

Biblioteca nazionale centrale di Firenze

Il nuovo Soggettario italiano: dallo studio al progetto *

a cura di

Anna Lucarelli

(BNCF, coordinatrice del Progetto per il rinnovamento del Soggettario)

Leda Bultrini

(Classification and Indexing Section, IFLA)

Alberto Cheti

(Gruppo di progetto per il rinnovamento del Soggettario)

1. Il nuovo Soggettario: i presupposti, il contesto, gli obiettivi del progetto

Il progetto

La Biblioteca nazionale centrale di Firenze e l'Agenzia bibliografica nazionale (BNI) si stanno occupando del rinnovamento del *Soggettario per i cataloghi delle biblioteche italiane*, lo strumento impiegato in Italia nell'indicizzazione semantica dalla maggior parte delle biblioteche.

Il *Soggettario*, pubblicato nel 1956, venne elaborato a partire dalle intestazioni del catalogo a soggetto della Biblioteca nazionale di Firenze; i suoi ideatori si ispirarono alle Subject Headings della Library of Congress, l'unico riferimento di cui i bibliotecari italiani potevano disporre in quel momento.

Il *Soggettario* prevede un'indicizzazione di tipo preordinato, basata su voci principali e suddivisioni. Presenta, in forma di vocabolario controllato, termini in sequenza alfabetica che hanno legami con altri termini sulla base dei loro significati. Le relazioni semantiche individuate non rispondono sempre a criteri uniformi e rigorosi: la Norma ISO 2788 sulla costruzione di tesauri monolingue sarebbe stata pubblicata molti anni dopo, nel 1986.

Nel *Soggettario* voci principali e suddivisioni si trovano di solito già collegate, ma le suddivisioni possono venire estratte da elenchi separati e combinate dal catalogatore per la costruzione delle stringhe di soggetto; in questo senso si dice che esso ha una natura sia enumerativa che sintetica. Non contiene norme esplicite né di tipo sintattico, né semantico, e per la costruzione delle stringhe di soggetto si limita a fornire esempi di possibili combinazioni di

* Questo documento presenta in forma sintetica il progetto sul rinnovamento del *Soggettario per i cataloghi delle biblioteche italiane*, dallo studio di fattibilità concluso nel settembre 2002 agli sviluppi che il lavoro ha avuto sino ad oggi (agosto 2004). Lo studio è stato pubblicato in *Per un nuovo Soggettario. Studio di fattibilità sul rinnovamento del Soggettario per i cataloghi delle biblioteche italiane*, commissionato dalla BNCF alla Ifnet Firenze, realizzato dal Gruppo di progetto per il rinnovamento del Soggettario. Milano, Editrice Bibliografica, 2002.

termini. L'ordine di citazione è stabilito essenzialmente sulla base dell'appartenenza categoriale dei termini; i principi dell'eshaustività e della specificità non sono sempre rispettati.

La terminologia del *Soggettario* nel corso degli anni è stata arricchita tramite *Liste di aggiornamento*, elenchi di termini introdotti dalla Bibliografia nazionale italiana; ma specialmente nei settori disciplinari che si sono maggiormente evoluti e in contesti documentari specialistici lo strumento continua a manifestare la sua inadeguatezza strutturale. Il suo invecchiamento si manifesta sia sul piano terminologico che sintattico.

La Bibliografia nazionale italiana ha adottato il linguaggio del *Soggettario* sin dalla sua nascita. L'agenzia bibliografica ha fornito l'esemplificazione di un metodo di indicizzazione in mancanza di norme precise ed esplicite, e ha rappresentato un riferimento per numerose biblioteche italiane, nonostante la sua prassi abbia subito evoluzioni, ripensamenti, a volte anche difformità.

Nel momento in cui la Biblioteca nazionale centrale di Firenze ha deciso di promuovere un progetto per il rinnovamento del *Soggettario*, ha avuto ben chiaro che il lavoro non avrebbe dovuto soddisfare soltanto le esigenze descrittive e catalografiche della BNI ma offrire un metodo nuovo, moderno e coerente di indicizzazione all'intero e complesso sistema dei servizi informativi e bibliografici italiani, a biblioteche di varia natura, a centri di documentazione che lavorano per mettere a disposizione della propria utenza il patrimonio e la ricchezza delle proprie raccolte. Rispettare questo presupposto significava allora non dare nulla per scontato, analizzare con precisione le esperienze che in questi anni sono state fatte, sia in Italia che a livello internazionale, nel mondo dell'indicizzazione, riconsiderare standard e normative, accogliere le linee emerse dai lavori sull'indicizzazione dell'International Federation of Library Associations and Institutions (IFLA), i requisiti individuati da FRBR per favorire le esigenze informative dell'utenza, le evoluzioni di altri linguaggi di indicizzazione.

Significava anche recepire quanto prodotto in Italia dall'attività del Gruppo di ricerca sull'indicizzazione per soggetto (GRIS) dell'AIB, le riflessioni elaborate all'interno della BNI, le esperienze dei poli del Servizio bibliotecario nazionale (SBN), i lavori di altre biblioteche, gli studi più recenti sulla terminologia e la costruzione di tesauri.

Per garantire questi presupposti, il lavoro ha preso avvio da uno studio di fattibilità il cui scopo è stato quello di produrre un progetto preliminare in cui, fra l'altro, fosse definito un preventivo di spesa sulla base del quale il Ministero per i beni e le attività culturali potesse finanziare l'effettiva realizzazione del lavoro.¹

La Biblioteca nel dare avvio allo studio ha inteso, come è stato più volte detto, "partire dalle fondamenta". Si è chiesta se mantenere il sistema attuale, con qualche aggiornamento oppure abbandonarlo in favore di un sistema diverso; se rinnovarlo profondamente, pur conservandone gli elementi tuttora validi; a quali esperienze straniere riferirsi; come rapportarsi alla tradizione delle *Library of Congress subject headings* (LCSH), ispiratrici del vecchio *Soggettario* e ampiamente diffuse soprattutto in alcuni ambiti geografici; quali altri linguaggi di indicizzazione avere come riferimento in funzione di eventuali progetti di cooperazione internazionale.

¹ Lo studio è stato affidato ad un gruppo di consulenti esterni, coordinato da Luigi Crocetti, e composto da Alberto Cheti, Daniele Danesi, Massimo Rolle e Stefano Tartaglia. La Biblioteca si è avvalsa della collaborazione di Carlo Revelli e del contributo scientifico di Diego Maltese, mentre per l'Istituto centrale per il catalogo unico e per le informazioni bibliografiche (ICCU) ha partecipato ad alcuni incontri Maria Lucia Di Geso. Il coordinamento del progetto è affidato dalla Biblioteca nazionale ad Anna Lucarelli che con Marta Ricci segue sin dall'inizio lo svolgimento dei lavori.

E' stata fatta la scelta del rinnovamento, con l'obiettivo di raggiungere un modello di indicizzazione più evoluto: il nuovo *Soggettario* non sarà l'adattamento di strumenti impiegati in altri contesti, né la traduzione italiana di un thesaurus in altra lingua.

Come verrà spiegato nelle pagine che seguono, terrà conto dei più recenti indirizzi elaborati in campo internazionale, ma allo stesso tempo sarà legato al contesto nazionale a cui è destinato.²

Il progetto si è posto con chiarezza alcuni obiettivi di base:

- un lavoro innovativo e non un semplice restauro
- la considerazione delle componenti migliori della nostra tradizione catalogafica;
- un linguaggio conforme agli standard internazionali;
- un sistema fondato su regole chiare e su strutture coerenti, adeguate a controllare l'evoluzione del linguaggio;
- uno strumento flessibile e modulare, destinato sia a contesti informativi generali che specializzati, capace di rappresentare soggetti di diverso livello di complessità;
- un linguaggio d'indicizzazione semplice da insegnare e da apprendere;
- un linguaggio condivisibile in un sistema integrato e cooperativo.

Lo studio di fattibilità

Lo studio di fattibilità sul rinnovamento del *Soggettario* è iniziato nel mese di settembre del 2000 e si è concluso nel settembre del 2002. E' stato pubblicato nel già citato volume che ne documenta i passaggi e l'evoluzione.

Lo studio ha permesso di:

1. indicare i principi e le caratteristiche del nuovo linguaggio di indicizzazione
2. definire modalità per raggiungere obiettivi realizzabili e soluzioni economicamente sostenibili;
3. elaborare un piano a breve e medio termine, un avanzamento per tappe;
4. calcolare un preventivo di spesa relativo alle risorse umane e tecnologiche necessarie alla realizzazione del progetto.

Sullo studio di fattibilità sono state date notizie in vari sedi, a cui si rinvia per informazioni di varia natura.³ La principale articolazione dello studio si può così sintetizzare:

² Principi comuni relativi a sistemi di indicizzazione diversi, vengono presentati nell'indagine IFLA su *Principles underlying subject heading languages. (SHLs)*, edit by Maria Ines Lopes and Julianne Beall. München, Saur, 1999.

³ A. LUCARELLI, *La Biblioteca nazionale centrale di Firenze e il rinnovamento dell'indicizzazione per soggetto*. «Bollettino AIB», 41 (2001) n. 3, p. 307-313; M. ROLLE, *Per un nuovo modello di linguaggio documentario*. «Bollettino AIB», 41 (2001) n. 3, p. 315-326; L. BULTRINI, *Subject indexing and Classification in Italy, 2001-2002*, «Newsletter. Classification and Indexing Section. Division of Bibliographic Control. IFLA», November 2002, nr. 26, p. 16-17, <http://www.ifla.org/VII/s29/pubs/ci26.pdf>; ALBERTO CHETI, *Il punto sul GRIS e gli sviluppi attuali*, «Bibliotime», 6 (2003), n.1, <http://www.spbo.unibo.it/bibliotime/num-vi-1/cheti.htm>; A. LUCARELLI, *La revisione del Soggettario*, «Bibliotime», 6 (2003), n.1, <http://www.spbo.unibo.it/bibliotime/num-vi-1/lucarell.htm>

- una fase preliminare, in cui è stato approfondito il quadro teorico e metodologico entro cui sviluppare la riflessione sul linguaggio del nuovo Soggettario. E' stata condotta un'analisi dei principali strumenti in uso, sono stati messi a punto i riferimenti culturali e scientifici del lavoro per arrivare ad una visione unitaria dei problemi dell'indicizzazione semantica. Particolare attenzione è stata rivolta all'analisi categoriale e all'organizzazione concettuale, ai modelli che ne hanno stabilito metodologie e standard. Sono stati analizzati i problemi della ricerca per soggetto e i requisiti che deve avere un sistema di indicizzazione per rendere efficace l'accesso ai documenti sulla base del loro soggetto e, in funzione di questo, il grado di efficienza di linguaggi preordinati e postordinati, la possibilità di distinguere la sfera semantica da quella sintattica.

- Una fase intermedia in cui sono state definite "Proposte" sulle caratteristiche del nuovo sistema. Le proposte riguardavano la precisione del linguaggio (sia da un punto di vista terminologico che sintattico), le sue unità di base (i termini e le stringhe), le modalità di controllo delle relazioni semantiche e sintattiche e relativamente all'ordine di citazione nelle stringhe, le strutture delle voci del vocabolario. Gli argomenti delle proposte sono stati discussi da esperti italiani e stranieri durante un seminario tecnico tenuto a Firenze nell'aprile del 2001.⁴

- A seguito delle riflessioni emerse durante il seminario, il lavoro si è orientato alla definizione delle strutture e delle norme del nuovo Soggettario. Lo studio di fattibilità non ha prodotto un vero e proprio testo di norme (prodotto che verrà realizzato nella fase esecutiva del lavoro), ma una guida, un insieme articolato di raccomandazioni utili alla sua elaborazione.

- Le ipotesi individuate sono state verificate con una sperimentazione di tipo terminologico su circa 500 termini e con una prima sperimentazione sulla componente sintattica del nuovo linguaggio, con particolare attenzione a quella che è stata definita "nota sintattica" o "nota applicativa", un insieme di istruzioni che corredano la struttura di alcuni termini del vocabolario.

- Lo studio si è concluso nel giugno del 2002 con la stesura di un progetto preliminare e la messa a punto di aspetti organizzativi. Sono state affrontate questioni gestionali relativamente allo smantellamento del vecchio sistema, alla definizione dei costi, e a modalità per realizzare il lavoro secondo piani a breve e medio termine. Il nuovo Soggettario dovrà avere costi di gestione e manutenzione contenuti ed essere competitivo rispetto a sistemi di ricerca automatici e post-coordinati che, benché molto poco efficaci dal punto di vista della ricerca, appaiono vincenti sul piano dell'economicità.

Stato dell'arte e nuove prospettive

Nel periodo l'anno che ha fatto seguito alla conclusione dello studio di fattibilità, in attesa che il Ministero finanziasse l'avvio del progetto, il lavoro è continuato in varie direzioni:

➤ La Bibliografia nazionale italiana ha sperimentato il sistema nuovo Soggettario, in particolare relativamente a questioni inerenti

a) le Categorie semantiche da attribuire ai termini del Thesaurus.

⁴ Al seminario di Firenze hanno partecipato: Lynn El Hoshy della Library of Congress, Mary Dykstra esperta di Precis, Max Naudi della Bibliothèque nationale de France, Marion Sedelmayer della Bayerische Staatsbibliothek, Maria Ines Cordeiro della Fundação Calouste Gulbenkian, Lisboa, Patrice Landry della Schweizerische Landesbibliothek; fra gli esperti italiani: Mauro Guerrini, Alberto Petrucciani, Riccardo Ridi, Paul Gabriele Weston.

b) alcuni aspetti sintattici.

➤ Alcuni principi terminologici del progetto nuovo Soggettario sono stati applicati alla revisione del lessico di una biblioteca italiana specializzata in enologia e viticoltura: la Biblioteca internazionale La Vigna di Vicenza. Il lavoro, finanziato dalla Regione Veneto, ha fatto emergere questioni interessanti, in particolare riguardo alle modalità di cumulazione di terminologie tecniche e settoriali già strutturate in un Thesaurus generale.

➤ La Bibliografia nazionale ha inoltre stabilito rapporti di collaborazione con gruppi che si occupano di terminologia sia in ambito generale che specialistico, e con esperti di indicizzazione di materiali vari (risorse elettroniche, audiovisivi, documenti archivistici, fotografie, ecc.) per verificare l'applicabilità del nuovo linguaggio anche a questi settori.

La Biblioteca nazionale centrale di Firenze sarà impegnata ad approfondire il suo impegno su vari versanti come: la componente informatica del sistema; l'ottimizzazione del nuovo linguaggio nell'OPAC; l'impiego del formato MARC per la strutturazione dei dati; il tipo di visualizzazione dei dispositivi del Thesaurus e le strategie di navigazione interna; lo sviluppo del sistema in una direzione multilingue (il Thesaurus italiano, infatti, se strutturato rigorosamente, potrà comunicare o interagire con strumenti in altre lingue predisposti sulla base dei medesimi standard); il raccordo fra il nuovo Soggettario e un eventuale sistema nazionale di Authority control ancora non realizzato.

Dovranno inoltre essere definite con precisione le modalità con cui il nuovo linguaggio potrà essere condiviso. La Bibliografia nazionale italiana, se potrà disporre di risorse adeguate, offrirà un prototipo del nuovo sistema, esemplificandone l'impiego a beneficio delle biblioteche che condividono risorse nel nostro Servizio bibliotecario nazionale (SBN). La BNI, in collaborazione con l'Istituto centrale per il catalogo unico e le informazioni bibliografiche (ICCU), potrà gestire l'implementazione on-line dello strumento, raccogliere proposte di aggiornamento del vocabolario, validare stringhe costruite secondo una sintassi coerente.

Riguardo alla componente informatica del sistema, è importante che il software di gestione del vocabolario, consenta soluzioni avanzate, efficienti e semplici al tempo stesso. Si dovranno potenziare gli aspetti di interoperabilità tecnica del nuovo strumento, nell'ottica di un'apertura al web. Lo strumento dovrà rispondere a protocolli standard, presentandosi anche dal punto di vista tecnico come un servizio ed essere integrato da altre applicazioni fondate su software di catalogazione di vario tipo.

2. Le scelte metodologiche

Il modello analitico sintetico

Guardando alle caratteristiche principali del linguaggio indicate nel progetto, che saranno illustrate nel capitolo successivo (normalizzazione, insegnabilità, flessibilità e modularità), si è individuato il riferimento adatto per la loro realizzazione nel modello analitico-sintetico, che consente la produzione di stringhe di soggetto a partire da termini di vocabolario, combinati fra loro secondo le regole della sintassi.

La prima scelta operata è stata, dunque, quella fra linguaggi di tipo enumerativo, nei quali le formulazioni di soggetto sono solo quelle specificate in una lista d'autorità (secondo il modello dei soggettari), e linguaggi sintetici, nei quali ciascun indicizzatore può produrre la stringa di soggetto che meglio rappresenta il tema del documento che sta indicizzando, a partire dai termini di un vocabolario controllato e da un corredo di norme sintattiche che lo guidano nella loro combinazione.

Sono prevalentemente enumerative le LCSH e lo è, appunto, il *Soggettario*, sebbene l'uno e le altre nel corso del tempo abbiano teso ad incorporare sempre di più elementi di sintesi. Sono sintetici i linguaggi a stringa come il PRECIS e come quello proposto dalla *Guida* GRIS. Si colloca sul versante dei linguaggi sintetici, sia pure in maniera non altrettanto rigorosa, il portoghese SIPORbase, a conferma di una tendenza non solo dottrinale.

Il modello analitico-sintetico distingue decisamente il piano della semantica da quello della sintassi. Al piano semantico appartengono le relazioni fra i termini, in quanto rappresentativi di concetti singoli, considerati ciascuno di per sé, entro un orizzonte di significati *a priori*, indipendentemente dal loro comparire all'interno di una combinazione di concetti in relazione logico-funzionale fra loro. Al piano della sintassi appartengono le relazioni funzionali che legano i concetti nel contesto di una combinazione logica tra loro, come avviene all'interno di una formulazione di soggetto.

La distinzione fra i due piani, già enunciata nella norma ISO 2788-1986 sulla costruzione dei tesauri monolingue, è fissata con chiarezza nel documento dell'IFLA sui *Principles underlying subject heading languages*⁵ nel quale si dice che le parti che compongono un'intestazione di soggetto dovrebbero essere legate da relazioni sintagmatiche e non semantiche (o paradigmatiche) e che la struttura semantica di un linguaggio di indicizzazione si esprime attraverso relazioni di equivalenza, gerarchiche e associative.

L'indicizzazione semantica ha a fondamento la gestione e il controllo dei due tipi di relazioni, semantiche e sintattiche.

Nel modello analitico-sintetico adottato il controllo delle relazioni semantiche avviene attraverso la costruzione di un vocabolario *strutturato* di termini, nel quale le relazioni di significato possibili sono ridotte convenzionalmente alle tre codificate dalla norma ISO (di equivalenza, gerarchiche, associative), che vanno a costituire il corredo semantico dei termini.

Il controllo delle relazioni sintattiche individua i rapporti logici esistenti fra i concetti all'interno dei concreti enunciati di soggetto (e della lingua in genere) e, attraverso la definizione

⁵ *Principles underlying subject heading languages (SHLs)*, cit. p.2.

di norme di costruzione (ordine di citazione), ne consente la riproduzione e la leggibilità all'interno di una *stringa di soggetto*.

Termini e stringhe vs. voci principali e suddivisioni

Il modello analitico-sintetico, dunque, fa sì che si possa lavorare con unità semplici (i concetti singoli o unitari e i termini che li rappresentano) e con un numero limitato di norme di costruzione per esprimere tutti gli enunciati, comunque complessi, che la varietà e ricchezza dei documenti da indicizzare producono.

Viene, dunque, abbandonata l'articolazione del linguaggio in *voce principale*, parola d'ordine, prima voce del soggetto, e quindi anche punto d'accesso e sostanziale espressione del contenuto del documento, e *suddivisioni*, con funzioni complete, che compaiono al seguito della voce principale nelle voci già costruite di un linguaggio prevalentemente enumerativo, il quale prevede solo una limitata libertà di costruzione, costituita dalla possibilità di aggiungere suddivisioni "libere" alle *voci* enumerate.

Parallelamente non ha più ragion d'essere la ricerca dei criteri di individuazione della voce principale in quanto elemento di maggiore importanza, intesa nel senso di maggiore *significatività*, misurata sull'appartenenza a categorie semantiche o tipi concettuali (entità vs. attività; nomi concreti vs. nomi astratti, cose/materiali/azioni...), criteri che entrano in crisi quando nella definizione del contenuto concettuale di un documento si trovano a coesistere più concetti appartenenti alla medesima categoria (da cui l'espedito della moltiplicazione del numero delle voci assegnate al singolo documento, nessuna delle quali rappresentativa del suo contenuto nella sua interezza, o l'applicazione di ulteriori, più discutibili criteri di comparazione di specificità fra concetti appartenenti a gerarchie semantiche differenti e quindi non confrontabili sotto il profilo dell'estensione del significato).

Il sistema combinatorio del modello analitico-sintetico pone tutti i concetti, e i termini che li esprimono, sullo stesso piano, ne controlla la forma, unica, indipendentemente dall'uso che se ne farà e dalla posizione che andranno ad assumere all'interno delle singole stringhe di soggetto, ne definisce il significato attraverso la struttura relazionale del vocabolario nel suo complesso, collocandoli all'interno di una rete di collegamenti semantici. Di quei termini si serve per costruire le stringhe di soggetto, individuando il ruolo logico che i concetti che essi esprimono svolgono, gli uni rispetto agli altri, all'interno del singolo enunciato di soggetto e ricostruendolo all'interno della stringa mediante una disposizione dei termini che è regolata da norme sintattiche, anch'esse di valore generale.

Specificità e coestensione

L'articolazione del linguaggio nei due piani distinti, semantico e sintattico, consente di dare piena attualizzazione al principio base dell'indicizzazione semantica, che è quello della specificità, nella sua più propria connotazione relazionale.

L'esigenza di creare accessi ai documenti tramite intestazioni specifiche attraversa, infatti, la storia dell'indicizzazione moderna fin dalle sue origini; ma nel tempo la specificità si è venuta

meglio definendo come una proprietà da declinare secondo le diverse unità linguistiche implicate dal processo di indicizzazione.

La *specificità* si realizza, innanzitutto, a livello delle unità di base, i termini, mediante la precisione con cui un concetto è espresso dal termine che lo rappresenta. Per ogni concetto deve essere impiegato il termine che lo esprime con maggiore esattezza.

La specificità deve essere garantita anche alla rappresentazione dei legami logici fra i concetti che esprimono il contenuto concettuale del documento: la sintassi del linguaggio di indicizzazione deve consentire di rappresentare, in modo chiaro e non ambiguo, tutte le relazioni logiche, le quali devono poter trovare espressione all'interno di una stessa stringa, che dia conto del tema del documento, senza che questo venga artificialmente spezzato in più rappresentazioni parziali. A livello delle stringhe di soggetto, dunque, la specificità si realizza più propriamente come *coestensione* della stringa con il contenuto del documento espresso dall'enunciato di soggetto.

Evidentemente, è necessario che l'enunciato di soggetto sia, a sua volta, un'espressione precisa del contenuto concettuale del documento, vale a dire che ne esprima il tema di base, individuando e dando conto di tutti i concetti essenziali a descriverne il soggetto complessivo. Questa forma di specificità, intesa come *esaustività* degli enunciati di soggetto, è il risultato di una corretta analisi concettuale ed è strettamente legata alla nozione di "soggetto" come "tema di base" del documento, definito non in base ad una struttura classificatoria preesistente al documento trattato, ma esclusivamente in rapporto al concreto contenuto di quel documento ed espresso in forma descrittiva e non attraverso l'individuazione di un "nome" ammesso in una lista di autorità di nomi di soggetti.

È l'idea di Coates che un argomento trattato in un documento possa essere espresso in una (sola) intestazione di soggetto, o, dall'angolazione opposta, che qualunque testo che si presenti come un'unità coerente abbia come proprietà necessaria un soggetto unitario. Queste idee sono state sviluppate nella "Guida GRIS" con l'individuazione di criteri (l'implicazione e la coerenza), che definiscono le condizioni logiche attraverso le quali si può stabilire la pertinenza di un documento ad un certo soggetto: un tema generale è implicato nell'insieme dei temi particolari da cui è ricavato, ogni tema particolare deve essere coerente con il tema generale al quale è riconducibile.

Specificità, coestensione ed esaustività sono la garanzia dell'efficacia e dell'efficienza del linguaggio, sul piano della sua precisione, quale strumento di ricerca: qualità irrinunciabile in un'epoca caratterizzata dalla mole incommensurabile di documenti virtualmente disponibili e difficilmente ritracciabili e dalla creazione ed interconnessione di basi dati bibliografiche sempre più vaste.

L'analisi categoriale per il controllo delle relazioni semantiche

L'altro parametro di funzionalità, la capacità di richiamo, è affidato invece (oltre che alla *predittività* in genere del linguaggio, vale a dire al suo essere fondato su norme certe ed esplicite, applicabili con coerenza) al piano della correlazione semantica tra i termini nel vocabolario.

Per il controllo delle relazioni semantiche, nel nuovo Soggettario si utilizza lo strumento dell'analisi categoriale, che consiste nell'individuare una serie di tipi concettuali fondamentali (categorie quali quelle di attività, materiali, oggetti, organizzazioni, persone, processi, strumenti

ecc.), che vanno a costituire le macrostrutture semantiche del vocabolario, e nel valutare l'appartenenza dei singoli concetti alle categorie individuate, a partire dal loro contenuto semantico di base, ossia dalle caratteristiche essenziali presenti nella definizione del significato base del concetto (principio della *definizione tipica*), utilizzando queste caratteristiche come criterio di riferimento per il riconoscimento dell'appartenenza categoriale. Dunque, tutti i concetti appartenenti ad una determinata categoria semantica presentano una stessa caratteristica fondamentale nel loro significato di base. Con criteri analoghi si individuano anche i principi di suddivisione di ciascuna categoria, espressi dalle *faccette*.

Nel modello a tre relazioni (di equivalenza, gerarchiche e associative) definito dalla norma ISO 2788, l'analisi categoriale è, dunque, lo strumento attraverso il quale si costruisce la struttura portante del vocabolario, costituita dalle relazioni *gerarchiche* (con particolare rilevanza di quella più tipica e di più ampia applicazione, la relazione generica). L'appartenenza dei termini ad una gerarchia è regolata anzitutto dall'appartenenza ad una medesima categoria semantica.

L'analisi categoriale, così intesa con riferimento alle *categorie semantiche*, supporta anche la scelta della morfologia del termine, soprattutto per quanto riguarda l'alternativa fra forma singolare e forma plurale, previste entrambe come possibilità dal vocabolario al fine di garantire la massima espressività del linguaggio. Il criterio della *numerabilità*, per cui la forma plurale è utilizzata per esprimere i concetti numerabili è applicabile, infatti, anzitutto alla categoria semantica in sé e quindi, secondo il criterio gerarchico, a tutti i termini che esprimono concetti riconducibili a quella categoria. È sempre l'analisi categoriale che scioglie l'interrogativo sulla possibilità di prevedere entrambe le forme di uno stesso termine, restringendola al caso in cui ci si trovi di fronte ad uno stesso lemma che nelle due forme rappresenti concetti appartenenti a due categorie semantiche distinte (criterio dello *scostamento categoriale*).

Ma il modello dell'analisi categoriale, sempre nell'ambito del controllo terminologico, non si applica solo alle categorie semantiche, ma anche alle categorie intese come funzioni logiche, cui si fa ricorso nei criteri di scomposizione dei termini composti, nel rispetto della logica di fondo della norma ISO 2788 che raccomanda di preferire termini che esprimono concetti semplici o unitari.

L'analisi dei ruoli per il controllo delle relazioni sintattiche

Il metodo della categorizzazione dei concetti favorisce l'applicabilità, l'omogeneità e la semplicità delle norme. Così, anche sul piano delle relazioni sintattiche, la categorizzazione presiede ai criteri per l'ordine di citazione dei termini nelle stringhe di soggetto. Un sistema di categorie logiche generali, i "ruoli", identificano infatti in maniera esaustiva le funzioni che i concetti possono svolgere nei soggetti.

L'ordine di citazione degli elementi di una stringa di soggetto, che nel linguaggio del vecchio *Soggettario* veniva stabilito secondo categorie lessicali (nomi propri/nomi comuni) e semantiche (entità/attività, concreto/astratto, ...), ancorate al criterio dell'importanza o significatività della prima voce, nel nuovo *Soggettario* si fonderà su categorie funzionali (azione/oggetto, azione/agente, intero/parte ecc.). Pertanto, la collocazione dei termini nella stringa dipenderà dal tipo di relazioni logiche che legano i concetti che compongono l'enunciato di soggetto.

Ne risulta un ordine di citazione esplicito e stabile, fondato su poche norme di validità generale (principio della relazione uno a uno e principio della dipendenza logica), attraverso il quale dare concretezza, anche sul piano sintattico, al principio dell'unicità della forma di rappresentazione del soggetto, garantendo il rispetto dei requisiti di *univocità* (le relazioni logiche debbono essere espresse in forma non ambigua) e di *uniformità* (una certa relazione logica deve essere espressa mediante un ordine di citazione codificato e costante), concorrenti, rispettivamente, alla realizzazione delle proprietà di precisione e richiamo.

3. L'architettura del nuovo Soggettario

I componenti

Conforme al modello analitico-sintetico, rispondente ad esigenze di flessibilità e di adattabilità ai diversi contesti documentari e in linea con lo sviluppo delle tecnologie, il nuovo Soggettario si configura come un sistema integrato, costituito da sottosistemi distinti e complementari: un insieme di componenti progettati per poter operare insieme al fine di supportare il processo di indicizzazione.

Il primo componente è costituito dalle *norme*, la grammatica del linguaggio documentario: norme esplicite ed elaborate a priori nella fase di implementazione del linguaggio, non implicite o derivate dall'uso prevalente; norme coerenti e semplici, essendo queste due condizioni necessarie per l'insegnabilità del linguaggio e per una sua estesa applicazione; norme organiche, ossia riguardanti tutte e tre le fasi del processo di indicizzazione.

Il secondo componente è costituito dal *vocabolario*: un vocabolario controllato e strutturato, al quale fare riferimento per la scelta dei termini di indicizzazione e per l'individuazione delle relazioni semantiche fra i termini stessi. Il vocabolario contiene anche altre informazioni, a carattere storico, gestionale, applicativo.

Il terzo componente è il cosiddetto *corredo sintattico-applicativo*. I termini del vocabolario possono essere collegati a istruzioni particolari, che hanno lo scopo di chiarirne l'impiego nell'indicizzazione dei documenti, facilitando l'applicazione delle regole generali. Queste istruzioni prendono la forma di *note sintattiche* o di *voci di manuale*, a seconda che riguardino singoli termini del vocabolario (p.e., **Contratti**, **Discipline**, **Malattie**), oppure aree semantiche generali (p.e., arte, letteratura). Tuttavia, nello studio di fattibilità è stata sviluppata solo la struttura della nota sintattica. Queste istruzioni, oltre a consentire l'applicazione delle norme generali, facilitano anche il passaggio dalla struttura sintattica precedente a quella del nuovo linguaggio documentario.

L'ultimo componente è l'*archivio delle stringhe di soggetto*. Un repertorio di stringhe non rappresenta più, in un linguaggio sintetico, l'elemento principale del sistema. Infatti, come nella lingua naturale le frasi si costruiscono sulla base della grammatica e del lessico, così nel linguaggio documentario sono le norme e il vocabolario gli strumenti fondamentali per costruire le stringhe con cui rappresentare il contenuto concettuale dei documenti. Tuttavia, le stringhe di soggetto prodotte dalla BNI nell'indicizzazione dei documenti possono costituire un'esemplificazione autorevole ed estesa del nuovo linguaggio documentario. Questo archivio nazionale delle stringhe di soggetto sarà implementato sulla base delle procedure di cooperazione; l'Agenzia bibliografica nazionale avrà il compito di validare tali stringhe.

Le interrelazioni

Le interrelazioni tra questi quattro componenti sono illustrate nella *Fig. 1*. Le norme presiedono alla costruzione sia del vocabolario che delle stringhe di soggetto: la forma dei termini, le loro relazioni semantiche, il loro ordine di citazione nelle stringhe sono stabiliti in

conformità alle regole generali del linguaggio di indicizzazione. Le informazioni di carattere storico e applicativo contenute nel vocabolario e, in particolare, le istruzioni rappresentate dalle note sintattiche e dalle voci del manuale facilitano l'applicazione delle regole generali. Queste interrelazioni funzionali sono indicate con le linee tratteggiate. Le linee continue indicano, invece, i collegamenti che si hanno, da un lato, tra i termini del vocabolario e le stringhe nelle quali sono presenti, dall'altro, tra i termini del vocabolario e le corrispondenti note sintattiche o le voci del manuale che possono riguardarli.

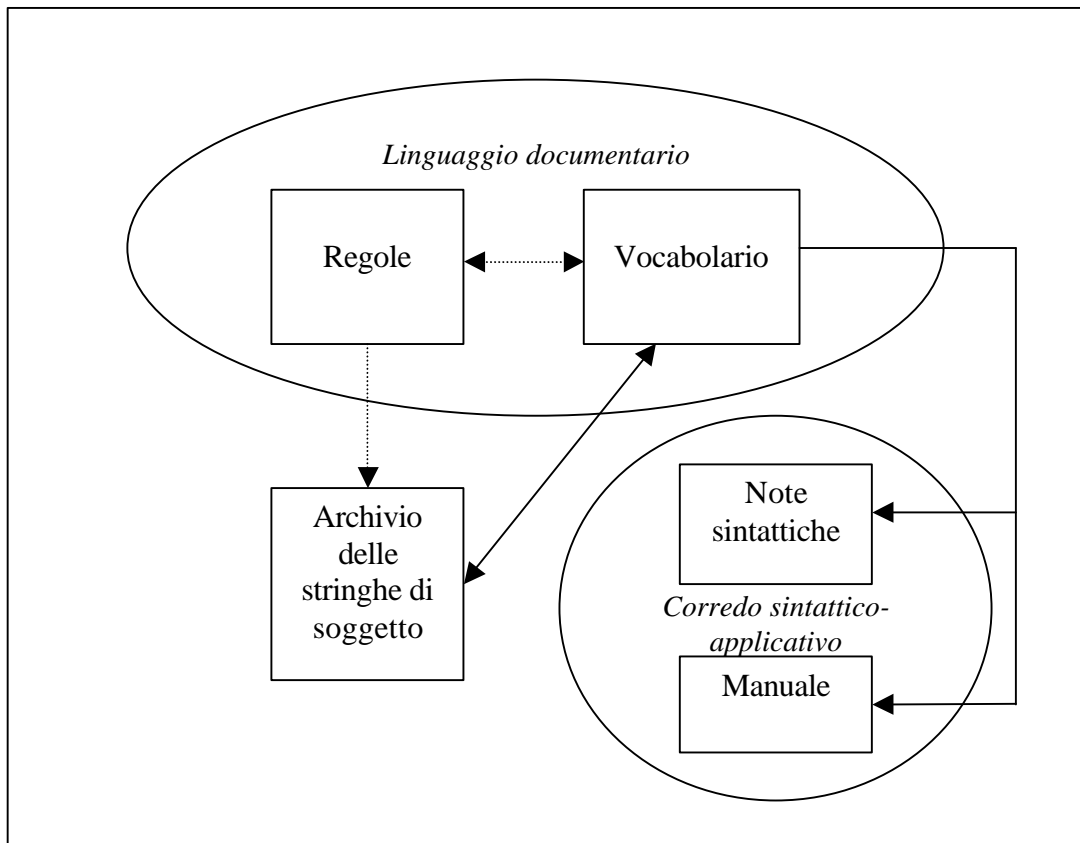


Fig. 1: Componenti del Sistema Nuovo Soggettario e loro relazioni

Il *vocabolario* costituisce la base terminologica del linguaggio documentario; il vocabolario e le *norme* costituiscono il linguaggio documentario vero e proprio; il linguaggio documentario, il *corredo sintattico-applicativo* e l'*archivio delle stringhe di soggetto* costituiscono il sistema di indicizzazione nel suo insieme. Dunque, quattro componenti, ma non sullo stesso piano: mentre le norme e il vocabolario, come elementi costitutivi del linguaggio documentario, sono requisiti indispensabili per l'impiego del sistema, gli altri due componenti, il corredo sintattico-applicativo e l'archivio delle stringhe di soggetto, sono elementi complementari, di supporto, di aiuto agli indicizzatori.

Le funzioni per l'utente

Questi componenti hanno anche ruoli distinti nello svolgimento delle funzioni per l'utente. Da questo punto di vista, due, in particolare, sono i componenti rilevanti: il vocabolario e l'archivio delle stringhe di soggetto.

Il primo, che dà visibilità ai termini di indicizzazione e alle loro relazioni semantiche, svolge la prima delle quattro funzioni utente indicate in FRBR: trovare un'entità. Un termine di indicizzazione è usato tipicamente come termine di ricerca per trovare un'entità, ossia il soggetto che interessa, oppure per indirizzare l'utente ad un'altra entità strettamente correlata, ossia un termine equivalente o un concetto sovraordinato, subordinato, associato.

L'archivio delle stringhe di soggetto svolge, invece, le funzioni di identificare e selezionare un'entità:

- a) ogni stringa di soggetto permette di identificare un soggetto;
- b) le stringhe di soggetto collegate ad uno stesso termine di ricerca permettono di identificare il soggetto che interessa tra un insieme di soggetti che presentano almeno una caratteristica (un concetto) in comune;
- c) la stringa di soggetto scelta permette di selezionare le opere attinenti a quel soggetto.

Questa modalità di ricerca è detta anche ricerca "a due stadi", nella quale le stringhe di soggetto fanno da "filtro", da "ponte" tra i singoli concetti e il contenuto concettuale dei documenti, di cui le stringhe sono indicatori significativi.

Modularità e flessibilità del sistema

La macrostruttura progettata per il nuovo Soggettario è caratterizzata da una modularità che lo rende flessibile sotto molti aspetti.

La prima forma di flessibilità è l'adattabilità del linguaggio alla rappresentazione di soggetti di diverso livello di complessità. Si richiede oggi di poter esprimere tutta la varietà e la complessità dei soggetti, di rappresentare cioè non più un "soggetto referenziale" (mediante un nome), ma un "soggetto descrittivo" (mediante una combinazione di nomi). La precisione, come è noto, è il parametro dell'efficienza del catalogo. Occorrono, dunque, stringhe "su misura": termini semplici e un'elevata capacità combinatoria. E questo è ottenibile non mediante un linguaggio fatto di espressioni preconfezionate, di formule rigide, che pretendono di rappresentare enumerativamente l'universo concettuale dei documenti, ma di un linguaggio sintetico, generativo, le cui espressioni (stringhe di soggetto) si possono ottenere mediante la combinazione dei termini del vocabolario secondo le regole della sintassi.

Un'altra forma di flessibilità riguarda l'applicabilità del sistema in contesti documentari diversi (generalisti, specializzati, preordinati, postordinati) e a risorse informative diverse (p.e., risorse elettroniche). Insomma, non esiste un "soggettario" unico, valido per tutte le biblioteche, se per "soggettario" s'intende l'elenco delle "voci" impiegate nell'indicizzazione dei documenti. Ogni biblioteca si costruisce il suo archivio di stringhe di soggetto, graduando esaustività, coestensione e specificità secondo le proprie esigenze, seppure adottando un linguaggio comune, fatto di regole condivise e di un vocabolario di riferimento. Non solo, essendo le due strutture, quella semantica e quella sintattica, integrate nel sistema come strutture

distinte, queste sono all'occorrenza separabili, rendendo il sistema adattabile anche ad un'indicizzazione di tipo post-coordinato.

Dunque, anche se lo studio di fattibilità era prioritariamente finalizzato al rinnovamento e alla trasformazione del *Soggettario*, tuttavia lo strumento prodotto ha caratteristiche che ne permettono l'impiego in un contesto più ampio, e cioè:

1. in tutte quelle biblioteche che tradizionalmente fanno riferimento alla BNI e al *Soggettario* come fonte di terminologia e come modello di indicizzazione per i loro cataloghi, impiegando e adattando la terminologia e il metodo anche alla produzione bibliografica non italiana;
2. in generale, nell'insieme delle reti e biblioteche che partecipano a SBN;
3. in tutte quelle realtà, bibliotecarie e documentarie, nelle quali si scelga di utilizzare un linguaggio controllato e strutturato pienamente aderente a principi e standard internazionali, per cataloghi, indici e altri strumenti di accesso all'informazione, qualunque sia il tipo e il supporto del documento;
4. in tutti i contesti di indicizzazione che usino un vocabolario controllato e strutturato, sia preordinati che postordinati.

Infine, la modularità può essere intesa anche come gradualità nella costruzione del sistema di indicizzazione: la distinzione logica tra i diversi componenti del sistema comporta anche una loro distinta realizzazione, in modo da ottenere una scalabilità dal punto di vista economico e temporale dell'attuazione del progetto del nuovo *Soggettario*.

4. Le norme

Riferimenti

Lo studio di fattibilità non contiene il testo delle norme, bensì un insieme di raccomandazioni, suggerimenti, opzioni che possano servire da guida per l'elaborazione delle norme. Il loro principale riferimento è costituito dalla *Guida all'indicizzazione per soggetto* del GRIS: le norme in essa contenute sono, infatti, l'espressione più organica e autorevole di un modello di indicizzazione del tutto omogeneo con quello proposto per il nuovo Soggettario. La *Guida* e il lavoro di ricerca del GRIS costituiscono, nel loro complesso, il presupposto teorico e metodologico, la fonte ispiratrice del nuovo Soggettario. Se ne possono individuare alcuni elementi essenziali, fondamentali, sul piano normativo, posti a base anche del nuovo sistema di indicizzazione:

1. l'aderenza ai due standard internazionali: ISO 5963-1985, per l'analisi concettuale dei documenti, e ISO 2788-1986, per il controllo terminologico;
2. la riformulazione dei principi generali dell'indicizzazione: *uniformità e univocità, esaustività, coestensione, specificità*;
3. la definizione di "soggetto" come *tema di base* del documento;
4. l'adozione dell'*analisi categoriale* per la strutturazione del vocabolario;
5. l'adozione dell'*analisi dei ruoli* per la costruzione delle stringhe di soggetto;
6. una politica di indicizzazione orientata alla "stringa unica coestesa".

Raccomandazioni generali

Le raccomandazioni contenute nello studio si pongono, in particolare, tre obiettivi generali, riconducibili ad altrettanti requisiti fondamentali:

- a) definire norme di validità generale: le regole che formano il contenuto della guida normativa devono avere un'applicabilità generale, ossia riferirsi al linguaggio nel suo insieme; mentre, le istruzioni relative a casi particolari, rappresentati da singoli termini o da categorie di termini, sono affidate alle note sintattiche e alle voci del manuale;
- b) preservare l'omogeneità dei criteri adottati: il ricorso ad eccezioni o a criteri di natura differente comporta, infatti, due conseguenze negative, una minore coerenza e una maggiore complessità delle norme;
- c) semplificare le norme: il compito più difficile per chi si appresta a progettare un linguaggio di indicizzazione è quello di perseguire la semplicità e la chiarezza delle norme, in modo da facilitarne la comprensione e l'applicazione, senza sacrificarne la funzionalità e la completezza.

Le norme dovranno avere un carattere prescrittivo, in funzione di tre distinte operazioni: la costruzione del vocabolario, con il suo corredo sintattico-applicativo; il suo aggiornamento e sviluppo; il suo impiego nell'indicizzazione dei documenti. In altri termini, le voci del vocabolario, le note applicative e le voci del manuale (redatte nella fase di costruzione del linguaggio), le proposte di aggiornamento e sviluppo (formulate nel contesto di una gestione

cooperativa) e le voci d'indice (create dagli indicizzatori nella catalogazione dei documenti) devono essere conformi alle norme generali.

Inoltre, dovranno essere organiche, ossia riferirsi a tutte e tre le operazioni fondamentali dell'indicizzazione (analisi concettuale dei documenti, controllo del vocabolario e costruzione delle stringhe di soggetto), così da fornire ai catalogatori una guida completa per l'identificazione e la rappresentazione del soggetto.

Le norme si applicano non solo ai nuovi soggetti, ma anche a quanto del linguaggio preesistente verrà trasferito nel nuovo. Il recupero della terminologia e delle relazioni del *Soggettario* è stato definito, infatti, come un processo guidato da norme. A tale proposito, è stato formulato il seguente criterio generale, relativo alla scelta della forma del termine: in caso di dubbio tra due o più forme, si tenderà a preferire la forma preesistente, se questa scelta non contrasta con i principi generali di formalizzazione del vocabolario. Quando, per coerenza con le norme, si dovrà preferire una forma nuova, sarà la struttura della voce di vocabolario a garantire, mediante relazioni, note e varianti, la corrispondenza tra la forma non più usata e quella nuova.

Sommario delle norme

Uno sguardo d'insieme dell'articolazione e del contenuto delle norme è dato dalla figura seguente:

Introduzione	Singolare e plurale
Sommario	Termini composti
Riferimenti	Nomi propri
	Omografia
Principi	Struttura del vocabolario
Unità	Relazioni di equivalenza
Uniformità	Relazioni gerarchiche
Univocità	Relazione associativa
Predittività	
Specificità	Costruzione delle stringhe di soggetto
Esaustività	Analisi dell'enunciato di soggetto
Coestensione	Tipi di relazioni
Specificità	I ruoli
	L'ordine di citazione
Analisi concettuale dei documenti	Punteggiatura e connettivi
Il soggetto	Il nucleo
La procedura	Principio della relazione uno a uno
Esame delle fonti	Principio della dipendenza logica
Selezione dei concetti	Complementi
L'enunciato di soggetto	
Controllo del vocabolario	Indice analitico
Morfologia	

Fig. 2: Sommario delle norme

Nei paragrafi seguenti non si intende sintetizzare le raccomandazioni per l'elaborazione delle norme contenute nello studio, bensì sottolinearne alcuni aspetti qualificanti e distintivi.

Principi

I principi generali che governano la forma di rappresentazione di un soggetto nel linguaggio di indicizzazione possono essere ridotti a tre fondamentali: uno riguarda l'unicità della forma scelta (uniformità e univocità), l'altro la sua aderenza a determinati criteri, il terzo la sua corrispondenza al contenuto di soggetto che deve esprimere. Questi principi si basano rispettivamente sulle nozioni di *unità*, *predittività*, *specificità*.

Il primo è un principio tradizionale e ben noto. Nello studio, coerentemente con l'individuazione delle due unità del linguaggio, i termini di indicizzazione e le stringhe di soggetto, si raccomanda che i requisiti di uniformità e univocità, e i criteri di controllo della sinonimia e della polisemia che ne discendono, siano formulati in modo da valere sia per i singoli termini che per le stringhe di soggetto.

Il principio di predittività ha ricevuto un'attenzione e una formalizzazione più recente, parallelamente al consolidarsi della normativa sull'indicizzazione per soggetto: stabilisce che le forme di rappresentazione dei soggetti (termini e stringhe) vengano scelte secondo criteri certi, logici e di applicabilità generale, i quali garantiscano coerenza e omogeneità del linguaggio, prevedibilità dei modi di espressione dei concetti e delle loro relazioni.

Il principio di specificità è il più noto e il più "specifico" tra i principi generali dell'indicizzazione per soggetto: raccomanda che ciascuna voce d'indice esprima nel modo più preciso possibile il contenuto di soggetto cui si riferisce. Qui se ne sottolinea la triplice valenza, in relazione alle tre operazioni fondamentali dell'indicizzazione per soggetto: come *esaustività* dell'analisi, come *coestensione* delle stringhe di soggetto, come *specificità* dei termini. Una ridefinizione rigorosa del principio di specificità è necessaria per superare le ambiguità che la sua applicazione ha avuto nella soggettazione tradizionale.

L'applicazione di questi principi generali evidenzia la duplice struttura del linguaggio, costituita dal vocabolario e dalla sintassi, e la corrispondente duplice natura delle relazioni tra i concetti, identificate come relazioni semantiche e relazioni sintattiche.

Analisi concettuale dei documenti

Il riferimento normativo è costituito dallo standard internazionale ISO 5963-1985, i cui criteri e procedure dovranno, tuttavia, essere precisati e adattati ad un linguaggio di indicizzazione di tipo preordinato, in particolare per quanto riguarda la definizione di soggetto e l'utilizzazione delle funzioni logiche nella selezione e nell'organizzazione dei concetti.

Se, infatti, in generale e a prescindere da contesti particolari, vale la definizione data dalla norma ISO, in base alla quale soggetto può essere considerato "ogni concetto o combinazione di concetti che rappresenti un tema all'interno del documento", nel contesto di un linguaggio sintetico e preordinato è pertinente considerare come soggetto il *tema di base*, ossia il tema complessivo o centrale del documento, nozione introdotta criticamente nella "Guida GRIS".

Analogamente, la *lista di controllo*, suggerita dalla norma ISO per favorire un approccio sistematico all'identificazione dei concetti che costituiscono gli elementi essenziali nella descrizione del soggetto del documento, può essere stabilita facendo riferimento alla

nomenclatura delle funzioni logiche che sottendono i ruoli impiegati per la costruzione delle stringhe di soggetto, tenendo presente che, sul piano dell'analisi, tali nozioni fungono da mezzi euristici per la selezione dei concetti e non da criteri regolatori del loro ordine di citazione nelle stringhe di soggetto.

Infine, la norma sull'analisi concettuale dei documenti dovrebbe suggerire la formulazione di un *enunciato di soggetto*, che esprima, mentalmente e nella forma del linguaggio naturale, il tema individuato quale soggetto del documento. L'enunciato consente, infatti, di controllare che i concetti e le loro relazioni siano espressi con precisione e in modo chiaro; inoltre, se da un lato esso rappresenta il punto di arrivo dell'analisi concettuale, dall'altro costituisce il punto di partenza per la traduzione del soggetto nel linguaggio di indicizzazione.

Controllo del vocabolario

Il principio della formalizzazione del vocabolario e delle relazioni semantiche è un principio essenziale dei linguaggi di indicizzazione che utilizzano vocabolari controllati. Le regole fondamentali sono le seguenti:

- a) un concetto è rappresentato da uno e un solo termine; a ogni termine corrisponde uno e un solo concetto;
- b) le relazioni che sono presentate nel vocabolario sono sempre relazioni a priori;
- c) queste relazioni devono essere ricondotte a tipi di relazione definiti in modo rigoroso e chiaro, cosicché ognuno individui una struttura classificatoria, il cui significato è omogeneo per tutte le relazioni ad esso riconducibili.

I criteri adottati per svolgere le funzioni di controllo del vocabolario sono sostanzialmente quelli contenuti nella norma ISO 2788-1986. Si tratta di principi logici e, dunque, omogenei e prevedibili. In particolare, essi presentano una duplice caratteristica:

- a) sono criteri di applicabilità generale, ossia che valgono per la generalità dei termini e non per una parte di essi;
- b) si basano sul metodo dell'analisi categoriale, che consiste nel riferire un termine, sulla base del suo significato, ad una categoria o tipo concettuale fondamentale (un'entità, un materiale, un'attività, una proprietà, ecc.).

Le norme dovrebbero, dunque, fornire un elenco di categorie generali, cui fare riferimento nell'esecuzione delle procedure di controllo terminologico. Le categorie comprese in questo elenco saranno le medesime nelle quali è organizzato il vocabolario.

Morfologia

Riguardo al **numero**, vengono proposti il *principio della numerabilità*, per la scelta tra singolare e plurale, e il *principio dello scostamento categoriale*, per la scelta di entrambe le forme.

Il primo stabilisce l'uso del plurale per i termini che rappresentano concetti *numerabili*, che cioè si possono contare, e l'uso del singolare per i termini che non sono numerabili, che cioè rappresentano concetti di massa. Il secondo ammette la compresenza nello stesso vocabolario delle due forme, singolare e plurale, per uno stesso lemma solo se i concetti rappresentati dalle due forme appartengono chiaramente a due categorie o faccette diverse.

Se il secondo principio fa direttamente riferimento all'analisi categoriale, anche l'applicazione del primo è facilitata dal ricorso alle categorie cui appartengono i concetti: in caso di dubbio tra singolare e plurale, si darà la precedenza al numero che è tipico della categoria di appartenenza del concetto.

Va tenuto presente che, pur rispettando il principio dello scostamento categoriale, i casi di coesistenza delle due forme per uno stesso lemma saranno più numerosi in un vocabolario generale che in uno specializzato. Ad ogni modo, le modifiche nella forma plurale-singolare dei termini presenti nel *Soggettario* potranno essere facilmente gestite nel nuovo strumento come forme varianti.

Riguardo ai **termini composti**, si seguono sostanzialmente il metodo di analisi e i criteri di scomposizione dettati da ISO 2788, allo scopo di:

- a) accrescere la coerenza del vocabolario di indicizzazione, favorendo una pratica uniforme nel trattamento dei termini composti;
- b) ridurre la complessità dei termini di indicizzazione, in accordo con il principio generale che i termini devono rappresentare concetti semplici o unitari;
- c) tenere sotto controllo le dimensioni del vocabolario, riducendone anche i costi di gestione.

Per ridurre l'impatto della scomposizione sui concetti rappresentati nel *Soggettario* con un termine composto, si stabilisce il criterio per cui, in caso di dubbio tra l'applicazione dei principi di mantenimento del termine composto o di scomposizione, si opta per il mantenimento del termine composto anche nel nuovo linguaggio.

Per la morfologia dei **nomi propri** si richiamano nello studio di fattibilità le raccomandazioni contenute nella "Guida GRIS", relativamente alla natura e alle funzioni dei nomi propri, ai criteri per stabilire la loro forma e alla necessità di perseguire il maggior livello possibile di uniformità con la forma autorizzata del catalogo per autori. Questa tematica sarà oggetto di ulteriori riflessioni, anche nella prospettiva della revisione degli attuali codici di catalogazione.

Le tecniche di **disambiguazione** servono ad eliminare l'ambiguità che si verifica quando ad uno stesso termine siano associati nel linguaggio naturale due o più significati. Sono, dunque, dispositivi necessari del linguaggio di indicizzazione; tuttavia, in un sistema preordinato, la disambiguazione può essere un elemento di appesantimento della stringa di soggetto e un ostacolo alla sua comprensione. Conseguentemente, accanto alla descrizione delle tecniche di disambiguazione si devono suggerire le circostanze e i metodi per limitarne l'uso, p.e., quando l'ambiguità di un termine, pur presente nel linguaggio naturale, non si verifica nel linguaggio di indicizzazione; oppure, disambiguando, in un gruppo di omografi, tutti i termini meno uno, scelto secondo determinati criteri. Altre limitazioni della disambiguazione consistono nell'evitare la disambiguazione con funzione esplicativa, ossia per chiarire il significato di un termine, o quella preventiva o quella omografica, ossia applicata ai significati accessori di un termine. Riguardo alle tecniche di disambiguazione, si raccomanda che la qualificazione, da aggiungere tra parentesi dopo il termine vero e proprio, sia espressa preferibilmente con un termine generico, gerarchicamente sovraordinato al termine da disambiguare, e non con il nome della

disciplina o del campo di attività a cui afferisce il concetto, che non appare idoneo ad eliminare l'ambiguità, dando spesso luogo ad una disambiguazione poco rigorosa e imprecisa.

Struttura del vocabolario

La struttura del vocabolario è costituita dalle relazioni semantiche. Si distinguono i sistemi a due e quelli a tre relazioni. Per il nuovo Soggettario viene proposto il modello a tre relazioni, perché:

- a) è dotato di una maggiore ricchezza e flessibilità nella rappresentazione delle relazioni;
- b) inoltre, in quanto più strutturato, presenta una sostanziale compatibilità con il modello a due relazioni, compatibilità che invece non si ha nel senso opposto.

Sulla base del modello proposto, le norme dovrebbero descrivere il sistema delle relazioni, costituito da: relazione di equivalenza, relazioni gerarchiche, relazione associativa.

La relazione di equivalenza si ha quando si assume che due o più termini abbiano lo stesso significato. I termini equivalenti formano il cosiddetto gruppo di equivalenza, nel quale un termine è quello preferito, tutti gli altri sono termini non preferiti. Si possono riconoscere essenzialmente due tipi di relazioni di equivalenza:

- a) la relazione di equivalenza propriamente detta, ossia tra termini che sono considerati sinonimi anche nel linguaggio naturale;
- b) la relazione di quasi sinonimia, ossia tra termini che nel linguaggio naturale hanno diverso significato, ma che ai fini dell'indicizzazione sono trattati come sinonimi.

Riguardo alle relazioni gerarchiche, che costituiscono la struttura principale del vocabolario, viene adottato il principio in base al quale due termini possono essere legati da una relazione gerarchica solo se appartengono alla stessa categoria: p.e., un termine che rappresenta un concetto di organismo non può avere come termine generico un termine che rappresenta un concetto di disciplina. Questo principio costituisce il fondamento della struttura del vocabolario, validando da un lato l'ammissibilità di una relazione gerarchica e dall'altro configurando questa struttura come una struttura categoriale.

Si riconoscono tre tipi di relazione gerarchica:

- a) relazione generica: è la relazione gerarchica tipica in quanto utilizzata nelle tassonomie. Stabilisce una relazione tra un genere e le sue specie. Per verificare la validità di una relazione generica si può ricorrere al test logico "tutti-alcuni" confrontato con "alcuni-alcuni";
- b) relazione tutto-parte: è considerata di solito una relazione di tipo sintattico; tuttavia, si possono riconoscere alcuni casi nei quali la parte implica l'intero; in questi casi la relazione va trattata come semantica;
- c) relazione esemplificativa: è il legame che esiste tra una classe, espressa da un nome comune, e un esempio individuale di questa classe, espresso da un nome proprio.

Riguardo alla struttura dei nomi propri, si raccomanda di attivare la relazione esemplificativa, espressa non come una relazione particolare, ma come relazione gerarchica.

La relazione associativa è considerata residuale rispetto alle prime due: si ha quando due termini sono legati da una relazione definitoria, diversa tuttavia dalla relazione di equivalenza e da quella gerarchica. La relazione associativa è una relazione reciproca.

La strutturazione del vocabolario del nuovo Soggettario è guidata, inoltre, da alcune scelte di fondo, che riguardano:

- a) la cosiddetta *poligerarchia temperata*: viene adottata la monogerarchia, lasciando tuttavia la possibilità della poligerarchia in alcuni casi tipici (p.e., strumenti musicali, organismi, organizzazioni ecc.), pur nel rispetto rigoroso dei principi della gerarchia, in particolare quello dell'appartenenza categoriale;
- b) l'ancoraggio stabile della relazione associativa al principio definitorio, evitando che qualsiasi associazione di idee tra concetti sia ammissibile come relazione associativa;
- c) la drastica limitazione delle relazioni termine-stringa e viceversa quasi esclusivamente ai casi di equivalenza;
- d) l'uso di etichette e di termini-etichetta per costruire, quando si renda necessario, classi mutualmente esclusive;
- e) l'inserimento, in fase di costruzione del vocabolario, di termini che, pur non essendo utilizzati nell'indicizzazione, servono a dare coerenza e completezza alla gerarchia.

Categorie e categorizzazione

La categorizzazione dei termini è un'operazione fondamentale per la strutturazione del vocabolario. Le categorie costituiscono, infatti, le strutture classificatorie fondamentali del vocabolario. Trattandosi di un'operazione complessa, lo studio di fattibilità dà alcune indicazioni sulla natura delle categorie e il loro impiego nel linguaggio di indicizzazione.

Inanzitutto, chiarisce la distinzione tra gerarchia, faccetta e categoria, concetti che a volte sono usati in modo intercambiabile: la gerarchia è una catena di termini tutti subordinati ad un top term; la faccetta è un principio di suddivisione di una classe (o di una categoria); la categoria è lo strumento analitico, al livello più astratto, che rende possibile l'analisi della terminologia e la sua strutturazione.

Per impiegare le categorie nella strutturazione del vocabolario è necessario compiere alcune operazioni preliminari, che consistono:

1. nella definizione chiara ed esplicita delle caratteristiche delle categorie utilizzate nel vocabolario (creazione di un profilo della categoria);
2. nella delimitazione degli ambiti di sovrapposizione delle categorie;
3. nell'individuazione delle faccette principali di ogni categoria.

La *Fig. 3* mostra lo schema delle categorie individuate, in via del tutto sperimentale, nello studio di fattibilità.

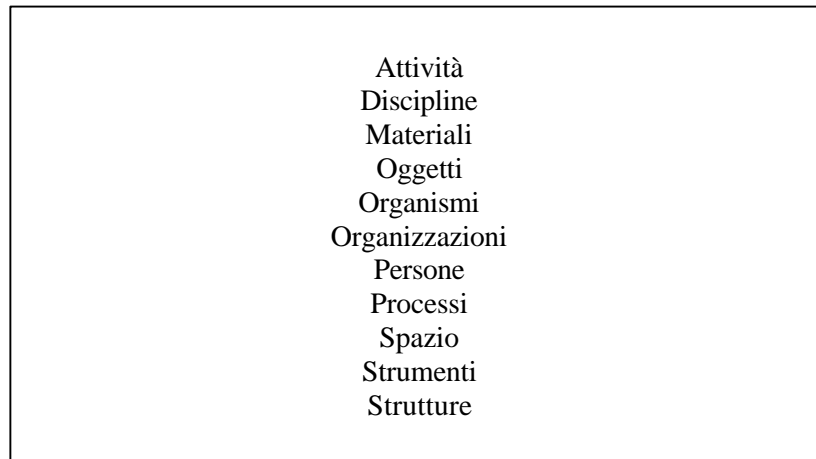


Fig. 3: Schema delle categorie

Il modello sintattico

Il modello sintattico proposto per il nuovo Soggettario consiste nell'analisi dei ruoli, ossia delle funzioni logiche svolte dai concetti nella definizione del soggetto (p.e., azione, oggetto, agente, ecc.), e nell'adozione di un ordine di citazione ad essi associato (p.e., l'azione transitiva dopo l'oggetto e prima dell'agente). E', dunque, un modello che implica due componenti: uno schema dei ruoli e un criterio per l'ordine di citazione.

Questo modello, coerentemente con una preordinazione di tipo sintetico, garantisce un ordine di citazione predefinito, in grado di soddisfare tre requisiti: l'uniformità, l'univocità e la prevedibilità delle stringhe. L'uniformità consente una rappresentazione coerente del soggetto, prevenendo fenomeni di dispersione dell'informazione con conseguente diminuzione del grado di richiamo; l'univocità ne assicura una formulazione chiara, garantendo un maggior grado di precisione; la prevedibilità implica che soggetti diversi, i cui concetti sono legati dalle medesime relazioni sintattiche, siano indicizzati in modo analogo, ossia con uno stesso ordine di citazione, facilitando agli utenti il riconoscimento dei modi di rappresentazione dei soggetti nel linguaggio di indicizzazione.

Di questo modello (tipi di relazione, ruoli e criteri per l'ordine di citazione) lo studio di fattibilità propone un'ipotesi di formalizzazione che intende perseguire l'obiettivo di una semplificazione del modello, senza sacrificarne la ricchezza.

Il processo di costruzione della stringa

Il processo di costruzione della stringa di soggetto è articolato in due fasi, distinte e successive, che corrispondono alla natura analitico-sintetica del linguaggio:

1. la prima fase è quella dell'analisi logica dell'enunciato che definisce il soggetto, volta ad individuare non solo i singoli concetti, ma anche, e soprattutto, le relazioni logiche tra i concetti e quindi le diverse funzioni logiche che quelle relazioni determinano per ciascun concetto o insieme di concetti.
2. la seconda fase è articolata in due sottofasi:

- i. la prima consiste nell'assegnazione dell'appropriato ruolo sintattico ad ognuno dei termini o sintagmi componenti la stringa, in base alle relazioni e alle funzioni logiche individuate nell'enunciato;
- ii. la seconda nell'ordinamento dei termini secondo la posizione stabilita per ciascun ruolo.

Le relazioni

Nello studio di fattibilità viene avanzata l'ipotesi di ricondurre le diverse relazioni presenti negli enunciati di soggetto a due tipi:

- a) le relazioni che sono mediate da un'azione transitiva, o ad esse assimilabili (*relazioni transitive*);
- b) le relazioni non mediate da un'azione transitiva (*relazioni di appartenenza*).

Ognuno di questi due tipi di relazione determina funzioni logiche diverse, attiva una distinta struttura logica, come è illustrato nella *Fig. 4*.

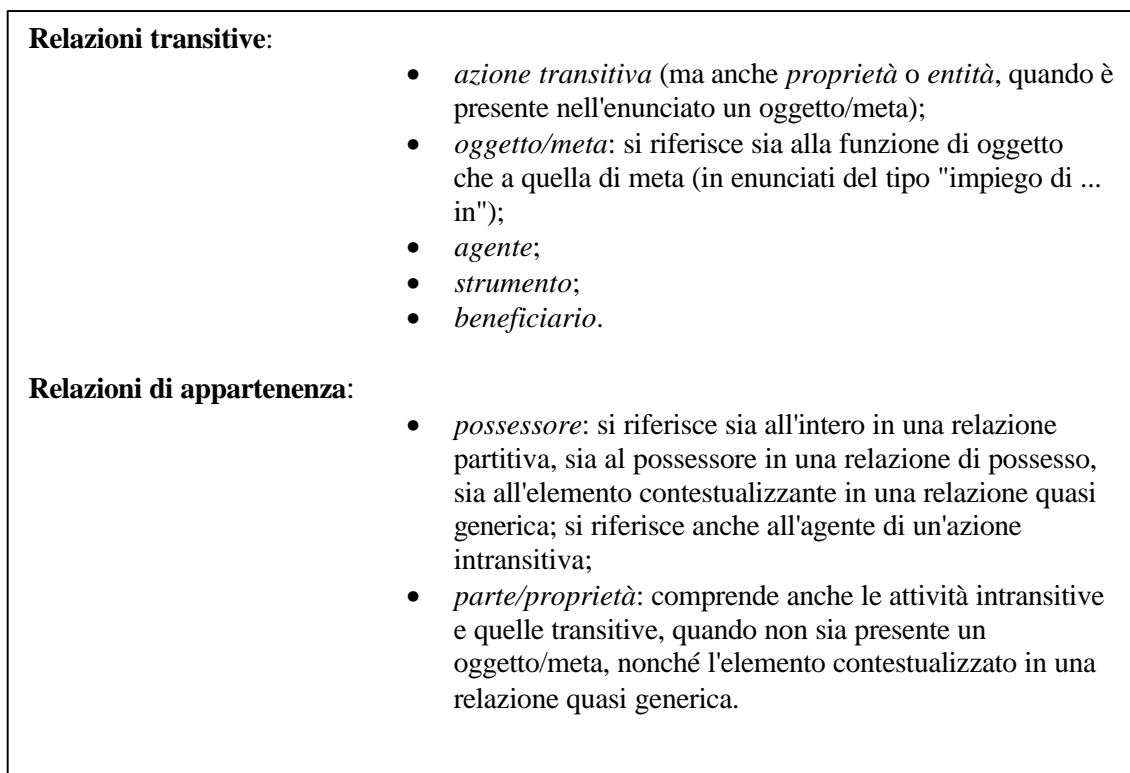


Fig. 4: La struttura delle relazioni

L'analisi delle relazioni presenti in un enunciato di soggetto si attua a partire dal riconoscimento dei concetti che hanno una valenza sintattica, da cui dipendono cioè le funzioni assunte dai concetti direttamente collegati e che, pertanto, determinano la struttura dell'enunciato o di un suo segmento: innanzitutto, il concetto di azione o altro tipo di concetto che, nel contesto dell'enunciato, svolge analoga funzione, ma anche i concetti di parte e di proprietà.

In uno stesso enunciato possono essere presenti più tipi di relazioni, p.e., una relazione transitiva e una partitiva, come nell'enunciato "formazione professionale del personale direttivo delle Unità sanitarie locali". Ad un primo livello, l'analisi dell'enunciato deve tendere a riconoscere la relazione sulla quale è incentrato l'enunciato, poiché è da questa relazione che derivano le funzioni logiche primarie ed è essa che determina, nella fase di ordinamento, l'applicazione appropriata dello schema dei ruoli. Individuate le funzioni primarie (p.e., l'azione "formazione professionale" e l'oggetto "personale direttivo delle Unità sanitarie locali"), si verificherà, ad un secondo livello di analisi, se una di queste funzioni (p.e., l'oggetto "personale direttivo delle Unità sanitarie locali") sia rappresentata da un sintagma ulteriormente scomponibile in elementi legati tra loro secondo determinate funzioni (p.e., la parte "personale direttivo" e l'intero "Unità sanitarie locali").

L'analisi dell'enunciato non può prescindere dall'introduzione di un'altra distinzione funzionale, quella tra *nucleo* del soggetto e concetti che fungono da *complementi* del nucleo. Il nucleo del soggetto è la componente fondamentale della definizione del soggetto, all'interno della quale i concetti sono legati da relazioni logiche dirette e reciproche, come quelle che legano un'azione all'oggetto o una proprietà al possessore. I complementi del nucleo hanno una condizione relazionale differente, poiché non sono reciprocamente correlati, ma ognuno singolarmente è correlato all'intero nucleo, di cui specifica il tempo, il luogo o la forma.

I ruoli

La fase successiva consiste nel passaggio dall'individuazione delle funzioni logiche all'assegnazione dei ruoli ai termini nella stringa di soggetto. In base alle relazioni e alle funzioni logiche individuate nell'enunciato, si assegna ad ogni termine l'appropriato ruolo sintattico. Per compiere questa operazione, occorre un elenco di ruoli, che li identifichi in modo preciso ed esaustivo. Nello studio di fattibilità viene proposto lo schema dei ruoli illustrato nella *Fig. 5*.

L'ordine di citazione

L'ordinamento dei termini nella stringa implica la definizione di principi generali e criteri specifici in base ai quali è regolata la successione dei diversi ruoli.

Innanzitutto, il nucleo, costituito da un insieme di concetti o anche da un concetto singolo, viene sempre citato prima dei complementi. Per ordinare i concetti all'interno del nucleo sono stati formulati criteri diversi nel nome, ma sostanzialmente convergenti nella sostanza. In particolare, questi criteri presentano tre caratteristiche comuni:

- a) la "relazione uno a uno", ossia il mantenimento, nell'ordinamento dei termini della stringa, delle loro relazioni dirette (*principio della relazione uno a uno*);

Ruoli nucleari

- *concetto chiave* [o *testa*]: è il ruolo assegnato al termine che rappresenta l'oggetto/meta o il possessore;
- *elemento transitivo*: è il ruolo assegnato al termine con il quale è espresso il concetto che stabilisce la relazione tra un oggetto/meta ed un agente/strumento o che comunque determina, per uno dei concetti espressi nell'enunciato, la funzione di oggetto/meta; tale funzione è svolta normalmente da concetti di attività (azioni, processi, discipline), ma può essere svolta anche da concetti di proprietà o entità a condizione che sia presente un oggetto/meta;
- *agente*;
- *strumento*;
- *beneficiario*;
- *parte/proprietà*: il ruolo è assegnato sia a termini che rappresentano parti o proprietà, sia a termini che rappresentano attività, quando non si verificano le condizioni per le quali la proprietà, l'entità o l'attività debbano essere considerate elemento transitivo; è assegnato anche all'elemento contestualizzato, in una relazione quasi generica.

Ruoli complementari

- *luogo*: ruolo da assegnare, probabilmente, solo ai nomi propri di territori e località, quando i relativi concetti non siano coinvolti in relazioni transitive o di appartenenza;
- *tempo*;
- *forma intellettuale*: è il ruolo da assegnare ai termini che precisano il tipo di approccio intellettuale al tema, indicando un aspetto generale del tema (es., Aspetti culturali) o esplicitamente la disciplina nel cui contesto il tema è studiato; a termini come "Teoria/e" o "Bibliografia/e", quando non compresi nel nucleo, dovrebbe essere assegnato questo ruolo;
- *caso*;
- *destinazione/forma bibliografica*: i termini che indicano la destinazione e la forma bibliografica del documento vengono introdotti nella stringa non per segnalare una particolare modalità di pubblicazione, ma per precisare ulteriormente il livello e il "taglio" della trattazione, quanto essa sia esaustiva ed in che modo sia organizzata;
- *forma materiale*: è il ruolo da assegnare ai termini che forniscono informazioni sulla natura fisica del documento.

Fig. 5: Schema dei ruoli

- b) la "dipendenza", ossia un ordinamento in base al quale un elemento dipendente segue direttamente l'elemento da cui dipende (*principio della dipendenza* o anche *principio della dipendenza dal contesto*);
- c) il "sistema passivo", ossia un ordinamento dei termini nella stringa analogo a quello delle parole di una frase in forma passiva (*principio della costruzione passiva*).

Il *principio della relazione uno a uno* è un principio di livello molto generale che, in caso di conflitto, prevale su qualsiasi altro principio sintattico. Tuttavia, esso non è sufficiente a definire un ordinamento dei termini. In PRECIS viene formulato un altro principio, quello della *dipendenza dal contesto*: ciascun termine stabilisce il contesto di quelli che seguono, delimita l'ambito nel quale essi sono considerati. Tale contesto può rappresentare una localizzazione (il luogo), un campo di applicazione (l'oggetto di un'azione transitiva o l'agente di un'azione intransitiva) o l'ambito di appartenenza (un intero o un possessore). Nella "Guida GRIS" i principi sono due, quello della *costruzione passiva* (per le relazioni incentrate su un concetto di

attività) e quello *della dipendenza dal contesto* (per le relazioni non incentrate su un concetto di attività).

Nello studio di fattibilità, fermo restando il principio della relazione uno a uno, l'altro principio è definito come *principio della dipendenza logica*, il quale trova attuazione in due criteri più specifici, relativi ai due tipi di relazione già individuati, la relazione transitiva e quella di appartenenza. Nel contesto delle relazioni transitive, il principio della dipendenza logica si attua in un “principio dell’obiettivo”, in base al quale è data la precedenza al termine che rappresenta l'oggetto o la meta. Questo principio può essere espresso come *principio della dipendenza dall'oggetto/meta*: in presenza di una relazione transitiva, i termini si dispongono nella stringa secondo il principio della relazione uno a uno, a partire dal termine che rappresenta l'oggetto o la meta verso cui è diretto l'elemento transitivo. Nel contesto delle relazioni di appartenenza, il principio della dipendenza logica si realizza in un principio in base al quale è data la precedenza al termine che rappresenta il possessore. Questo principio può essere espresso come *principio della dipendenza dall'intero/possessore*: in presenza di una relazione di appartenenza, i termini si dispongono nella stringa nell'ordine *possessore – parte/proprietà*.

Anche per l'ordine di citazione dei complementi sarebbe corretto indicare un criterio logico, non essendo appropriato né il principio della relazione uno a uno né il principio della dipendenza logica. Tale criterio sostanzialmente c'è (la decrescente attinenza al tema), ma si dovrà verificare se la sua espressione e spiegazione risulta efficace oppure troppo complessa o troppo vaga. L'ordine di citazione dei complementi è comunque invariabile: – *luogo – tempo – forma intellettuale – caso – destinazione/forma bibliografica – forma materiale*.

La sintassi, dunque, è basata su una sorta di dualismo, nel senso che è articolata su due nozioni complementari: due tipi di relazioni, due tipi di valenza sintattica, due principi generali, due criteri specifici. Questa impostazione, che risponde ad un'esigenza di semplificazione nella formalizzazione delle relazioni e dei ruoli sintattici, è illustrata nella *Fig.6*.

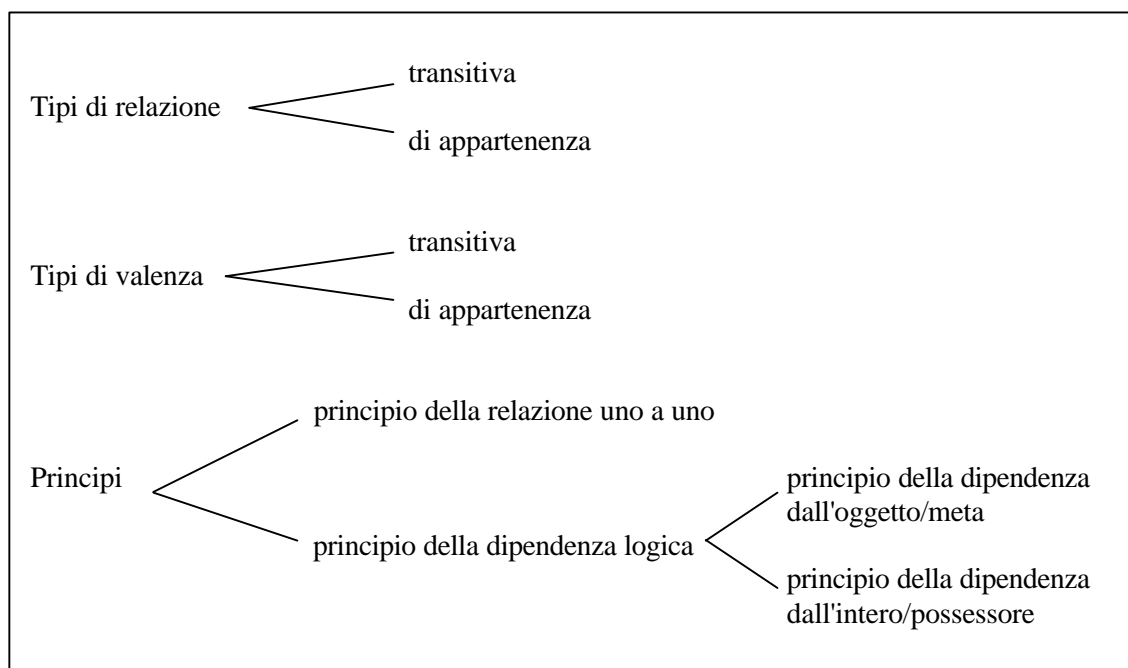


Fig. 6: Carattere dualistico del modello sintattico

L'impatto delle norme sul *Soggettario*

In vari punti dello studio di fattibilità vengono individuati gli elementi di continuità e di discontinuità tra il *Soggettario* e il nuovo strumento, risultanti dall'applicazione delle norme per il controllo del vocabolario e per l'ordine di citazione. Si segnalano qui le principali differenze.

Biunivocità concetto/termine

Si tratta di uno degli aspetti di più evidente distacco del *Soggettario* dai principi della formalizzazione dei linguaggi documentari moderni, che si manifesta nella forte presenza di quasi sinonimi, conseguenza indesiderabile del principio dell'aderenza al linguaggio del documento o dell'autore. Inoltre, lo scarso impiego di note definitorie o di ambito rende problematica l'individuazione del significato specifico attribuito a ciascuno dei quasi sinonimi, individuazione che è le più volte affidata alla contestualizzazione sintattica.

Le ridondanze si presentano soprattutto in due casi: a) uso sia del singolare che plurale dello stesso lemma (**Donna/Donne**, **Scuola/Scuole**, etc.); b) impiego di più forme dello stesso concetto mediante disambiguazioni ridondanti o risolvibili sintatticamente (uso in parentesi di **Concetto**, **Vocabolo**, **Toponimo**, **Nome**, etc.). Inoltre, le liste separate delle suddivisioni (suddivisioni di soggetti geografici, suddivisioni di soggetti biografici e suddivisioni formali) favoriscono la moltiplicazione delle forme e dei possibili significati: p.e., il concetto "casa" figura come **Case** nella sequenza principale e nelle suddivisioni dei soggetti geografici, come **Casa** nei soggetti biografici; si ha **Enciclopedie** nella sequenza principale, con un rinvio, per l'uso come suddivisione, a **Enciclopedie e dizionari**, che non è elencato nella lista delle suddivisioni biografiche, mentre nelle altre due sì; né sotto la voce **Enciclopedie** né sotto la suddivisione **Enciclopedie e dizionari** figura una nota esplicativa o applicativa, che separi con chiarezza le varie sfumature di significato.

Nonostante che il numero dei casi di ridondanza sia piuttosto ampio, la relativa obsolescenza del vocabolario e il fatto che le ridondanze si presentino spesso per termini generici rendono relativamente semplice il compito di risolvere il problema. Nel nuovo *Soggettario*, tutti i termini di indicizzazione, indipendentemente dalle loro funzioni nella descrizione del soggetto (precedentemente identificate con la distinzione tra voce principale e suddivisione), saranno individuati come elementi omogenei del vocabolario, collegati tra loro mediante la struttura delle relazioni e un opportuno corredo di note esplicative.

Riguardo all'impiego delle due forme, singolare e plurale, si potrà fare ricorso nel nuovo strumento ad artifici di gestione automatizzata, per rendere compatibili le precedenti soluzioni con quelle corrette, come verrà illustrato nel capitolo sulla struttura della voce di vocabolario. Infine, la regolarizzazione dei significati resi ridondanti dalla disambiguazione parentetica non presenta particolari problemi.

Morfologia

La morfologia è uno degli aspetti critici del *Soggettario* e una delle maggiori fonti di disomogeneità.

La disambiguazione viene eseguita prevalentemente per mezzo della contestualizzazione disciplinare, tranne in alcuni casi in cui si usa il nome della classe di appartenenza. Questo

comporta spesso effetti indesiderati: p.e., nel caso di "società" come organizzazione, il termine **Società** viene disambiguato con il nome di due discipline, **Società (Diritto e economia)**.

Inoltre, in un sistema preordinato i casi di disambiguazione parentetica dovrebbero essere ridotti al minimo indispensabile, per non costituire elemento di disturbo nella lettura della stringa; mentre nel *Soggettario* si hanno casi non compatibili con un uso razionale del vocabolario, che derivano da:

- a) disambiguazione esplicativa: p.e., **Dama (Giucoco)** e **Genio (Arma)**;
- b) disambiguazione sintattica: p.e., **Colpa (Diritto penale)** e **Colpa (Diritto civile)**;
- c) ridondanza di significati: p.e., **Arte, Arte (Concetto), Arte (Vocabolo)**.
- d) ridondanza di qualificazioni: p.e., **Tex Willer (Fumetto)**.

Le soluzioni a questo tipo di disomogeneità sono varie, in quanto comportano un'analisi accurata delle casistiche e l'impiego di strumenti sintattici e di riduzione delle ridondanze.

Riguardi ai termini composti, si potrebbero citare molti casi di disomogeneità, anche se si è già accennato ai criteri per ridurre l'impatto della scomposizione sulla terminologia del *Soggettario*. In ogni caso, nella fase di passaggio al nuovo vocabolario, è necessaria un'analisi accurata di tutti i termini composti per procedere alla loro validazione o alla scomposizione.

Struttura

Il *Soggettario* è un vocabolario a due relazioni, anche se in realtà la sua struttura classificatoria non è riconducibile a un modello facilmente circoscrivibile, perché le relazioni stesse non sono molto formalizzate. Per esempio, possiamo cioè dire che la relazione *v.a./*** corrisponda alla relazione gerarchica (NT) e/o a quella associativa (RT), sapendo tuttavia che spesso si tratta di relazioni gerarchiche o associative in senso molto lato (p.e., con un termine più generico, ma non appartenente a quella che si dovrebbe considerare la stessa gerarchia; o sulla base di pure associazioni di idee). Per esempio, alla voce **Trasporti** si ha *v.a.* **Diritto postale e ** Commercio, Comunicazioni, Economia, Geografia economica, Ingegneria civile, Locazione d'opera**. Non si può certo sostenere che **Trasporti** abbia come classe generica **Geografia economica** o **Ingegneria civile** o **Locazione d'opera**. Tuttavia, la correttezza delle relazioni risulta più frequente a livello di voci più specifiche.

Con le medesime cautele, si può dire che nel *Soggettario* sono presenti le tre relazioni gerarchiche previste dallo standard dei thesauri: la relazione generica, la relazione esemplificativa, che prende il simbolo ***es.*, e la relazione parte tutto, non distinta dalle altre, cioè trattata con i due simboli normali (*v.a./ ***).

In una qualche misura, le considerazioni svolte sopra valgono anche per la relazione di equivalenza. Infatti, nel *Soggettario* questa relazione, oltre alle funzioni proprie, ne svolge altre spurie. Intanto, un termine non preferito, nel *Soggettario*, può rinviare a più termini preferiti, pratica da evitare. Per esempio: Cibi *v.* **Alimenti; Generi alimentari**.

In altri casi, il rinvio è usato per:

- a) rinviare dalla forma con inversione alla forma diretta (p.e., Euclide, teorema di *v.* **Teorema di Euclide**);
- b) rinviare dal termine nascosto di una stringa formata da due termini uniti dalla congiunzione (p.e., Animali e uomo *v.* **Uomo e animali**);
- c) rinviare da una stringa al termine composto corrispondente e viceversa (p.e., Antenati – Culto *v.* **Culto degli antenati**; Diritto nobiliare *v.* **Nobiltà – Diritto**);

- d) rinviare dal termine nascosto di una stringa di concetti coordinati (p.e., Genotipo v. **Fenotipo e genotipo**);
- e) rinviare da uno specifico non preferito al suo generico (p.e., Soldatini di piombo v. **Giocattoli**); ma anche il contrario (p.e., Parlamentari v. **Deputati; Senatori**).

Alcuni di questi artifici vengono resi obsoleti dall'evoluzione dei sistemi di gestione del linguaggio di indicizzazione; altri vanno corretti in modo da renderli compatibili con relazioni correttamente formalizzate.

Le differenze sul piano sintattico

I principi per l'ordine di citazione determinano una costruzione sostanzialmente passiva della stringa di soggetto, nella quale è posto in prima posizione l'oggetto/meta dell'azione o della funzione agentiva o strumentale.

Questo ordine di citazione è sicuramente prevalente anche nella sintassi del *Soggettario*, in particolare quando la relazione tra l'azione e l'agente o strumento è diretta. Vi compaiono, tuttavia, delle eccezioni, soprattutto in presenza di soggetti biografici, che sono ovviamente incompatibili con l'applicazione coerente del *principio della dipendenza dall'oggetto/meta* e quindi non dovrebbero essere ammesse nel nuovo linguaggio.

Molto meno uniforme è, invece, nel *Soggettario* l'espressione delle relazioni in presenza di un'agente indiretto, cioè quando l'agente è introdotto da un concetto che ne chiarisce la funzione rispetto all'oggetto/meta (p.e., **Atteggimento, Impiego, Giudizi, Partecipazione** ecc.). In questi casi, la sintassi varia a seconda del termine che specifica la funzione agentiva: ne risulta la scarsa coerenza dell'ordine di citazione, causata dall'adozione di criteri lessicali deboli (la presenza di una certa "parola"), anziché di criteri basati sulle relazioni profonde che intercorrono tra gli elementi che definiscono il soggetto da rappresentare. Per il *principio della dipendenza dall'oggetto/meta*, che attiene a quelle relazioni profonde, anche in presenza di un agente indiretto l'oggetto/meta va comunque citato in posizione di concetto chiave, qualsiasi sia il termine che esprime la relazione.

Nella *Fig. 7* sono riportati alcuni esempi di stringhe di soggetto, che esprimono relazioni transitive, rappresentativi delle somiglianze e delle differenze tra vecchio e nuovo *Soggettario*.

Nell'espressione delle relazioni di appartenenza (intero/parte, possessore/proprietà) l'aspetto sintattico si interseca, nel *Soggettario*, con quello terminologico, poiché non raramente queste relazioni sono espresse con termini composti (p.e., **Eliche aeree, Sterilità femminile**). Nella loro rappresentazione puramente sintattica il *Soggettario* rispetta, di norma, il *principio della dipendenza dall'intero/possessore*, per il quale il concetto che esprime l'intero/possessore è citato prima del concetto che esprime la parte/proprietà (es., **Trattori agricoli – Pneumatici, Atmosfera – Trasparenza**). Esistono tuttavia molte eccezioni, sia palesi sia celate sotto metonimia (Fanciulli/Infanzia, Anziani/Senilità). Anche in tal caso è da prefigurare, per il nuovo linguaggio, una maggiore coerenza e predittività delle costruzioni sintattiche.

Nella *Fig. 8* sono mostrate le somiglianze e le differenze tra vecchio e nuovo *Soggettario*, riguardo alla costruzione di stringhe di soggetto che esprimono relazioni di appartenenza.

Coestensione

La coestensione delle stringhe rispetto al tema da rappresentare è un obiettivo da perseguire con il nuovo linguaggio, poiché costituisce un fattore determinante di precisione e, quindi, di efficienza del sistema di indicizzazione e recupero.

L'adozione di criteri di ordinamento fondati sulle relazioni sintattiche tra i concetti che definiscono il tema da indicizzare, e sulle funzioni logiche che ne derivano, permette, ordinariamente, la formulazione coerente di stringhe coestese e intelligibili, come è evidenziato dagli esempi della *Fig. 9*, nei quali, per ciascuna coppia di voci assegnata nella BNI è ipotizzata una stringa unica e coestesa.

<i>Soggettario</i>	Nuovo Soggettario
Acque – Inquinamento da metalli pesanti	Acque – Inquinamento da metalli pesanti
Merci – Trasporto marittimo con container	Merci – Trasporto marittimo con container
Uccelli acquatici – Migrazioni	Uccelli acquatici – Migrazioni
Musica – Storia - Insegnamento agli adulti	Musica – Storia - Insegnamento agli adulti
Leopardi, Giacomo - Concezione della musica	Musica – Concezione di Leopardi, Giacomo
Infezioni - Sistema nervoso centrale	Sistema nervoso centrale - Infezioni
Tumori - Apparato digerente	Apparato digerente – Tumori
Cattolici – Atteggiamento verso gli ebrei	Ebrei - Atteggiamento dei cattolici
Regione – Funzioni in materia di agricoltura	Agricoltura – Competenze delle regioni
Elaboratori elettronici – Impiego nei servizi sanitari	Servizi sanitari – Impiego degli elaboratori elettronici
Elettroforesi - Applicazioni in virologia	Virologia – Applicazioni dell'elettroforesi
Sturzo, Luigi - Giudizi sulla guerra mondiale 1939-1945	Guerra mondiale 1939-1945 – Giudizi di Sturzo, Luigi
Resistenza – Italia – Giudizi degli adolescenti	Resistenza – Italia – Giudizi degli adolescenti
Ayroldi, Antonio - Partecipazione alla Resistenza	Resistenza – Partecipazione di Ayroldi, Antonio
Resistenza – Partecipazione dei cattolici	Resistenza – Partecipazione dei cattolici
Fascismo – Italia – Influssi della civiltà classica	Fascismo – Italia – Influssi della civiltà classica
Acque – Inquinamento – Effetti sull'agricoltura	Agricoltura – Effetti dell'inquinamento delle acque
Funghi – Effetti delle radiazioni	Funghi – Effetti delle radiazioni

Fig. 7: L'espressione delle relazioni transitive

<i>Soggettario</i>	Nuovo Soggettario
Apparato digerente – Mammiferi	Mammiferi - Apparato digerente
Cuore - Feto - Malattie	Feto - Cuore - Malattie
Ghiacciai - Monte Bianco	Monte Bianco – Ghiacciai
Economia - Popoli primitivi	Popoli primitivi – Economia
Miti - Indiani pueblos	Pueblos - Miti
Personalità – Disturbi - Infanzia	Bambini - Personalità – Disturbi
Ipertensione arteriosa - Senilità	Anziani - Ipertensione arteriosa

Fig. 8: L'espressione delle relazioni di appartenenza

1. Il controllo da parte della Corte dei conti sui conti consuntivi degli enti locali
 1. Corte dei conti
 2. Enti locali – Bilancio – Controllo

Enti locali – Bilancio – Controllo della Corte dei conti
2. Gli effetti della sentenza dichiarativa di fallimento nei confronti dei creditori
 1. Fallimento - Dichiarazione – Giurisprudenza
 2. Creditori - Tutela

Creditori – Tutela – Effetti della dichiarazione di fallimento – Giurisprudenza
3. Manuale di laser terapia tricologica
 1. Capelli – Malattie – Terapia
 2. Laser – Impiego terapeutico

Capelli – Malattie – Terapia – Impiego del laser
4. Come farsi le ossa nella vita : il ruolo del calcio nella prevenzione dell'osteoporosi dalla nascita alla terza età
 1. Osteoporosi - Prevenzione
 2. Calcio - Impiego terapeutico

Osteoporosi – Prevenzione – Ruolo del calcio
5. Metodi variazionali per il controllo ottimo
 1. Controlli automatici
 2. Calcolo delle variazioni – Applicazioni

Controlli automatici – Applicazioni del calcolo delle variazioni
6. Il volto trinitario di Dio : nel libro dell'Apocalisse
 1. Bibbia. Nuovo Testamento. Apocalisse – Interpretazione
 2. Trinità - Concezione biblica

Bibbia. Nuovo Testamento. Apocalisse – Temi: Trinità

Fig. 9: Applicazione del principio della coestensione

5. Il vocabolario

Contenuto e caratteristiche fondamentali

Il vocabolario contiene tutti i termini di indicizzazione da impiegare nella descrizione del soggetto dei documenti. I termini, che esprimono concetti singoli o unitari, costituiscono le unità di base del linguaggio, la cui combinazione, secondo le regole di sintassi, dà luogo alle stringhe di soggetto che rappresentano i soggetti composti.

Nella fase di implementazione del nuovo linguaggio, saranno recuperati tutti i termini presenti nel *Soggettario* e nelle *Liste di aggiornamento*, tranne i termini mai utilizzati, e non necessari per la struttura, e i termini obsoleti. Tale recupero, tuttavia, si configura come un processo guidato dalle norme, per cui la terminologia e le relazioni subiranno quelle modifiche morfologiche e strutturali indicate nel capitolo sulle norme. La struttura della voce è predisposta anche per assicurare la compatibilità tra vecchie e nuove forme dei termini.

Ciascun termine costituisce una voce controllata e strutturata del vocabolario, che contiene l'indicazione delle relazioni semantiche, oltre ad altre informazioni di carattere storico, gestionale, applicativo.

Prima di illustrare la struttura della *voce*, è opportuno riassumere le caratteristiche principali del nuovo vocabolario, identificabili nei seguenti punti:

- a) abbandono della distinzione tra "voce principale" e "suddivisioni": ogni termine, indipendentemente dalle diverse funzioni che può svolgere nella descrizione dei soggetti e dalla posizione che assume nelle relative stringhe, costituisce una voce del vocabolario; conseguentemente, sotto ciascuna voce non comparirà più alcuna suddivisione e i termini che nel *Soggettario* compaiono unicamente come suddivisioni saranno trasformati in voci;
- b) aderenza alle norme internazionali per il controllo terminologico e l'indicazione delle relazioni semantiche: vecchi e nuovi termini saranno controllati e strutturati in accordo alle norme stabilite; conseguentemente, le voci del *Soggettario* saranno soggette ad un processo di revisione per renderle compatibili con le nuove norme;
- c) adozione del sistema a tre relazioni: la formalizzazione delle relazioni semantiche prevede l'indicazione delle tre relazioni (di equivalenza, gerarchiche, associativa) sul modello dei thesauri; conseguentemente, le voci del *Soggettario*, costruite sul sistema a due relazioni, saranno ristrutturare sulla base nuovo modello;
- d) definizione di alcune scelte strategiche relative alla struttura classificatoria (categoriale) del vocabolario, alla monogerarchia/poligerarchia (poligerarchia temperata), alla procedura di costruzione del vocabolario (dal basso verso l'alto).

La struttura della voce

Nello studio di fattibilità la voce di vocabolario è stata definita come la cellula del nuovo organismo: è, infatti, l'unità fondamentale, la più piccola parte funzionante del sistema. In analogia con la cellula vivente, la voce di vocabolario svolge una funzione "riproduttiva" e "predittiva": serve cioè a costruire, in modo omogeneo e coerente, tutte quelle relazioni tra i

concetti, sia sul piano semantico che su quello sintattico, di cui c'è bisogno per descrivere i soggetti dei documenti e poterli ricercare in modo efficace ed efficiente. La voce assolve a questa funzione "riproduttiva", avvalendosi di tre tipi di informazioni, che riguardano la natura delle relazioni, le norme che le governano e le procedure per applicarle. Per esempio, le relazioni gerarchiche sono definite come relazione generica, relazione parte-tutto e relazione esemplificativa, sono regolate dal criterio dell'appartenenza dei termini di una gerarchia alla medesima categoria, sono (o possono essere costruite) dal basso verso l'alto, ossia dallo specifico al generale. I simboli BT/NT (e, analogamente, gli altri simboli) contengono tutte queste informazioni, che fanno della voce non una struttura statica, ma dinamica. La struttura della voce può essere definita, dunque, come un "insieme di relazioni", che, operando in associazione con alcune regole, costituisce un modello per costruire "insiemi di concetti interrelati".

E' stata configurata per rispondere ad una molteplicità di esigenze:

- a) di controllo semantico, mediante l'indicazione delle relazioni semantiche;
- b) di strutturazione del vocabolario, mediante l'indicazione della categoria e della classe di appartenenza;
- c) di applicazione del termine nella descrizione dei soggetti, mediante la nota sintattica;
- d) di compatibilità all'indietro con il *Soggettario*, mediante la nota storica, la nota sui cambiamenti all'interno della nota sintattica, l'indicazione della fonte e delle varianti.

La struttura della voce è illustrata nella *Fig. 10*, seguita dalla descrizione dei singoli campi che la compongono.

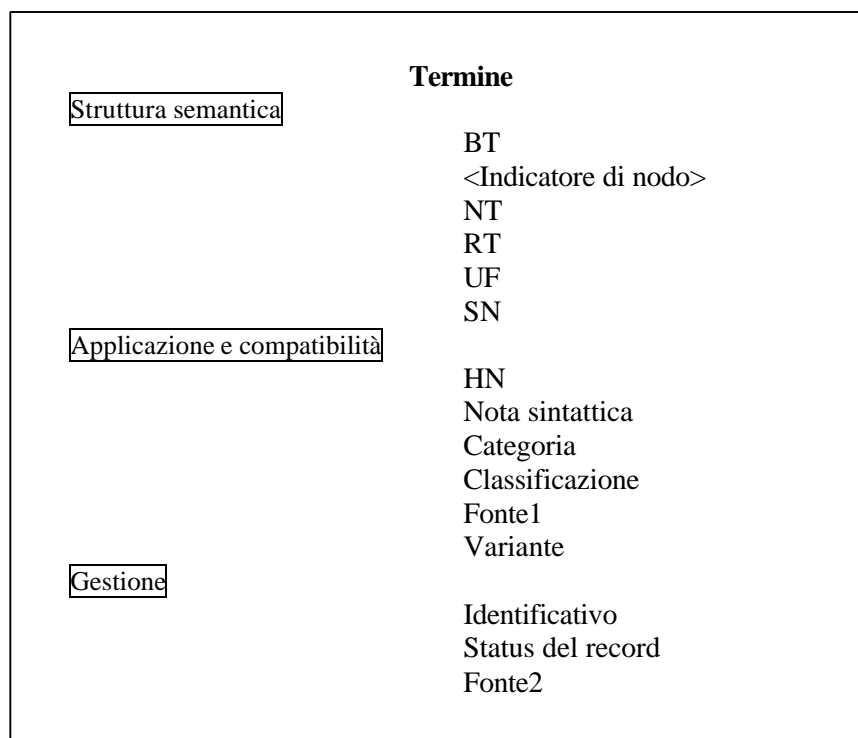


Fig. 10: Struttura della voce relativa ad un termine preferito

Termine. Si usa "termine" nel senso impiegato nello standard ISO 2788-1986; "voce" rappresenta il termine con il suo corredo di relazioni semantiche e note dello stesso tipo; "record", o registrazione, l'insieme di tutti gli elementi (semantici, di applicazione, gestionali).

BT. *Broader term*, cioè termine generico. Il simbolo può essere sostituito con l'equivalente in italiano, non standard, TG. Ripetibile (per i casi di poligerarchia).

Indicatore di nodo. L'*indicatore di nodo*, inserito tra parentesi uncinate (o in alternativa tra parentesi quadre), serve a garantire che da una data classe vengano derivate sottoclassi mutuamente esclusive. Viene visualizzato sempre dall'alto verso il basso (da un termine generico per suddividere i suoi specifici) e in una eventuale presentazione gerarchica o classificata del vocabolario; non è mai visualizzato in senso inverso, cioè dallo specifico al generico; non è un vero e proprio termine e quindi, in linea di massima, non costituisce una voce. Ripetibile.

NT. *Narrower term*, cioè termine specifico. Il simbolo può essere sostituito con l'equivalente in italiano, non standard, TS. Ripetibile.

RT. *Related term*, cioè termine associato. Il simbolo può essere sostituito con l'equivalente in italiano, non standard, TA. Ripetibile.

UF. *Use for*, cioè usare per. Il simbolo può essere sostituito con l'equivalente in italiano, non standard, *Usato per* o *Usare per*. Ripetibile.

SN. *Scope note*. SN può essere sostituito con l'equivalente in italiano, non standard, *Nota* o *Nota d'ambito* o *Nota di definizione*. La forma più breve, nel caso che non si voglia utilizzare la forma standard, è quella consigliata in quanto questa è la nota per eccellenza a carattere semantico. In linea di massima non dovrebbe essere ripetibile.

HN. *History note*, nota storica. HN può essere sostituito con *Nota storica*. Fornisce all'indicizzatore i dati fondamentali sull'utilizzo di un termine quando, nel corso della vita del vocabolario, il suo ruolo, significato, struttura o forma, per qualsiasi motivo, subiscano cambiamenti rilevanti; per esempio, quando un termine sostituisce, in tutto o in parte, un altro termine, oppure quando il significato di un termine viene ampliato o ridotto. Per altro la funzione della nota, come è ovvio, non è strutturale: permette di ricostruire la storia del vocabolario e informa l'indicizzatore sull'evoluzione del linguaggio. La nota storica è inserita tra gli elementi applicativi, da non visualizzare per l'utente, ma solo per l'indicizzatore, perché non si ritiene che contenga informazioni rilevanti per le attività di recupero dell'informazione da parte dell'utente finale. Queste ultime dovrebbero essere date nei campi relativi alla relazione di equivalenza o di variante.

Nota sintattica. La *Nota sintattica* è del tutto non standard. E' una nota per l'indicizzatore, che nasce dall'esigenza di prevedere, in un contesto di indicizzazione preordinata, istruzioni per l'applicazione delle regole sintattiche al livello del singolo termine. La struttura e i tipi di nota sintattica sono illustrati più avanti. In linea di massima non ripetibile.

Categoria. *Categoria* (o anche *Faccetta*) in un vocabolario convenzionale ha la funzione di raggruppare i termini, sia quelli completamente strutturati, sia quelli "orfani" (cioè privi di un termine generico) in base ad un principio generale di tipo classificatorio. Il suo impiego è

indipendente dalla struttura del vocabolario. Ovviamente in un vocabolario basato sull'analisi a faccette questo campo assume una funzione classificatoria in senso stretto. Non ripetibile.

Classificazione. Questo campo può servire per attribuire ad un termine la notazione di una classificazione (p.e., il numero interdisciplinare di DDC, o la notazione di una classificazione usata localmente). Non ripetibile.

Fonte1. Individua il repertorio o i repertori, il vocabolario di indicizzazione o la classificazione dal quale un termine è stato estratto o per mezzo del quale è stato validato. Può servire, per esempio, a segnare tutti i termini derivati dal *Soggettario* o dalle singole *Liste di aggiornamento*. Ripetibile.

Variante. Questa proposta è chiaramente non standard e, in parte, deriva dal campo ALT di AAT, anche se qui se ne propone un uso più ampio. Dovrebbero essere inserite qui tutte le varianti di un termine che non sono considerate relazioni di equivalenza e che quindi non devono rispondere necessariamente alla regola del rapporto biunivoco termine-significato: quindi, varianti puramente morfologiche (singolare/plurale, forma aggettivale, etc.), varianti locali, varianti sintattiche (termini impiegati nella stringa in forma diversa da quella contemplata nel vocabolario). Inoltre, questo campo ha la funzione di ricostruire la compatibilità all'indietro del vocabolario, inserendo tutte le forme obsolete del *Soggettario* che abbiano uno o più termini corrispondenti nel nuovo vocabolario. Inoltre, dovrebbe corrispondere alla variante un meccanismo di "switching" automatico (se l'utente cerca con un termine variante gli vengono presentati i termini o i documenti rilevanti) e non essere visualizzata nella voce del vocabolario. Può essere opportuno differenziare in due relazioni diverse questo campo (una puramente morfologica, una di compatibilità). Un'altra caratteristica può essere quella di permettere alle biblioteche, in particolari condizioni, di utilizzare una variante non accettata di un termine come forma preferita, retrocedendo la forma accettata a variante. Ripetibile.

Fonte2. Serve ad individuare l'agenzia catalografica che ha proposto un termine o a "segnare" il sottoinsieme di termini che appartengono a una specifica agenzia. Ripetibile.

Esempi di termini strutturati

Gli esempi sono tratti dal *Soggettario* e dalle *Liste di aggiornamento*. È stato utilizzato, nel primo esempio, l'artificio dell'NT2 ecc. per dare profondità gerarchica, senza dover necessariamente strutturare tutta la gerarchia. .

Va tenuto presente che gli esempi che seguono sono stati costruiti partendo da voci del *Soggettario* e delle *Liste di aggiornamento*, nei limiti del possibile, senza aggiunte. Non compaiono negli esempi quei termini correlati, presenti nel *Soggettario*, che sono stati ritenuti non legati semanticamente, almeno in maniera diretta, al lemma. Aggiunte sono state fatte, ovviamente, per i termini sovraordinati, indispensabili per completare la gerarchia, quando non presenti negli strumenti utilizzati. Infine, in questi esempi il nome della categoria e del TT coincidono, ma non è detto che sia così nella struttura definitiva.

Monumenti

- TT Strutture
 BT Costruzioni
 <secondo la forma>
 NT Archi di trionfo
 <secondo la funzione>
 NT Monumenti commemorativi
 NT2 Monumenti ai caduti
 NT Monumenti nazionali
 NT Monumenti sepolcrali
 NT2 Mausolei
 NT2 Piramidi
 RT Opere d'arte
 RT Edifici
 SN Usare sia nel significato specifico di costruzione consacrata alla memoria di fatti e personaggi importanti; sia nel senso, collettivo, di insieme di costruzioni che testimoniano la storia e l'arte di un luogo, di una nazione, etc., per es., i monumenti di Firenze. È essenziale che il significato di "costruzione" sia sotteso al significato individuato; non si deve cioè considerare monumento ciò che non è costruito (p.e., le tradizioni, gli oggetti, etc., in quanto tali)
 HN Usato nel *Soggettario 1956*; le *Liste di aggiornamento 1986-1998* hanno: Monumenti ai caduti dei campi di concentramento
 Categoria: Strutture
 Classificazione: 725.94 (DDC21)
 Fonte1: *Soggettario 1956*
 Varianti: Monumento
 Identificativo: M000001
 Status del record: Esempio
 Fonte2: Gruppo

Genealogia

- TT Discipline
 BT Discipline storiche
 RT Araldica
 RT Famiglie
 Categoria: Discipline
 Classificazione: 929.1 (DDC21)
 Fonte1: *Soggettario 1956*
 Identificativo: G000001
 Status del record: Esempio
 Fonte2: Gruppo

Monti di pegno

TT Organizzazioni
 BT Istituti di credito
 RT Polizze di pegno
 RT Prestiti bancari su pegno
 UF Monti di piet a
 UF Monti di credito su pegno
 Categoria: Organizzazioni
 Classificazione: 332.34 (DDC21)
 Fonte1: *Soggettario 1956*
 Varianti: Monte di pegno; Monte di piet a; Monti di credito
 Identificativo: M000002
 Status del record: Esempio
 Fonte2: Gruppo

Malattie

TT Processi
 BT Processi patologici
 <*secondo l'agente*>
 NT Malattie parassitarie
 <*secondo il modo di trasmissione*>
 NT Malattie infettive
 NT Malattie ereditarie
 <*secondo gli organi e le parti*>
 NT Malattie cerebrovascolari
 NT Malattie gastrointestinali
 NT Distrofia muscolare
 <*secondo il paziente*>
 NT Malattie infantili
 RT Malati
 RT Patologia
 Categoria: Processi
 Classificazione: 616 (DDC21)
 Fonte1: *Soggettario 1956*
 Identificativo: M000003
 Status del record: Esempio
 Fonte2: Gruppo

Montagne

TT Spazio
 BT Rilievi orografici
 <parti>
 NT Pendici montane
 RT Abitazioni di montagna
 RT Alpinismo
 RT Autostrade di montagna
 RT Comuni di montagna
 RT Ghiacciai
 RT Strade di montagna
 RT Zone montane
 UF Monti

Categoria: Spazio

Classificazione: 551.432 (DDC21)

Fonte1: *Soggettario 1956*

Varianti: Montagna; Monte

Identificativo: M000003

Status del record: Esempio

Fonte2: Gruppo

Rilievi orografici

TT Spazio
 BT Aree geomorfologiche
 NT Altopiani
 NT Bacini montani
 NT Catene montane
 NT Colline
 NT Valli
 RT Orografia
 UF Rilievi montani
 HN Rappresenta il termine generico che raggruppa tutti i tipi di elevazione o depressione orografica; non presente nel *Soggettario 1956*, dove probabilmente era sostituito dal termine che sta per la disciplina (Orografia)

Categoria: Spazio

Classificazione: 551.43 (DDC21)

Fonte1: non in *Soggettario 1956*

Identificativo: R000001

Status del record: Esempio

Fonte2: Gruppo

6. Il corredo sintattico-applicativo

Le funzioni del corredo sintattico-applicativo

Quando il catalogatore si trova a indicizzare documenti su "Responsabilità dei datori di lavoro in materia di assegni familiari", "Appalti di servizi pubblici negli enti locali", "Insegnamento della matematica nelle scuole elementari", può avere dei dubbi o delle incertezze nell'analisi dell'enunciato, nel riconoscimento di alcuni ruoli e, dunque, nell'applicazione delle regole per l'ordine di citazione e nella costruzione della stringa. A quale tipo di struttura sintattica dà luogo un determinato termine (p.e., **Appalti**), quando si trova, nella definizione del soggetto, unito ad altri termini (p.e., **Servizi pubblici, Enti locali**)? Quali funzioni logiche devono essere attribuite ai termini ad esso associati? Conseguentemente, quale ordine di citazione deve essere adottato per la costruzione della relativa stringa?

La nota sintattica aiuta a sciogliere questi dubbi e incertezze, mediante un'istruzione che ha la funzione di guidare l'indicizzatore nell'analisi dei ruoli e nell'applicazione dell'ordine di citazione.

Dubbi e incertezze, casi di complessità, strutture sintattiche particolari possono presentarsi non solo per soggetti specifici che hanno in comune uno stesso concetto (p.e., **Appalti**), ma per un ampio numero di soggetti che hanno in comune una caratteristica tipica di un'area semantica generale. Per esempio, nel campo della letteratura, sono frequenti soggetti che riguardano temi, caratteri, motivi letterari o di un determinato genere letterario o di una singola opera letteraria: queste caratteristiche determinano strutture sintattiche identiche per tutti i soggetti che, per esempio, trattano di un tema letterario in un genere letterario in una determinata lingua, indipendentemente dallo specifico tema, genere, lingua. Analogamente sono tipici della letteratura soggetti che includono "relazioni di fase" (come influssi, confronto ecc.). Il modo di tradurre queste strutture nella stringa di soggetto non si discosta dalle regole generali, la cui applicazione, tuttavia, può essere facilitata da istruzioni contenute sotto una "voce" del manuale (p.e., **Letteratura**), che ha lo scopo di aiutare l'indicizzatore nell'analisi e nell'indicizzazione delle opere di ambito letterario. Nella "voce" del manuale potrà trovare posto anche la trattazione di aspetti terminologici caratteristici di un'area semantica, p.e., la forma corretta per indicare la produzione letteraria, in una lingua, di un gruppo di autori caratterizzati dall'appartenenza ad una etnia o religione o sesso ecc.

Nello studio di fattibilità è stata sviluppata solo la nota sintattica, di cui nei paragrafi seguenti si indicheranno la natura, le funzioni e la struttura.

La nota sintattica

La nota sintattica, che correda alcune voci del vocabolario, svolge una funzione di aiuto nella costruzione delle stringhe di soggetto.

Essa si basa sulla nozione di *valenza sintattica*: la capacità di un termine di creare legami, di stabilire relazioni con altri termini. Questa proprietà si esprime mediante la "cornice" o "schema dei ruoli" associati ad un termine. Si tratta di un metodo coerente con quello che governa le

regole generali, anzi ne è l'applicazione al singolo termine, un'applicazione che in molti casi risulterà ovvia, e dunque non richiederà nessuna istruzione particolare, mentre in altri potrà incontrare dubbi e incertezze, che la nota aiuterà a sciogliere.

La valenza sintattica è propria principalmente dei termini che possono svolgere la funzione di azione, ai quali è associato appunto uno "schema dei ruoli". Per esempio, di un'azione transitiva, come quella espressa dal termine **Insegnamento**, si dice che può avere una relazione con un *oggetto*, un *beneficiario*, un *agente*, uno *strumento*: queste funzioni (oggetto, beneficiario, agente, strumento) formano lo schema dei ruoli del concetto di **Insegnamento**, illustrato nella Fig. 11.

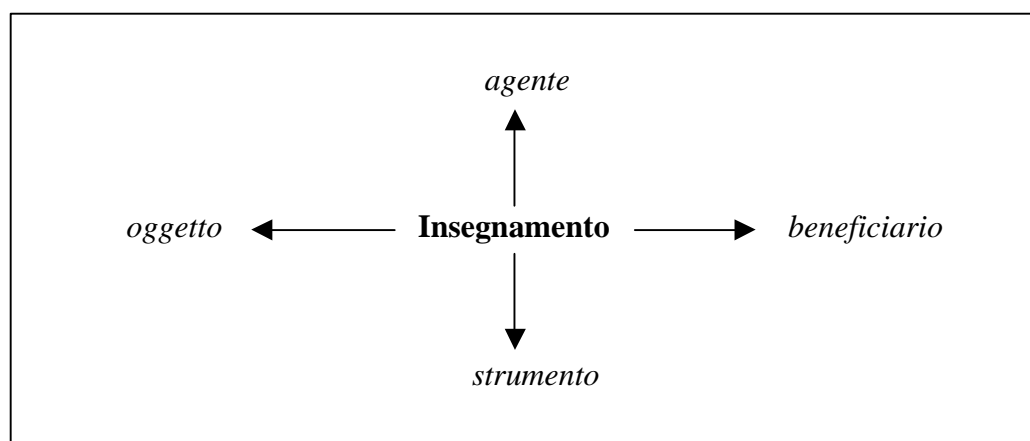


Fig. 11: Schema dei ruoli di Insegnamento

Tuttavia, la valenza sintattica è propria anche di termini che rappresentano *parti* o *proprietà*, in quanto possono stabilire una relazione con un termine che rappresenta l'intero o il possessore.

La nota sintattica svolge una funzione di aiuto, di guida nelle seguenti operazioni:

- a) analisi dell'enunciato di soggetto;
- b) applicazione delle procedure e delle norme sintattiche a singoli termini del vocabolario.

Il punto di partenza dell'analisi consiste nel riconoscimento del concetto di *azione*, se presente: dalla sua valenza sintattica dipendono, infatti, le funzioni assunte dagli altri concetti. È il concetto di azione che determina la struttura sintattica dell'enunciato, definendo i tipi di elementi ad esso associati. Inoltre, il termine che rappresenta il concetto di azione costituisce anche la "testa" dell'enunciato, ossia l'elemento iniziale del sintagma, a partire dal quale si dispongono gli altri elementi. Per questi motivi, la nota sintattica è formulata principalmente per i termini che rappresentano concetti di azione. L'indicizzatore, identificato nell'enunciato il concetto di azione e scelto nel vocabolario il termine che lo rappresenta, è guidato dalla nota nell'identificazione dei ruoli svolti dagli altri elementi e, conseguentemente, nella costruzione della stringa. In mancanza di un concetto di azione, saranno termini che rappresentano proprietà o parti a costituire il punto di partenza dell'analisi. Dunque, anche per questi termini può rendersi utile la formulazione della nota sintattica.

Del resto, sono queste valenze sintattiche (azioni, proprietà e parti) che danno luogo, anche nel *Soggettario*, alla nota "*anche suddivisione di particolari soggetti es.*".

Tuttavia, la nota non si troverà nel corredo di tutti i termini che rappresentano azioni e proprietà o che possono essere considerati come parti di altri concetti, ma solo di quelli che

richiedono un'interpretazione della loro valenza sintattica o una chiarificazione delle strutture sintattiche che stabiliscono. Questa situazione può verificarsi, per esempio, quando ad uno stesso termine si attribuiscono, a seconda del contesto, più valenze.

Per esempio, il termine **Appalti** funge da azione transitiva a tre argomenti e, dunque, può avere un oggetto, un agente e un beneficiario, come in "Appalti dei servizi delle biblioteche comunali alle cooperative di lavoro" o in "Appalti di servizi pubblici da parte degli enti locali". Mentre, è da interpretare come una proprietà quando sia espresso l'appaltatore, ma non l'oggetto dell'appalto, come in "Appalti delle Ferrovie dello Stato". Il termine **Responsabilità** rappresenta una proprietà e, come tale, può essere unito a termini che rappresentano il possessore, per esempio in "Responsabilità dei datori di lavoro". Tuttavia, in certi casi, può essere diretto verso un oggetto ed avere anche un agente, analogamente ad un'azione, come in "Responsabilità dei datori di lavoro in materia di assegni familiari".

Un'altra funzione accessoria della nota può essere quella di segnalare i cambiamenti nella costruzione della stringa di soggetto, come indicato negli esempi seguenti dalla segnalazione "*precedentemente*".

La struttura della nota sintattica

La struttura della nota può essere illustrata con i seguenti esempi:

Appalti

TT Strumenti
 BT Contratti
 NT Appalti comunali
 NT Appalti edilizi
 NT Subappalti
 RT Appaltatori
 RT Capitolati di appalto
 RT Contratti di appalto

Nota sintattica: **Elem. trans.** *Segue il termine che rappresenta l'oggetto/meta e precede il termine che rappresenta il beneficiario, lo strumento o l'agente, p.e.* Biblioteche comunali - Servizi - Appalti alle cooperative di lavoro; Servizi pubblici - Appalti degli enti locali [*precedentemente* 1. Lavori pubblici - Appalti, 2. Enti locali - Funzioni in materia di lavori pubblici]; Lavori pubblici - Appalti - Ruolo delle associazioni temporanee di imprese - Legislazione [*precedentemente* 1. Associazioni temporanee di imprese - Legislazione, 2. Lavori pubblici - Appalti - Legislazione]; Lavori pubblici - Appalti - Impiego dell'autocertificazione - Legislazione [*precedentemente* 1. Lavori pubblici - Appalti; 2. Autocertificazione - Legislazione]. **Parte/Propr.** *Quando non sia citato un oggetto/meta, segue il termine che rappresenta il possessore (appaltatore), p.e.,* Ferrovie dello Stato - Appalti.

Categoria: Strumenti
 Fonte1: *Soggettario 1956*
 Identificativo: A000002
 Status del record: Esempio
 Fonte2: Gruppo

Malattie

TT Processi

BT Processi patologici

<secondo l'agente>

NT Malattie parassitarie

<secondo il modo di trasmissione>

NT Malattie infettive

NT Malattie ereditarie

<secondo gli organi e le parti>

NT Malattie cerebrovascolari

NT Malattie gastrointestinali

NT Distrofia muscolare

<secondo il paziente>

NT Malattie infantili

RT Malati

RT Patologia

Nota sintattica: **Parte/propr.** Segue il termine che rappresenta il possessore (singoli individui, gruppi di persone, organismi e loro parti), p.e., Leopardi, Giacomo - Malattie; Adolescenti - Malattie; Gatti - Malattie; Apparato digerente - Malattie; Bambini - Sistema nervoso - Malattie [precedentemente Sistema nervoso - Malattie - Infanzia]; Laringe - Vasi sanguigni - Malattie.

Categoria: Processi

Classificazione: 616 (DDC21)

Fonte1: *Soggettario 1956*

Identificativo: M000003

Status del record: Esempio

Fonte2: Gruppo

Affreschi

TT Oggetti

BT Dipinti

Nota sintattica: **Parte/propr.** Segue il termine che rappresenta l'intero (classe di edifici o singolo edificio) o il possessore (autore), p.e., Ville Venete - Affreschi; Abbazia di Grottaferrata - Affreschi; Benozzo Gozzoli - Affreschi - Castelfiorentino. Quando siano citati entrambi i concetti (intero e possessore), segue il termine che rappresenta l'intero e precede il termine che rappresenta il possessore, p.e., Convento di San Marco, Firenze - Affreschi di Beato Angelico - Restauro [precedentemente Beato Angelico - Affreschi - Firenze - Convento di San Marco - Restauro].

Categoria: Oggetti

Classificazione: 751.73 (DDC21)

Fonte1: *Soggettario 1956*

Identificativo: A000001

Status del record: Esempio

Fonte2: Gruppo

Commercio

TT Attività
 UF Scambi commerciali
 BT Attività economiche
 NT Commercio internazionale
 NT Scambio

Nota sintattica: **Parte/Propr.** *Segue il termine che rappresenta il possessore (gruppo sociale o individuo), preceduto da un elemento contestualizzante, p.e. Costantini <famiglia> - Attività economiche: Commercio [precedentemente Costantini <famiglia> - Attività commerciale]; non può essere citato come Parte/Propr. di un luogo, quindi, p.e. Commercio - Milano - 1952-1961 [precedentemente Milano - Commercio - 1952-1961]*

Categoria: Attività
 Classificazione: 380.1 (DDC21)
 Fonte: *Soggettario 1956*
 Identificativo: C000001
 Status del record: Esempio
 Fonte2: Gruppo

Nel corso della sperimentazione è stata formulata anche "una nota sintattica generale" sotto termini che indicano una categoria o, comunque, una classe molto generale. La peculiarità di questo tipo di nota consiste nel fatto di riferirsi non tanto al termine sotto il quale compare, ma a tutti i termini appartenenti a quella categoria o classe, che condividono le valenze sintattiche indicate nella nota. La sua funzione è quella di riproporre, sotto un'altra angolazione, le norme sintattiche generali.

Attività

NT Attività artistiche
 NT Attività economiche
 NT Attività religiose
 NT Attività ricreative
 NT Attività tecniche
 NT Giochi

Nota sintattica generale: **Elem. trans.** *I termini della faccetta Attività seguono il termine che rappresenta l'oggetto/meta e precedono il termine che rappresenta il beneficiario, lo strumento o l'agente, p.e. Catechismo - Insegnamento agli adolescenti; Alimenti - Conservazione - Impiego dell'acido para-ossi-benzoico; Paesi in via di sviluppo - Assistenza economica della Comunità europea.*
Parte/Propr. *Quando non sia citato un oggetto/meta, i termini della faccetta Attività seguono il termine che rappresenta il possessore (agente), p.e. Anziani - Attività ricreative; Popoli primitivi - Agricoltura.*

Categoria: Attività
 Fonte1: Struttura
 Identificativo: A000003
 Status del record: Esempio
 Fonte2: Gruppo

Come mostrano gli esempi, la nota è costituita dalle seguenti parti:

1. indicazione della valenza sintattica, mediante una sigla in neretto corsivo (*Elem. trans.*, *Parte/Prop.*);
2. istruzione, costituita da: indicazione della posizione del termine nella stringa, introdotta dalle espressioni "segue ..." e/o "precede ..." (in corsivo); indicazione dei ruoli associati o "schema dei ruoli" (in corsivo).
3. esempi, introdotti dalla sigla "p.e." (in corsivo), separati dal segno del punto e virgola;
4. nota sui cambiamenti (in parentesi quadre), introdotta dall'espressione "precedentemente" (in corsivo), seguita da un esempio di stringa costruita alla maniera della precedente edizione, che può consistere in un diverso ordine di citazione dei termini o in un diverso modo di esprimere il medesimo soggetto, p.e., più stringhe al posto di un'unica stringa coestesa.

Confronto con la nota del *Soggettario*

Il modello sintattico tradizionale è basato sulla distinzione tra voce principale e suddivisioni, distinzione che a sua volta poggia su criteri lessicali, semantici o pragmatici. In che modo si riflette questo modello nella struttura sintattica del *Soggettario*? La presenza del termine come voce principale segnala la sua funzione preminente nella rappresentazione del soggetto. L'elenco esemplificativo delle suddivisioni che compaiono sotto la voce principale indica alcuni aspetti o parti limitatamente alle quali può essere trattato il soggetto espresso dalla voce principale. La nota segnala l'impiego del termine come suddivisione, per esempio:

Appalti

...; *anche suddivisione di particolari soggetti es. Lavori pubblici - Appalti*

La nota ci informa che la voce **Appalti**, oltre a fungere da voce principale, eventualmente seguita da altre voci che ne specificano un aspetto o una parte, quali le suddivisioni indicate nell'elenco (**Controversie**, **Onerosità**, **Prezzi** ecc.), può comparire, come suddivisione, sotto particolari soggetti, rappresentati dall'esempio, ad indicarne un aspetto.

Questa struttura, incentrata sulla distinzione voce principale/suddivisioni, presenta, dunque, per un termine i seguenti elementi:

- a) la sua funzione nella rappresentazione del soggetto, che varia a seconda che sia impiegato come voce principale o come suddivisione;
- b) la sua posizione nella stringa: in quanto voce principale, precede termini che fungono da suddivisioni; in quanto suddivisione, segue termini che fungono da voce principale;
- c) la funzione degli altri termini associati, nel ruolo di suddivisioni o di voci principali.

La nota proposta per il nuovo *Soggettario* esplicita e reinterpreta questi elementi, alla luce dei nuovi criteri logico-sintattici e nel contesto di una preordinazione di tipo sintetico, segnalando:

- a) la funzione (intesa come *valenza sintattica*);
- b) la posizione nella stringa (intesa come *ordine di citazione*);
- c) le funzioni dei termini associati (intese come *ruoli*).

Agosto 2004