

TeKaMoLo ovvero le lingue come KOS, e viceversa

Claudio Gnoli*

Tra i sistemi per l'organizzazione della conoscenza (KOS), soggettari, tesauri e classificazioni sono stati chiamati nella seconda metà del Novecento anche linguaggi di indicizzazione, e ancora oggi in francese vengono detti *langages documentaires*. Queste espressioni ci ricordano che, dopotutto, i KOS sono un tipo particolare di linguaggi, parzialmente artificiali perché destinati a particolari scopi, ma pur sempre dotati di molte caratteristiche comuni alle lingue parlate.

D'altra parte, già le lingue sono il primo dei modi con cui gli umani organizzano la conoscenza, suddividendo l'esperienza in un numero finito di concetti espressi da parole. Parlare è infatti prima di tutto assegnare a ciascun fenomeno un nome (*febbre, poiana, assemblea condominiale...*), il quale appartiene ad una rete implicita di gerarchie e associazioni. Studiata dalla semantica lessicale (Jezek 2005), essa assomiglia parecchio ai reticoli più formalizzati di concetti e relazioni che sottendono una classificazione o un'ontologia. I sostantivi della lingua vengono poi collegati con verbi, casi, preposizioni e congiunzioni che esprimono relazioni, analoghe alle combinazioni di concetti nei KOS.

Queste osservazioni suggeriscono tra l'altro che l'organizzazione della conoscenza si possa applicare anche alle reti semantiche studiate dalla linguistica, e magari ad una didattica basata sul riconoscimento delle strutture grammaticali come quella di M. Michela Sebastiani ed Emanuela Valenzano (2016): direzioni che probabilmente meritano di essere esplorate più a fondo.

Quello che nelle lingue naturali non c'è, ed è il motivo per cui risulta necessario sviluppare dei KOS, sono soprattutto tre cose:

(1) il controllo terminologico, che elimina le ambiguità dovute alle sinonimie e alle omonimie nella maggior parte dei KOS (con l'eccezione delle

* Biblioteca della scienza e della tecnica, Università degli Studi di Pavia, Pavia, Italia. claudio.gnoli@unipv.it.

parole-chiave sciolte o tag) associando ciascun concetto ad un termine univoco e offrendo opportuni rinvii (*bozzaghi*: vedi *poiane*);

(2) relazioni paradigmatiche esplicite, che permettono di percorrere a piacere le reti di concetti verso l'alto (termini più generici, classi superiori), verso il basso (termini più specifici, sottoclassi) o lateralmente (termini associati, faccette);

(3) un ordine sistematico, implicito o controllato da una notazione, che permette di visualizzare e scorrere i concetti in modi coerenti e utili, anziché secondo le bizzarrie dell'alfabeto.

Queste tre componenti permettono di passare dalle lingue naturali a quei linguaggi per scopi particolari che sono appunto i KOS. Ma la natura linguistica di questi ultimi rimane, e sebbene spesso trascurata può riemergere in diversi aspetti: ben lo sapevano autori come Maurice Coyaud (1966), W.J. Hutchins (1975) o Jacques Maniez (1987) sul quale è apparsa recentemente una voce dell'enciclopedia ISKO.

Anche Gérard Cordonnier (1951), teorico delle classificazioni sintetiche, si muoveva a cavallo fra lingue artificiali e naturali immaginando un sistema capace di funzionare come una lingua parlabile, caratteristica vantaggiosa anche per memorizzarne più facilmente la notazione. A questo scopo occorrerebbe che i simboli, che solitamente possono consistere di qualunque stringa di caratteri, risultassero pronunciabili. Questa funzione è stata realizzata da Douglas e Joy Foskett nella London Education Classification, prevedendo simboli che siano sempre sequenze consonante-vocale/i-consonante: così *RidMenLim* significa "metodo diretto, francese, scuola secondaria". Un'alternativa è progettare liberamente la notazione ottimizzandola per le sue consuete funzioni di ordinamento, e solo successivamente associare a ciascun simbolo un fonema o una sillaba che risulti pronunciabile in qualsiasi combinazione, perché ad esempio termina sempre per vocale: così nella Integrative Levels Classification (ILC) la faccetta *sb2e* "cibi, ristorazione" si può pronunciare /sabantœ/. Non a caso questo KOS è stato ispirato anche dalla precedente esperienza di sviluppo di una lingua logica chiamata liva. Simili problemi sono stati affrontati nel passato dagli sviluppatori di lingue ausiliarie internazionali, come l'esperanto, o di lingue filosofiche quale quella del vescovo seicentesco John Wilkins, che ha aspetti molto vicini alle classificazioni bibliografiche come già notò il grande Brian Vickery (1953).

Come abbiamo accennato, le relazioni che nelle lingue naturali sono espresse dai diversi tipi di complementi corrispondono sostanzialmente alle faccette e ai ruoli dei KOS sintetici: si può rendersene conto leggendo la guida Nuovo Soggettario (Biblioteca Nazionale Centrale di Firenze 2021). In questi KOS infatti ricorrono certe categorie fondamentali (parti, proprietà, processi, agenti, spazio, tempo...) che hanno degli equivalenti nei sistemi di casi di mol-

te lingue flessive, e meglio ancora nei casi profondi identificati dal linguista Charles Fillmore.

Talvolta nelle lingue naturali si ritrova perfino il principio classificatorio dell'ordine di citazione delle categorie: come le faccette vanno elencate secondo varianti della famosa formula PMEST di Ranganathan, così scrivendo in tedesco occorre seguire la formula equivalente TeKaMoLo (tempo, causa, modo, luogo) nel disporre i complementi indiretti e gli avverbi. Sono ammesse eccezioni solo quando si vuole enfatizzare uno dei complementi: una possibilità che ha senso anche nei KOS, dove promuovere una faccetta più vicino all'inizio della stringa ha l'effetto di renderla più rilevante ai fini dell'ordinamento sequenziale delle voci: così l'ordine standard tipo-parte-processo di "cibi, con uova, fritti" si può modificare in "cibi, fritti, con uova" nel caso che in un determinato documento la faccetta processuale risulti più rilevante.

Già Paul Otlet immaginava che la sua Classificazione decimale universale potesse essere impiegata linguisticamente, per esempio nei telegrammi, costruendo espressioni con classi qualificate da suffissi per soggetti, verbi e complementi come *35.633⁰⁰¹*; *629.12313^{063.043}*; *35.521.6⁰⁰²*; *615.935⁰⁶⁴*. che vorrebbe significare "il servizio sanitario trasporterà con imbarcazioni a vela i lavoratori volontari colpiti da dissenteria" (Dousa 2010; Institut international de bibliographie 1906)¹.

Una differenza importante rispetto alle lingue naturali è che solitamente i KOS non esprimono affermazioni, del tipo "il cibo fritto aumenta il colesterolo", bensì classi senza tempo di soggetti, che vengono infatti espresse al plurale: "cibi fritti", dove tuttalpiù potremmo esplicitare il significato sottinteso "questo documento tratta di cibi fritti". Nell'esempio linguistico invece il tema ossia ciò di cui si parla ("il cibo fritto") è completato da un rema ("aumenta il colesterolo"), cioè da un'affermazione relativa ad un fatto osservato nella realtà: oltre alla classe si esprime dunque una sua relazione denotabile concretamente (deissi) perché si verifica qui o altrove o è relativa a un individuo particolare ("questo cibo") magari dotato di un nome proprio ("la torta di compleanno di Diana"). Nell'indicizzazione solitamente si intende esprimere soltanto il tema, ritenendo che sia esso ad interessare gli utenti che consultano gli indici. Tale assunto è stato però messo in discussione da un articolo di Bella Hass Weinberg (1988), secondo cui anche le specifiche affermazioni che un testo fa sul proprio tema possono essere rilevanti per gli studiosi.

Fra le relazioni espresse nelle classificazioni, Rick Szostak ritiene che in una prospettiva interdisciplinare le più importanti per gli utenti siano quelle di tipo causale. In fin dei conti, la conoscenza cerca innanzitutto di identificare i nessi causali fra i fenomeni: da che cosa un fenomeno è influenzato, da che cosa è derivato, ecc. Ad esempio, nella ILC è possibile costruire espressioni

¹ Sono grato a Tom per le segnalazioni.

come *fps85m(30sb(5oy))* che letteralmente significa “steroli, aumento, causato da cibi, friggere” e può essere recuperata sia cercando la notazione per “steroli” che cercando quella per “friggere”.

Per la verità, questo esempio non è ancora propriamente equivalente ad una frase del linguaggio naturale: per renderlo tale occorre aggiungere un deittico ossia un riferimento variabile alla realtà esterna, come *-V* “questi” o *-Y* “effettivamente”: si può allora avere *fps85m(30Ysb(5oy))* “steroli, aumento, è causato da cibi, friggere”. Di questo passo potremmo realizzare gli auspici di Wilkins e di Otlet e metterci a parlare in una lingua classificata.

Riferimenti bibliografici

- Biblioteca Nazionale Centrale di Firenze. 2021. *Nuovo soggettario: guida al sistema italiano di indicizzazione per soggetto*, 2nd ed. Roma: Associazione Italiana Biblioteche; Firenze: Biblioteca nazionale centrale di Firenze. <https://doi.org/10.53263/9788878123465>.
- Cordonnier, Gérard. 1951. “Classification, classement, rangement, et sélection.” *CNOF: revue mensuel de l'organisation* 25 (5-6): 19-31.
- Coyaud, Maurice. 1966. *Introduction à l'étude des langages documentaires*. Paris: Klincksieck.
- Dousa, Thomas. 2010. “IORG Paul Otlet the UDC and the Origins of the Concept of A.” YouTube video. youtu.be/hQaF5xJTgRc.
- Hass Weinberg, Bella. 1988. “Why Indexing Fails the Researcher.” *The Indexer: The International Journal of Indexing* 16: 3-6.
- Hutchins, William John. 1975. *Languages of Indexing and Classification: a linguistic study of structures and functions*. Stevenage, Eng: P. Peregrinus.
- Institut international de bibliographie. 1906. “Bulletin de l'Institut international de bibliographie.” no. 11: 92-99.
- Jezek, Elisabetta. 2005. *Lessico*. Bologna: Il Mulino.
- Maniez, Jacques. 1987. *Les langages documentaires et classificatoires*. Paris: Les éditions d'organisation.
- Sebastiani, M. Michela, e Emanuela Valenzano. 2016. *Educare alla lettura con i marcatori visivi*. Roma: Aracne.
- Vickery, Brian C. 1953. “The significance of John Wilkins in the history of bibliographical classification.” *Libri* 2: 326-43.