to the information man–theory*. In: *Information seeking in context*, edited by Pertti Vakkari, Reijo Savolainen, Brenda Dervin. London: Taylor Graham, 1997, p. 67-80: p. 70-71.

52 U. Eco, *Kant e l’ornitorinco* cit., p. 197-200.

relazioni semantiche) e delle stringhe di soggetto (organizzazione delle relazioni sintattiche), sia come componente 'cognitiva' dei processi inferenziali di comprensione e interpretazione dei significati di parole, enunciati e testi.

Possiamo anche rileggere i criteri e le regole che informano il linguaggio di indicizzazione (per esempio, riguardo alla forma dei termini, alla scelta del numero singolare/plurale, alle decisioni circa la scomposizione dei termini composti ecc.) dal punto di vista dell'intreccio esistente tra queste tre dimensioni, così da vedervi non uno slittamento incoerente e ambiguo dal piano concettuale a quello linguistico, come ritiene Cavalieri, bensì il contributo che ciascuna componente apporta alla regola.

Torniamo alle definizioni di concetto e termine:

o.3.2. Concetto: unità di pensiero (o idea singola) rappresentata con un termine di indicizzazione (vedi o.3.5.). [...] o.3.5 Termine di indicizzazione: rappresentazione lessicale di un concetto, preferibilmente sotto forma di un nome o di un sintagma nominale<sup>53</sup>.

In questa forma elementare, le due definizioni hanno una limitata espressività. Ci sono informazioni latenti che possiamo cercare di esplicitare? Partiamo dal 'termine'. Possiamo aggiungere che il termine di indicizzazione è la rappresentazione lessicale di un concetto, in quanto al termine è associato un determinato significato. L'associazione al termine di un significato 'determinato' è, infatti, la condizione per cui un termine è un 'termine di indicizzazione' e ha la funzione di rappresentare un concetto. Dunque, il processo di 'significazione' (associazione termine-significato) precede quello della 'rappresentazione' (associazione termine-concetto), o meglio, nel secondo è presupposto il primo. Questo 'significato determinato' è quello descritto nelle pagine precedenti: il significato come istruzione, come potenziale semantico, come insieme di relazioni e insieme di proprietà, il significato come frutto di un processo concreto di negoziazione, il significato stabilizzato, condiviso ecc.; in una parola, il significato registrato nel *thesaurus*. 'Gatti', 'Gatti di biblioteca', 'Anagrafe commerciale', 'Tessere', 'Locomozione' rappresentano concetti, nel linguaggio di indicizzazione, in quanto a ciascuno di questi termini è associato un significato come stabilito nella voce del *thesaurus* e attualizzato nelle stringhe di soggetto di cui è componente semantico.

Il termine di indicizzazione è l'unione di un significato (stabilito) con una forma lessicale (il termine preferito con una forma preferita). Dunque, il 'concetto' è strettamente legato alle vicende attraverso le quali si stabilisce il significato dei termini, tanto che 'termine' e 'concetto', si dice, sono spesso usati in modo intercambiabile.

Un'altra informazione implicita è la duplice valenza di 'unità' nell'espressione 'unità di pensiero' contenuta nella definizione di 'concetto'. La prima valenza riguarda il significato e specificamente la questione della polisemia dei termini: la relazione univoca termine-concetto è dipendente dal giudizio sulla relazione tra il termine e i significati cui è associato nel linguaggio naturale: se questi significati si configurano come un significato generale, che copre l'applicazione del termine in tutti i possibili contesti, allora siamo in presenza di un unico termine e di un unico concetto; se, invece, uno o più contesti selezionano significati differenti del termine, che non possono essere coperti da un significato più generale, allora si avranno termini e concetti distinti<sup>54</sup>. Dunque, la relazione termine-concetto è innanzitutto delimitata paradigmaticamente, ossia rispetto al significato del termine, attraverso gli strumenti di controllo della polisemia. Invece, rispetto al contesto dell'enunciato di soggetto

<sup>53</sup> AIB. GRIS, *Guida all'indicizzazione per soggetto cit.*, p. 6-7.

di cui il termine fa parte, la relazione è delimitata sintagmaticamente, attraverso gli strumenti di composizione/scomposizione dei termini. In altre parole, l'unità di pensiero che definisce il concetto è tale (ossia costituisce un'unità) sia riguardo al significato (generalità vs. ambiguità del termine), sia riguardo al significante (termine composto vs. termini separati).

Dunque, nell'ambito dell'indicizzazione per soggetto, possiamo considerare il 'concetto' come un insieme di conoscenze e di informazioni (relative al significato di un termine e alle sue condizioni di applicazione nelle stringhe di soggetto), immagazzinate e registrate nel 'termine di indicizzazione': un insieme delimitato sia paradigmaticamente (sotto l'aspetto del significato) sia sintagmaticamente (sotto l'aspetto del significante).

Appare sufficientemente chiaro che la funzione di rappresentazione attribuita ai termini di indicizzazione nei confronti dei concetti non implica qui alcun assunto circa l'idea di rappresentazione mentale. Le due nozioni non sono isomorfe: la seconda è un'entità psicologica, la prima una funzione semantica, che possiamo esprimere anche in forma di inferenza. Se 'fumo' appare nei contesti a, b, allora 'gas prodotto dalla combustione di un combustibile' (+ cause, effetti ecc.), ma allora 'fumo' = gas di combustione; se, invece, 'fumo' appare nei contesti y, z, allora 'atto di fumare come comportamento di individui' (+ agenti, effetti ecc.), ma allora 'fumo' = fumo. Dunque, 'gas di combustione' sta per 'fumo' nei contesti a, b; mentre 'fumo' sta per 'fumo' nei contesti y, z. In altre parole, 'gas di combustione' rappresenta il concetto 'gas prodotto dalla combustione di un combustibile' (un materiale), mentre 'fumo' rappresenta il concetto 'atto, abitudine di fumare' (un'attività). E viceversa: se 'gas di combustione', allora 'gas prodotto dalla combustione di un combustibile', indipendentemente dal fatto che nei contesti a, b il concetto sia espresso da 'gas di combustione', 'fumo', 'gas a fuel', 'gaz', 'gas de combustión' ecc. Quest'ultima conseguenza spiega la ragione per cui, in alcune definizioni degli standard, il 'concetto' venga definito come 'un'entità astratta indipendente dai termini utilizzati per esprimerla'.

È in questione, dunque, una posizione di preminenza o un ordine di priorità tra termine e concetto? Come mostra il ragionamento sopra, nel linguaggio di indicizzazione l'espressione 'fumo', che compare nei contesti a, b, è tradotta con 'gas di combustione', a rappresentare il 'fumo come gas di combustione', che è appunto il significato di 'fumo' in quei contesti, distinto da 'fumo come abitudine di fumare': l'inizio, sul piano processuale, spetta a 'fumo', la fine a Gas di combustione; la priorità, sul piano semantico, è quella che, se entra del fumo, quello che esce non sia, in questo caso, il fumo delle sigarette.

I significati di 'fumo' (gas di combustione o atto/abitudine di fumare) sono rappresentazioni mentali di che tipo? Che struttura ha la rappresentazione? Compositiva, prototipica ecc.? È presente nella nostra memoria a lungo termine «un dizionario mentale composto da una serie di entrate a ciascuna delle quali è associata una parola, e sotto ciascuna delle quali è immagazzinato un certo corpus di informazioni relative al suo significato, alla sua pronuncia, alle sue caratteristiche sintattiche e morfologiche, che ci permettono di farne un uso competente», entrate che chiamiamo 'concetti lessicali'<sup>55</sup>? Queste domande e, più in generale, quelle riguardanti lo stato

54 Ivi, p. 83. Cfr. anche la distinzione tra 'general' e 'ambiguous' in A. D. Cruse, *Lexical semantics* cit., p. 51-52.

55 Per un bilancio sui concetti lessicali cfr. Luca Gasparri, *Concetti lessicali*, «APhEx», n. 8 (2013), p. 81-111: p. 84.

ontologico (natura e struttura) dei concetti sono oggetto di studio della filosofia, della psicologia e della linguistica, poiché implicano considerazioni sul funzionamento della mente e del linguaggio e sui loro rapporti.

Concepts are the constituents of thoughts. Consequently, they are crucial to such psychological processes as categorization, inference, memory, learning, and decision-making. This much is relatively uncontroversial. But the nature of concepts – the kind of things concepts are – and the constraints that govern a theory of concepts have been the subject of much debate. This is due, at least in part, to the fact that disputes about concepts often reflect deeply opposing approaches to the study of the mind, to language, and even to philosophy itself<sup>56</sup>.

Invece, per quanto riguarda la riflessione in ambito biblioteconomico, se intendiamo per ‘concetto’ il significato associato a un ‘termine di indicizzazione’, nel senso tecnico precisato sopra, è cruciale intendersi sulle caratteristiche e sulle pratiche di questa associazione, riguardo alle quali mi pare che le note che precedono mostrino una sostanziale convergenza con l’assunto di Cavaleri:

L’assunto fondamentale, riguardo il significato, è che un termine significhi un’entità culturale e linguistica, autonoma da ogni parlante, un’entità che per la sua prima caratteristica, ‘culturalità’ è contingente, per la seconda, ‘semioticità’ è oggetto di interpretazione e per la terza, ‘autonomia’, sfugge al problema della soggettività. La relazione che ogni soggetto ha con un termine è quella di un utilizzatore che segue regole pragmatiche, cioè atte ad ottenere degli scopi comunicativi, socialmente modificabili, ma il cui significato non è disponibile per l’individuo singolo<sup>57</sup>.

La caratteristica contingente si può riferire alla rappresentazione, ossia all’associazione ‘termine di indicizzazione’-‘significato’ come registrata nel vocabolario di indicizzazione (equivalente alla relazione ‘termine di indicizzazione’-‘concetto’), in quanto implica un aspetto sincronico (il significato stabilizzato) e un aspetto diacronico (il significato mutevole), che potremmo qualificare anche come ‘convenzionale’. La caratteristica ‘essere oggetto di interpretazione’ si può riferire alla ‘significazione’, ossia all’associazione ‘termine’-‘significato’ come interpretata nel testo del documento e poi tradotta nel/nei termini di indicizzazione. La caratteristica ‘non disponibilità del significato per l’individuo singolo’ si può riferire alla ‘contestualizzazione’, ossia ad un uso coerente del termine nell’indicizzazione dei documenti, che equivale alla costanza del significato del termine di indicizzazione in tutti i contesti in cui compare. Significazione, rappresentazione e contestualizzazione rappresentano appunto i tre processi fondamentali dell’indicizzazione per soggetto.

<sup>56</sup> Eric Margolis; Stephen Laurence, *Concepts*. In: *The Stanford encyclopedia of philosophy*, edited by Edward N. Zalta, 2014, <<http://plato.stanford.edu/archives/spr2014/entries/concepts/>> (ultima consultazione: 01/10/2015).

<sup>57</sup> P. Cavaleri, *La biblioteca crea significato* cit., p. 207.

### Inferenze da schemi

Se in un testo compare il termine ‘lettura ad alta voce’, allora è probabile che ci si riferisca a una o più di queste attività, con i loro rispettivi ‘partecipanti’ (genitori, insegnanti, animatori, bibliotecari, attori, doppiatori ecc.; finalità, modalità, tecniche, spazi, tempi ecc.):

1. attività rivolta a bambini in età prescolare e scolare in un contesto familiare;
2. attività rivolta a bambini in un contesto scolastico;
3. attività di promozione della lettura rivolte a bambini in età prescolare e scolare da parte delle biblioteche;
4. attività svolte in istituti per anziani, ospedali, reparti pediatrici ecc. da parte di associazioni di volontariato;
5. attività rivolte a persone che non sono in grado di leggere da sé (persone malate, persone con disabilità);
6. attività svolte da professionisti nel contesto della comunicazione pubblica, delle arti e dello spettacolo.

Questo elenco di possibili inferenze indica, con molta approssimazione, i principali contesti di riferimento del termine ‘lettura ad alta voce’, ossia i normali contesti nei quali ci aspettiamo che compaia significativamente quel termine. Un termine non ha solo un ‘significato tipico’, ma anche ‘contesti tipici’ che si attivano al momento in cui viene selezionato. Si tratta di conoscenze di carattere esperienziale, culturale, intertestuale, conoscenze condivise e condivisibili, descritte e descrivibili, dunque prevedibili.

È possibile condensare queste scene in scene più inclusive, per esempio riunendo in un’unica scena gli elementi che fanno riferimento alla lettura ad alta voce in un contesto educativo, fino ad uno schema più astratto, del tipo: lettura ad alta voce (attività) a bambini, anziani, malati, persone con disabilità ecc. (beneficiario) di fiabe, racconti, poesie, romanzi ecc. (oggetto) da parte di genitori, insegnanti, animatori ecc. (agente), effettuata con determinate modalità (metodi, tecniche, strumenti). Queste funzioni – beneficiario, oggetto, agente, metodi, tecniche, strumenti – rappresentano la cornice dei ‘casi’ o dei ‘ruoli’ associabili alla lettura ad alta voce. Se, al contrario, questa cornice si espande, con l’aggiunta di conoscenze specifiche, otteniamo le scene descritte sopra, ulteriormente espandibili in base al grado di conoscenza che abbiamo degli elementi che ne fanno parte. A loro volta, i ‘casi’ o ‘ruoli’ attivano o presuppongono determinati schemi: per esempio, un’azione transitiva a tre argomenti attiva uno schema i cui nodi sono rappresentati da un oggetto, un agente e un beneficiario; mentre, il beneficiario presuppone uno schema in cui sia presente un’azione transitiva a tre argomenti.

I casi associabili a qualsiasi tipo di azione (e ad altri tipi di concetti con valenza sintattica, come ‘parti’ e ‘proprietà’) sono individuati nello ‘schema dei ruoli’, quale quello descritto nella Guida GRIS e nella Guida al *Nuovo soggettario*.

‘Scene’, ‘cornici di casi’, ‘schema dei ruoli’ rappresentano, a livelli diversi di astrazione, tre modalità di rappresentazione schematica (o *frame*) delle relazioni concettuali sul piano sintagmatico: a livello semantico (conoscenze specifiche relative ad azioni, soggetti, oggetti, eventi ecc., che fanno parte di una ‘scena’ normalmente associata a una particolare espressione), a livello lessicale (la ‘cornice di casi’ associabili a una particolare espressione), a un livello più astratto, lo ‘schema dei ruoli’ (come struttura adattabile a qualsiasi situazione specifica).

Nell’ambito dell’analisi concettuale dei documenti, questi schemi svolgono una funzione euristica e, in particolare, funzioni di supporto nella:

- a. formulazione di inferenze che aiutano l’indicizzatore nella comprensione dei testi:

se si parla della lettura ad alta voce in un contesto familiare, allora è ragionevole aspettarsi esperienze, azioni, soggetti, oggetti, modalità ecc. appropriati al contesto (la 'scena' della lettura ad alta voce nel contesto familiare);

b. formulazione di inferenze che aiutano l'indicizzatore nella selezione dei concetti ritenuti essenziali per la descrizione del soggetto: se si parla della lettura ad alta voce, poiché un'azione transitiva a tre argomenti come 'lettura ad alta voce' attiva una specifica 'cornice di casi', allora occorre controllare se siano presenti anche l'oggetto, l'agente e il beneficiario dell'azione;

c. integrazione delle informazioni (le informazioni riferibili a un unico tema non sono registrate come unità discrete, ma organizzate in un'unità più complessa): sono integrabili (in un unico enunciato e in un'unica stringa di soggetto) tutte quelle conoscenze, riguardo ad azioni, soggetti, oggetti, modalità ecc. della lettura ad alta voce, che si possono strutturare secondo le relazioni di ruolo comprese nello 'schema dei ruoli'.

Un esempio. Premesso che «nessun testo vien letto indipendentemente dall'esperienza che il lettore ha di altri testi»<sup>58</sup>, tuttavia si può supporre che si disponga di scarsi elementi conoscitivi riguardo agli scenari adombrati dal titolo del libro di Nadia Venturini *Neri e italiani ad Harlem: gli anni Trenta e la guerra d'Etiopia* (Roma: Edizioni Lavoro, 1990). Dall'analisi del titolo nascono allora una serie di domande. La 'e' che congiunge neri e italiani indica la coordinazione di due elementi oppure, come sembra più probabile, una relazione? Quale tipo di relazione? Un'interazione, come quella espressa con l'enunciato 'rapporti tra neri e italiani'? Oppure, un confronto tra le condizioni economiche e sociali dei due gruppi? Harlem dovrebbe indicare il contesto spaziale, come 'anni Trenta' quello temporale: è così? In che rapporto sta la guerra d'Etiopia con gli altri termini? Fortunatamente, un'ampia e articolata introduzione autoriale fornisce molti elementi per rispondere a queste domande. Il primo elemento è una smentita (rispetto alla previsione) riguardo alla funzione svolta da Harlem, scelta qui come 'studio di caso' (p. 14). Dunque, si devono verificare le condizioni che soddisfano il *frame* dello studio di caso (un soggetto generale + una sua esemplificazione, in questo caso l'area studiata; la relazione tra soggetto generale ed esemplificazione stabilita esplicitamente nel documento; il nome citato nell'esemplificazione come termine di accesso utile). Poco più avanti, si chiarisce il tipo di relazione tra neri e italiani: «Il raffronto tra neri ed immigrati può essere condotto su due piani: si può compiere un'analisi comparativa, oppure si possono esaminare le interrelazioni. A mio parere, questi due piani di analisi sono interdipendenti, perché non è possibile esaminare la natura dell'interazione fra i gruppi, senza averli raffrontati sul piano economico e sociale; d'altronde la situazione dei gruppi non dipende solo dal contesto economico e culturale della società americana, ma anche dai rapporti che si instaurano fra di loro, soprattutto sui luoghi di lavoro e nei quartieri» (p. 16). Questa duplice chiave di lettura (studio comparativo e analisi dei rapporti) è riproposta in molti passaggi dell'introduzione. Qual è il *frame* appropriato per esprimere la relazione tra neri e italiani descritta nel documento? Più complessa appare la relazione della guerra d'Etiopia con tutti gli altri elementi. È esprimibile nella forma di una generica correlazione, oppure di un nesso causale del tipo 'influssi di ... su ...' o 'effetti di ... su ...'? Una volta chiarita questa relazione, è possibile organizzare tutti gli elementi in un enunciato che esprima il tema dominante del libro, per esempio: 'influssi della Guerra d'Etiopia sui rapporti tra neri e immigrati italiani negli Stati Uniti durante gli anni Trenta. Il caso di Harlem'.

Dunque, il documento (a partire dal titolo) suscita domande e sollecita ipotesi interpretative, che si avvalgono dell'attivazione di *frames*, come quelli dello 'studio di caso', dell'interazione', del 'confronto', della 'correlazione', dell'agente indiretto'. Così, l'esame

58 U. Eco, *Lector in fabula* cit., p. 81.

dell'informazione contenuta nel testo (processo dal basso) interagisce con l'informazione depositata negli schemi attivabili (processo dall'alto), dando luogo a ipotesi interpretative riguardo alle relazioni instaurate tra neri, italiani, Harlem, guerra d'Etiopia e a ipotesi di organizzazione degli elementi interrelati in uno o più enunciati di soggetto. L'indicizzatore potrà formulare ipotesi (sia di interpretazione che di organizzazione dei concetti e delle loro relazioni) con maggiore sicurezza e probabilità di successo quanto più disponga di una competenza intertestuale relativa alle molte e complesse 'scene' che quelle espressioni possono evocare: le condizioni economiche e sociali dei neri negli Stati Uniti, l'emigrazione italiana negli Stati Uniti, le dinamiche tra i gruppi etnici nei centri urbani degli Stati Uniti, gli avvenimenti della guerra d'Etiopia ecc.

Com'è noto, lo schema dei ruoli di PRECIS, da cui sono derivati quelli di GRIS e del Nuovo soggettario, si ispira alla 'grammatica dei casi' del linguista Fillmore: la 'struttura di base della frase' che «consiste di un verbo e di uno o più sintagmi nominali, ciascuno associato al verbo in una particolare relazione di caso»; le 'relazioni di caso' che «comprendono concetti come quelli di agente, strumentale, oggettivo, fattitivo, locativo, benefattivo e forse vari altri ancora»; la nozione di 'cornice casuale' come insieme dei casi in cui un determinato verbo può essere inserito<sup>59</sup>. L'interpretazione di 'lettura ad alta voce' come un deverbale che prevede uno schema a tre argomenti (agente, oggetto, beneficiario) è l'esemplificazione di una rappresentazione schematica che ha una sostanziale analogia con la nozione di 'cornice casuale' di Fillmore, come lo è la struttura della 'nota sintattica' di cui sono corredati alcuni termini del *Thesaurus* del Nuovo soggettario.

La 'cornice casuale' sarà poi sviluppata da Fillmore nella *Frame semantics*: «il *frame* non è altro che lo sviluppo della rappresentazione schematica dei ruoli connessi a un predicato, con l'aggiunta di un più ricco background di conoscenza enciclopedica, oltre alla possibilità di estendere lo schema anche ad altre parti del discorso»<sup>60</sup>.

A particularly important notion, figuring especially in recent work in linguistics, cognitive psychology, and artificial intelligence, is the notion that goes by such names as "frame", "schema", and "scenario". Briefly, the idea is that people have in memory an inventory of schemata for structuring, classifying, and interpreting experiences [...]. The concept of frame does not depend on language, but as applied to language processing the notion figures in the following way. Particular words or speech formulas, or particular grammatical choices, are associated in memory with particular frames, in such a way that exposure to the linguistic form in an appropriate context activates in the perceiver's mind the particular frame-activation of the frame, by turn, enhancing access to the other linguistic material that is associated with the same frame<sup>61</sup>.

Una ricca, chiara e suggestiva spiegazione della «natura schematica del significato», a partire dalla *Frame Semantics* di Fillmore, e più in generale del «ruolo del contesto

59 Charles J. Fillmore, *Il caso del caso*. In: *Gli universali nella teoria linguistica*, a cura di Emmon Bach, Robert T. Harms. Torino: Boringhieri, 1978, p. 27-131 (ed. orig. 1968). Cfr. anche Jutta Sørensen; Derek Austin, *PRECIS in multilingual context. Part 2, a linguistic and logical explanation of the syntax*, «Libri», 26 (1976), n. 2, p. 108-139.

60 P. Violi, *Significato ed esperienza* cit., p. 283.

61 Charles J. Fillmore, *Frame semantics and the nature of language*. In: *Origins and evolution of language and speech*, edited by Stevan R. Harnad, Horst D. Steklis, Jane Lancaster. New York: The New York Academy of sciences, 1976, p. 20-32: p. 25.

nella determinazione del significato» si trova in *Significato ed esperienza* di Violi, da cui è tratta la seguente citazione:

Poiché quel certo termine è stato selezionato cercherò nell'«intorno», o costruirò ex novo, un contesto di interpretazione possibile. Le parole operano così come potenti dispositivi abduttivi, istruzioni per la costruzione della situazione e per la sua interpretazione: la parola forza l'attivazione di un contesto di interpretabilità creando le sue proprie condizioni di applicabilità. In questa prospettiva, invece di pensare al contesto come elemento che, dall'esterno, sottrae proprietà al termine (o comunque le modifica), si vede il termine come attivatore di contesto; è come se il termine, una volta inserito nella catena sintagmatica del discorso, ci costringesse a 'guardarci attorno' per individuare le condizioni che ne giustifichino l'uso<sup>62</sup>.

Note sono le «sceneggiature» descritte da Eco come dispositivi inferenziali nell'interpretazione dei testi narrativi<sup>63</sup>. Il ruolo strategico dei «*propositional schemata*» nei processi di comprensione dei testi e nella costruzione delle «macrostrutture» è illustrato da van Dijk e Kintsch: «many of the inferences made in discourse comprehension are schema driven»<sup>64</sup>. Per una presentazione sintetica dell'«azione degli schemi» nell'interpretazione dei testi e nella formazione delle macrostrutture, con alcune indicazioni felici anche per l'ambito dell'indicizzazione per soggetto, si rinvia a Job e Rumiati<sup>65</sup>.

Queste annotazioni e i riferimenti alla linguistica hanno soprattutto lo scopo di sottolineare la 'struttura profonda' dello schema dei ruoli: prima che uno strumento per ordinare i termini nella stringa di soggetto, è uno strumento per identificare le funzioni semantiche fondamentali attorno alle quali si organizza il contenuto concettuale del testo ad un livello globale e, al tempo stesso, uno strumento per analizzare le relazioni e le funzioni semantiche svolte dai concetti nella definizione del soggetto<sup>66</sup>.

Se le due applicazioni dello schema sono accomunate, sul piano semantico, da una medesima struttura, tuttavia differiscono sul piano operativo. Nell'applicazione all'analisi dell'enunciato di soggetto, lo schema opera su un'unità semantica definita e compiuta (l'enunciato di soggetto) e, in unione con alcune regole, determina l'ordine dei termini nella stringa di soggetto. Mentre, in funzione dell'analisi del testo, svolge un ruolo euristico, manifesta un carattere dinamico, come si è visto nel caso di *Neri e italiani ad Harlem*, guidando nel corso dell'esame la formazione di ipotesi su azioni, eventi, partecipanti, loro ruoli e relazioni, spazio, tempo ecc., che, se confermate, andranno a costituire l'enunciato di soggetto, dal momento che «a full macroproposition – that is, a topic or theme in our terms – can be derived only with a clause expressing a proposition, so that possible world, time, place, and kind of predicate (action, event), or other participants – and participant roles – are known»<sup>67</sup>.

62 P. Violi, *Significato ed esperienza* cit., p. 279 (cfr. tutto il cap. 9, p. 271-302).

63 U. Eco, *Lector in fabula* cit., in part. p. 79-84.

64 T. A. van Dijk; W. Kintsch, *Strategies of discourse comprehension* cit., p. 48 (cfr. anche il cap. 4: *Propositional strategies*).

65 R. Job; R. Rumiati, *Linguaggio e pensiero* cit., p. 89-104.

66 Cfr. AIB. GRIS, *Guida all'indicizzazione per soggetto* cit., p. 16; anche BNCF, *Nuovo soggettario* cit., p. 104-105.

67 T. A. van Dijk; W. Kintsch, *Strategies of discourse comprehension* cit., p. 205.

### Processi di sommarizzazione

Di fronte al titolo *Neri e italiani ad Harlem. Gli anni Trenta e la guerra d'Etiopia*, l'indicizzatore formula la domanda 'di che cosa si parla?', cui segue la risposta 'si parla probabilmente dei rapporti tra neri e immigrati italiani ad Harlem negli anni Trenta', che presuppone l'interpretazione della congiunzione 'e' di 'neri e italiani' non come coordinazione, ma come un tipo specifico di interrelazione. E la guerra d'Etiopia? Risposta: 'probabilmente si tratta di una qualche influenza esercitata da questo avvenimento bellico sui rapporti tra i due gruppi sociali'. Oppure, su quest'ultima domanda si sospende ogni ipotesi, in attesa di ulteriori indizi. Supponiamo che l'indicizzatore, attraverso l'analisi del titolo, sia arrivato a questo punto, supportato da conoscenze di vario tipo, come quelle descritte nei paragrafi precedenti (significati lessicali, scene ecc.). Tuttavia, il suo compito non è quello di «tenere in esercizio la semiosi», come direbbe Eco<sup>68</sup>, ma di interpretare il testo allo scopo di individuare il tema o i temi del documento. In altri termini, il suo interesse non è quello di sapere di che cosa si parla nel titolo (interesse comunque non secondario, dato che il titolo è il principale indicatore tematico), ma di che cosa parla il testo del documento. Si cercano così conferme, convalide attraverso l'esame dei «discorsi di scorta»<sup>69</sup>, ossia di quelle parti del documento dove è più probabile che appaiano espressioni che indicano di che cosa tratta il testo, oppure attraverso la 'progressione tematica' che il testo manifesta e la riorganizzazione dei temi in un tema principale che li ingloba. Man mano che si procede nell'esame del documento, conferme, smentite, precisazioni, integrazioni, indotte dagli elementi semantici che si succedono nel testo, costringono a riformulare l'ipotesi iniziale. In questo caso, le conferme riguardano il tipo di relazione tra i due gruppi sociali, la smentita il ruolo di Harlem, che diventa 'il caso di Harlem', le precisazioni l'impatto della guerra d'Etiopia sui rapporti tra neri e italiani ad Harlem.

L'enunciato 'influssi della Guerra d'Etiopia sui rapporti tra neri e immigrati italiani negli Stati Uniti durante gli anni Trenta. Il caso di Harlem' sottende una molteplicità di temi, trattati nel testo: alcuni si possono considerare inclusi a vario titolo nel tema di base, altri ne restano esclusi, perché collaterali o di rilevanza marginale. La maggior parte di essi rappresentano uno sviluppo o una specificazione dei due 'predicati' principali (rapporti di ... con ...; influssi di ... su ...), su cui è costruito l'intero enunciato. Del primo fanno parte temi riconducibili al confronto tra i due gruppi etnici (situazione abitativa, occupazionale, lavorativa ecc.) e alla loro interazione (conflitti razziali, competizione economica, pregiudizi, segregazione ecc.). Del secondo, temi riconducibili all'impatto della guerra d'Etiopia sui loro rapporti (sommosse, boicottaggi, alleanze, manifestazioni, campagne di propaganda ecc., con i loro principali protagonisti, i partiti, le organizzazioni politiche, gli organi di stampa ecc.). Collaterali si possono considerare temi come quelli riguardanti la posizione del governo degli Stati Uniti o dell'Unione Sovietica nei confronti della guerra. Marginale è, per esempio, il tema della presenza della comunità ebraica ad Harlem negli stessi anni.

Tanti enunciati particolari sono sostituiti da due enunciati generali, che definiscono il tema in termini di 'rapporti' tra due specifici gruppi etnici e di 'influssi' determinati da un evento bellico. Questa sostituzione corrisponde a una delle macro-operazioni di sommarizzazione, la 'costruzione', che è appunto la sostituzione di enunciati particolari, che esprimono singoli temi, con un enunciato più generale che li comprende<sup>70</sup>.

<sup>68</sup> U. Eco, *Lector in fabula* cit., in particolare p. 60.

<sup>69</sup> P. Cornea, *Introduzione alla teoria della lettura* cit., p. 174-177.

<sup>70</sup> Sulle macro-operazioni cfr. AIB. GRIS, *Guida all'indicizzazione per soggetto* cit., p. 16.

Tuttavia, operativamente si è trattato piuttosto di una ‘selezione’, poiché il tema di base è, come si è visto, enunciato esplicitamente nel documento.

‘Selezione’ da una parte, ‘costruzione/generalizzazione’ dall’altra, rappresentano due strade per l’individuazione del tema di base, a seconda che il tema sia indicato dal documento stesso, oppure sia da ricostruire. «Sovente il segnale è esplicito: il titolo appunto, o una espressione manifestata che dice di cosa appunto il testo si vuole occupare. Talora invece il *topic* è da cercare»<sup>71</sup>.

Ma i titoli possono essere ingannevoli e l’ipotesi iniziale fallire il bersaglio, come avverrebbe se, nel caso di *La minaccia dei pesticidi* di Van den Bosch e J.-P. Aeschlimann (Padova: Muzzio, 1989), si rispondesse alla domanda ‘di che si parla’ con ‘si parla probabilmente dei danni all’ambiente naturale provocati dall’uso di pesticidi’, ipotesi avvalorata dal senso figurato di ‘minaccia’ associato a ‘pesticidi’, che evoca la scena dell’impatto ambientale o degli agenti inquinanti (+ effetti, ambienti, monitoraggio, controllo, prevenzione, tutela ecc.). Lo sviluppo dei temi nel documento conferma questa interpretazione? Se nel testo si parla dei pesticidi come causa di inquinamento, questo tema può essere considerato come il tema dominante del testo? Il titolo italiano adombra, dei pesticidi, il pericolo per l’ambiente, quello francese (*L’engranaage des pesticides*) l’impatto sul controllo degli animali nocivi alle piante coltivate, quello originale americano (*The Pesticide conspiracy*) l’aspetto ‘politico’, ossia «gli intrighi politici, i tradimenti, la corruzione e gli inganni» che hanno permesso lo sviluppo dell’impiego di pesticidi in agricoltura a scapito della lotta integrata. Tutti e tre i titoli si possono considerare ‘rematici’ (descrivono ciò che si dice nel testo) e ‘metaforici’ (lo fanno in modo figurato), ma ognuno ne coglie un aspetto o una delle tesi, dalle quali si può ripartire per individuare i rispettivi temi:

1<sup>a</sup> (tesi) Esistono ceppi di insetti nocivi alle colture sempre più resistenti a una vastissima varietà di pesticidi; lo stesso può dirsi di altri organismi nocivi (erbe infestanti).

1<sup>b</sup> (domanda) Quali effetti hanno i pesticidi sugli organismi nocivi alle colture?

1<sup>c</sup> (tema) Effetti dei pesticidi sugli organismi nocivi alle colture.

2<sup>a</sup> (tesi) La resistenza ai pesticidi degli organismi nocivi alle colture costringe a un impiego crescente di queste sostanze in agricoltura (frequenza dei trattamenti + concentrazione/combinazione di sostanze attive).

2<sup>b</sup> (domanda) Quali conseguenze ha la resistenza ai pesticidi degli organismi nocivi alle colture sull’impiego di queste sostanze in agricoltura?

2<sup>c</sup> (tema) Conseguenze della resistenza ai pesticidi degli organismi nocivi alle colture sull’impiego di queste sostanze in agricoltura.

3<sup>a</sup> (tesi) Produttori agrofarmaceutici, poteri pubblici, coltivatori, consumatori, organi di stampa esercitano pressioni per mantenere le cose allo stato attuale in fatto di lotta agli organismi nocivi alle colture.

3<sup>b</sup> (domanda) Quale ruolo svolgono produttori agrofarmaceutici, poteri pubblici ecc. nelle politiche di lotta antiparassitaria?

3<sup>c</sup> (tema) Ruolo dei produttori agrofarmaceutici, dei poteri pubblici ecc. nelle politiche di lotta antiparassitaria.

4<sup>a</sup> (tesi) L’impiego dei pesticidi in agricoltura provoca gravi danni all’ambiente.

4<sup>b</sup> (domanda) Quali effetti ha sull’ambiente l’impiego dei pesticidi in agricoltura?

4<sup>c</sup> (tema) Effetti dell’impiego dei pesticidi in agricoltura sull’ambiente.

71 U. Eco, *Lector in fabula* cit., p. 91 (sul *topic* cfr. il § 5.2, p. 87-92).

5<sup>a</sup> (tesi) Nonostante vi siano molteplici esempi di lotta integrata efficace contro gli organismi nocivi alle colture (nella coltivazione della soia, del cotone, nell'agrumicoltura ecc.), le politiche di produttori agrofarmaceutici, poteri pubblici ecc. (favorevoli all'impiego dei pesticidi in agricoltura) ne hanno limitato la diffusione.

5<sup>b</sup> (domanda) Quali conseguenze hanno avuto le politiche di produttori agrofarmaceutici, poteri pubblici ecc. sulla diffusione della lotta integrata?

5<sup>c</sup> (tema) Conseguenze delle politiche di produttori agrofarmaceutici, poteri pubblici ecc. sulla diffusione della lotta integrata.

Riferite al titolo, si possono considerare le tesi  $1^a + 2^a = engranage$ ,  $3^a + 5^a = conspiracy$ ,  $4^a = minaccia$ . In *Testo e contesto*, si proponeva di riassumere i temi nel macrotema 'influssi di gruppi di pressione sulle politiche relative all'impiego di pesticidi in agricoltura negli Stati Uniti', supponendo che:

- 'impiego dei pesticidi in agricoltura' copra  $1^c$  e  $2^c$  (i pesticidi sono impiegati in agricoltura per combattere gli organismi nocivi alle colture e di questo impiego l'autore del documento ci dirà probabilmente se e quanto sia efficace);

- 'influssi di gruppi di pressione sulle politiche relative a ...' rispecchi  $3^c$ ;

-  $4^c$  sia una possibile conseguenza di 'impiego di pesticidi in agricoltura' e  $5^c$  degli 'influssi di gruppi di pressione sulle politiche relative a ...'.

I temi individuati sono, a loro volta, il risultato di operazioni di generalizzazione ('insetticidi' e 'diserbanti' diventano 'pesticidi'; produttori agrofarmaceutici, poteri pubblici, organi di informazione ecc. diventano 'gruppi di pressione') e di costruzione (gli interventi di lotta chimica contro gli organismi nocivi alle colture e i processi da essa innescati sono sussunti sotto 'impiego dei pesticidi in agricoltura'; intrighi, tradimenti, corruzione e inganni sono sostituiti da 'politiche').

Negli atti del convegno *L'ape come sentinella dell'ambiente* (Firenze: Regione Toscana, Giunta regionale, 1993), i diversi contributi sviluppano temi finalizzati a illustrare le 'possibilità di impiego delle api per il monitoraggio dell'inquinamento dell'ambiente naturale' (patologia e mortalità delle api, residui di prodotti fitosanitari, di radionuclidi e di elementi in traccia nei prodotti delle api). Rimanendo aderenti al linguaggio del titolo e, al tempo stesso, utilizzando una delle funzioni dello schema dei ruoli (la classe quasi generica) potremmo riesprimere il tema dominante con l'enunciato 'api come indicatori biologici (sentinelle) impiegati nel monitoraggio dell'inquinamento dell'ambiente naturale'.

Uno dei contributi al convegno verte sulle problematiche relative alla distribuzione e alla vendita di prodotti antiparassitari e diserbanti nel territorio della USL n. 2. Il tema è cancellato in quanto segna una discontinuità tematica rispetto al tema scelto come tema di base, anche se tra i fatti denotati dai due temi esiste una connessione, poiché i prodotti antiparassitari e i diserbanti costituiscono una delle cause di inquinamento dell'ambiente naturale, di cui le api sono un possibile indicatore biologico. Analogamente, un tema marginale come la 'presenza della comunità ebraica ad Harlem negli anni Trenta' è cancellato, sebbene sia connesso al tema di base in quanto questo fatto co-occorre (stesso luogo e periodo) con il fatto denotato dal tema di base (neri e immigrati italiani ad Harlem negli anni Trenta). Ma anche i temi 'effetti dei pesticidi sugli organismi nocivi alle colture', 'effetti dell'impiego dei pesticidi in agricoltura sull'ambiente' e 'politiche relative alla diffusione della lotta biologica contro gli organismi nocivi alle colture' sono soggetti a forme di cancellazione, in quanto connessi al tema di base da più strette relazioni logiche, che, nell'ordine, manifestano gradi diversi di implicitazione. «These relations are of various types, differing in

'strictness'; they range from possibility, via probability, to different kinds of necessity»<sup>72</sup>. «One aspect that distinguishes text-based inferences is the degree of certainty with which they can be made. Some inferences appear to be necessary consequences of the text, whereas others we are less certain about and still others we would regard as no more than plausible or possible conjectures»<sup>73</sup>.

Possiamo, dunque, considerare i vari tipi di 'cancellazione' (un'altra delle macro-operazioni di sommarizzazione) come gradi diversi di 'implicitazione' di un tema rispetto al tema di base. Ma, «'delete' here does not mean 'delete from memory' but 'delete from the macrostructure'. Thus, a given text proposition – a microproposition – may be deleted from the text's macrostructure but, nevertheless, be stored in memory and subsequently recalled as a microproposition»<sup>74</sup>. Unendo quest'ultima osservazione con quelle immediatamente precedenti e trasportandole sul piano dell'indicizzazione, potremmo concludere che un tema cancellato è candidato a essere considerato un eventuale tema particolare da selezionare e che questa candidatura è da avanzare con maggior forza quanto minore è il grado di implicitazione del tema particolare nel tema di base, sulla base della presupposizione che «by inverse application of the macro-rules at least part of detailed information will available by inference procedures»<sup>75</sup>.

Pertanto, i processi descritti, compreso l'impiego delle macro-operazioni, non escludono che

a part of a text may be neither a logically central theorem nor a hypothesis, neither the premise nor the conclusion of a central argument, and it may not state arguments or conclusions that are even peripherally related to the logically central theme. Yet it may state a proposition that is significant in the context of the author's other works or in the context of scholarship on the subject<sup>76</sup>.

I processi di sommarizzazione, dunque, si possono avvalere sì di regole, come le macro-regole, ma queste non presuppongono che «le regole mentali di riduzione semantica [...] non ammettano l'influenza del contesto sociale, politico, economico e culturale» e che «il significato del testo sia fisso e immutabile»<sup>77</sup>; non implicano una «valenza

72 Teun A. van Dijk, *Complex semantic information processing*. In: *Natural language in information science: perspectives and directions for research*, edited by Donald E. Walker, Hans Karlgren, Martin Kay. Stockholm: Scriptor, 1977, p. 127-163: p. 131.

73 T. A. van Dijk; W. Kintsch, *Strategies of discourse comprehension* cit., p. 49.

74 Walter Kintsch; Teun A. van Dijk, *Toward a model of text comprehension and production*, «Psychological review», 85 (1978), n. 5, p. 363-394: p. 372.

75 T. A. van Dijk, *Complex semantic information processing* cit., p. 140.

76 Bernd Frohmann, *Rules of indexing: a critique of mentalism in information retrieval theory*, «Journal of documentation», 46 (1990), n. 2, p. 81-101: p. 96.

77 Teresa Grimaldi, *025.47 Catalogazione per soggetto (Teoria dell'indicizzazione semantica)*. In: *Biblioteconomia: guida classificata*, diretta da Mauro Guerrini, condirettore Gianfranco Crupi; a cura di Stefano Gambari, collaborazione di Vincenzo Fugaldi; presentazione di Luigi Crocetti. Milano: Bibliografica, 2007, p. 625-628. Su *L'indicizzazione dal punto di vista cognitivo* si rinvia agli omonimi saggi di Teresa Grimaldi pubblicati in «Il bibliotecario», 12 (1995), n. 1, p. 277-301 (prima parte) e 13 (1996), n. 2, p. 259-293 (seconda parte).

esplicativa in senso mentalista e a-contestuale»<sup>78</sup>, né una «natura strettamente algoritmica del processo cognitivo che viene svolto quando una mente interagisce con un documento»<sup>79</sup>, né «the explanatory power of internally realised and tacitly known mental rules»<sup>80</sup>. Infatti, convenzioni (sul significato dei termini), schemi (per la costruzione di enunciati), regole (di sommarizzazione) sono applicati non attraverso algoritmi, bensì attraverso strategie. Detto in altri termini, essi operano in maniera dinamica, dialettica, nell'ambito di un uso euristico delle conoscenze che formano il contesto dell'indicizzazione e che vengono attivate nel corso dell'analisi concettuale dei documenti. A quest'uso euristico concorrono anche le 'proprietà' del contesto comunicativo, compresa la valutazione della pertinenza dei temi rispetto a determinati interessi informativi. Del resto, lo stesso Frohmann, che alla natura delle regole dell'indicizzazione e alla critica del mentalismo ha dedicato alcuni fortunati saggi<sup>81</sup>, attesta quest'uso euristico delle regole di sommarizzazione, quando riconosce che

We may, for example, agree with Beghtol about the significance to information science of van Dijk's five 'macrorules' [...], but our justification is their utility in deriving subject statements from texts and not [...] their status as hidden rules of our mental processes<sup>82</sup>.

Ma se la regola ha un'applicazione strategica/euristica e non algoritmica, allora essa presuppone anche un contesto di pratiche comunitarie entro cui questa applicazione si attualizzi, come avviene nella costruzione cooperativa del *Thesaurus* del *Nuovo soggettario* e come dovrebbe avvenire tutte le volte che le biblioteche di una stessa rete assegnano stringhe di soggetto a un medesimo documento. Solo così, infatti, le regole assumono «la natura di una tecnica che costituisce lo strumento di una pratica sociale»<sup>83</sup>, la funzione di «instruments of training and instruction [...] instruments used in public life»<sup>84</sup>, e viene messa alla prova la loro effettiva efficacia nel contribuire a determinare una prassi (un uso regolare) e una pratica sociale (un uso condiviso).

Associato alle regole di sommarizzazione viene ravvisato, sia da Frohmann che da Ventura, un altro indizio di 'mentalismo': l'equazione della 'rappresentazione del testo' con la 'rappresentazione della conoscenza'. Esiste questa equazione? Se la rappresentazione di un testo effettuata dall'indicizzatore è la rappresentazione dei temi di cui il testo tratta, ai vari livelli della sua struttura testuale, dai temi particolari

78 R. Ventura, *La biblioteca tra linguaggio e rappresentazione* cit., p. 53.

79 P. Cavaleri, *La biblioteca crea significato* cit., p. 75-76.

80 B. Frohmann, *Rules of indexing* cit., p. 84.

81 Oltre al saggio citato, cfr. di Bernd Frohmann: *The power of images: a discourse analysis of the cognitive viewpoint*, «Journal of documentation», 48 (1992), n. 4, p. 365-386; *An investigation of the semantic bases of some theoretical principles of classification proposed by Austin and the CRG*, «Cataloguing & classification quarterly», 4 (1983/1984), n. 1, p. 11-27. Per una discussione su quest'ultimo contributo di Frohmann, cfr. Diego Maltese; Alberto Petrucciani, *Un'esperienza di indicizzazione per soggetto: materiali per la versione italiana del PRECIS*. Roma: Associazione italiana biblioteche, 1990, p. 120-126.

82 B. Frohmann, *Rules of indexing* cit., p. 95.

83 R. Ventura, *La biblioteca tra linguaggio e rappresentazione* cit., p. 53.

84 B. Frohmann, *Rules of indexing* cit., p. 93.

al tema di base, poiché il tema è una conoscenza metatestuale (questo testo tratta di ...), allora la rappresentazione del testo è sì una conoscenza, ma di natura diversa dalla rappresentazione della conoscenza che il lettore si forma attraverso la fruizione del testo (le conoscenze acquisite con la lettura, che possono incidere sulle sue opinioni e credenze, sulle quali formula giudizi, attribuisce gradi di rilevanza, a partire dalle quali elabora altre conoscenze ecc.).

Lo spazio delle conoscenze metatestuali è anche quello in cui può avvenire un'interazione, forse meno 'fumosa' di quanto ritenuto da Ventura, tra la comprensione dell'indicizzatore e quella dell'utente, tra il modello cognitivo dell'uno e quello dell'altro. E' lo spazio che il catalogo offre all'utente. Infatti, un termine (con il suo corredo di relazioni, informazioni, istruzioni), le stringhe di soggetto di cui è un componente, le notizie bibliografiche collegate a una stringa di soggetto, l'insieme delle stringhe di soggetto di una notizia bibliografica, altri termini correlati al termine di partenza ecc. costituiscono il 'testo' sulla base del quale l'utente può effettuare progressivamente inferenze probabilistiche sui possibili temi di interesse trattati nei documenti, analogamente a quanto successo all'indicizzatore nel costruire questo testo. Ma se il testo costruito dall'indicizzatore non è connesso e coerente, o se una parte di esso rimane nascosta, come si riscontra talvolta nei cataloghi, allora si restringe lo spazio deputato ai processi cognitivi dell'utente, si riduce il grado di certezza delle sue inferenze e l'interazione perde gran parte della sua efficacia, venendole a mancare il terreno sotto i piedi. Quando poi, alla fine del processo, l'utente del catalogo si fa lettore, a quel punto inizia un'altra storia.

Articolo proposto il 23 ottobre 2015 e accettato il 23 novembre 2015.

---

## ABSTRACT

AIB studi, vol. 56 n. 1 (gennaio/aprile 2016), p. 7-39. DOI 10.2426/aibstudi-11407

---

ALBERTO CHETI, 50053 Empoli (FI), e-mail chetialberto@infinito.it.

### **I processi cognitivi nell'analisi concettuale dei documenti: una postilla tra biblioteconomia e linguistica**

L'articolo descrive i principali processi cognitivi che operano nell'ambito dell'analisi concettuale dei documenti, illustrandone la natura e le implicazioni operative. I processi cognitivi descritti consistono in alcuni processi inferenziali basati sull'interazione testo-contesto, e finalizzati all'interpretazione delle parole e delle loro relazioni, all'individuazione della rilevanza dei temi all'interno di un testo, alla valutazione della pertinenza dei temi individuati rispetto a determinati bisogni informativi.

Il modello di analisi concettuale che ne consegue è un modello interattivo, che implica due componenti: a) un processo dal basso (*bottom-up*) o percettivo, che consiste nell'esame dell'informazione contenuta nel documento; b) un processo dall'alto (*top-down*) o cognitivo, che consiste nell'attivazione di vari tipi di conoscenze contestuali (linguistiche, extralinguistiche, intertestuali ecc.) che svolgono una funzione euristica nell'interpretazione delle parole e degli enunciati, nella selezione dei concetti e nella scelta dei termini di indicizzazione.

L'articolo propone anche una riflessione sulle nozioni di 'significato', 'concetto', 'schema', 'sommario' e sulla loro valenza nell'ambito dell'indicizzazione per soggetto, avvalendosi di riferimenti relativi sia agli strumenti di indicizzazione per soggetto sia alla linguistica.

**The cognitive processes in conceptual analysis: a marginal note between library science and linguistics**

The article describes the main cognitive processes involved in document analysis by illustrating their nature and practical implications. These processes consist of a number of inferences based on text-context interaction, and aimed at the interpretation of words (as well as of relationships between words), at understanding the relevant topics within a text, and at identifying their significance in relation to specific information needs.

The conceptual analysis model can be described as an interactive model which involves two components: a) a bottom-up – or perceptual – process, which examines the information within the document; b) a top-down – or cognitive – process, which turns on different kinds of contextual knowledge (linguistic, extra-linguistic, intertextual etc.) essential for understanding words and sentences, for selecting concepts, and choosing indexing terms.

The article also investigates the notions of ‘meaning’, ‘concept’, ‘frame’, and ‘summarization’, and their corresponding significance in subject indexing, with some references to linguistics.