

## シソーラス

会誌編集部

### I. シソーラスとは

シソーラスって、検索をする上で一応理解して使ってはいるが、もう一つよくわかった気がしない人はいないだろうか。ディスクリプタ、非ディスクリプタ、階層構造、上位語 (Broader Term : BT)、下位語 (Narrower Term : NT)、カテゴリー、などなど。

表1はシソーラスの階層構造である。

表1 シソーラスの階層構造

- H 1 : 自然科学
  - H 1-10 : 化学
    - H 1-10-10 : 生化学
      - ：
  - H 1-40 : 生物化学
    - H 1-40-10 : 解剖学
      - ：
- H 2 : 保健医療業務
  - H 2-10 : Evidenced-Based Practice
  - H 2-20 : 医学
    - H 2-20-10 : アレルギーと免疫学
      - ：
  - H 2-100 : 看護
    - H 2-100-10 : 看護学
      - ：

「H」はMeSHや医中誌Webで使用している15のカテゴリーの1つで、医中誌Webでは「学問分野と専門分野」を示している。それらが下に向かって樹木のように枝分かれしていく。でも実は図書館員なら、これはすでに完璧につかひこなせている。

表2 NDC図書分類

- 490 : 医学
  - 491 : 基礎医学
    - ：
  - 494 : 外科学
    - 494.11 : 外科解剖学
      - ：
    - 494.7 : 整形外科学
      - 494.74 : 骨折
      - 494.75 : 脱臼
      - 494.76 : 捻挫

表1と表2はまったく同じ階層構造である。シソーラスはNDCなど図書分類と同じ構造なのである。

「ディスクリプタ」とは統制語(Controlled Term)のことだが、NDCではそれが言葉でなく数字になっているので、直ちにそれが何を表すかはわからない。それだけである。分類では「分類番号」や「分類記号」と呼ばれるものが、シソーラスだと「ディスクリプタ」という大層な言葉になる。分類と件名は図書館学で習う際に並列されるが、シソーラスはそうではないので、別モノのように思うけれども、これらはすべて「索引言語システム」の中の「統制語システム」に含まれる、同じ仲間である<sup>1)</sup>。統制語システムでは、上に述べたように同義語・類義語は何らかの名辞や記号で代表させられる。その反対は非統制語システムで、自然語(Natural Term)で検索するもの。雑誌記事索引がそうで、「癌」を検索する際、「悪性腫瘍」、「悪性新生物」など関係する単語を網羅的に検索しなくてはならず、統制語がないと検索の際は不便である。

## II. 図書分類とシソーラス

分類とシソーラスの歴史を見てみると<sup>2)</sup>、シソーラスという用語が最初に用語のリストの名称に使用されたのは、1852年ロジェー(P. M. Roget)の“Thesaurus of English Words and Phrases”の初版である。でもこれは作文の参考用に用語間の意味的關係を品詞の種類別に示したものであった。現在情報検索の分野では、シソーラスとは、文献の主題索引・検索用に統制された用語のリストを意味する。シソーラスという用語が後者の意味で使用されはじめたのは比較的最近であり、1950年代にIBM社のルーン(H. P. Luhn)が使用したのが最初である。一方分類は、DDC(デューイ十進分類法)が1876年に出ていて、これが現在普及している分類法の最初期のものとされているから、分類の方が古いと言える。

シソーラスが新しいということは、分類とシソーラスとの違いを考えればわかる。シソーラスと何が違うかと言えば、シソーラスでは主題以外のキーワードもたくさん取り上げられる。一方、分類は主題を1つにするのを基本とする。図書のラベルに与えられる分類は物理的に1つに限られるからである。これは主題分析法において「要約化」と呼ばれる。分類の規則には「○○と○○」というタイトルの場合、どういう分類を付与するかなどという決まりが細かく決められていたことを思い出す人も多いだろう。分類あるいは件名においては特に複数のテーマが示されていることもあるが、それらの表現の仕方がきちっと決められている。なぜかといえば、カード目録において分類や件名では与えられる主題を1列に並べる必要があるから、同じ主題のものを同じ場所におかないと意味がない。一方、シソーラスは、各キーワードの順番なんてまったく気にする必要がない。これは、コンピュータのデータベースで使われることを念頭においているからである。すなわち目録や件名とシソーラスの違いは、カード目録とコンピュータのデータベースへのそれぞれ相性によるものなのである<sup>1)</sup>。

では実際にNDCとシソーラスを比べてみよう。NDCでは整形外科学は「494.7」である。494.7という「ディスクリプタ」のもとに、整形外科関係の図書が並ぶ。「非ディスクリプタ」というのは同義語、異表記語である。同じ物事を違う言い方で言ったり、表記が揺れたりすると、さきほどの「癌」の場合のように検索に支障を来す。この場合「整形外科学」や「整形外科」、「Orthopaedics」といった書名に現れた「整形外科学」を指す個々の言葉にあたる。ディスクリプタと非ディスクリプタとの関連づけについては、相関索引(表3)がそれにあたる。相関索引を調べれば、個々の単語に対する分類がわかるから、すなわち非ディスクリプタからディスクリプタを探すことができるというわけである。蔵書検索する際、「整形外科」という言葉を入力して検索し、ヒットした検索結果から「494.7」がある書架へ行けば、実際多くの用は済むだろうが、別置図書も考慮して精密に検索するなら494.7で再度分類検索を

表3 関連索引の例

税	金	345
聖	具 (キリスト教)	196.4
整	経 (織物工業)	586.76
整形学校	(教育)	378.3
整形外科学		494.7
生計費		365.4

するだろう。これはまさにシソーラス検索と同じ作業である。

また階層構造であるが、NDCでは「整形外科学」の上に「外科学」があって、その上に「医学」がある。下には「骨折」も「脱臼」も「捻挫」もある。PubMedや医中誌 Webにある下位語を含む検索(エクスプロード検索)はシソーラス用語の下位語を含んだ検索であるが、これも通常の蔵書検索において、「整形外科学」の494.7で分類検索した場合、494.74の「骨折」も494.75の「脱臼」もヒットするのが普通で、ヒットしてくれないと困る感じも実感できるだろう。こういった上位・下位の概念だけでなく、NDCにも「を見よ参照」や「をも見よ参照」がある。

さらにMeSHや医中誌 Webでは副標目が使える。特に医中誌 WebではVer.5から付いた機能であるが、副標目(Subheadings)というのは、特定のディスクリプタについて「診断」や「治療」など組み合わせ可能なリストが表示され、それらによって絞り込めるようになっている。この副標目についても、NDCには助記表というのがある。02が歴史、03が辞典、08が叢書というアレである。地理区分や言語区分なんて付与してよいディスクリプタが限られているところまでよく似ている。助記表は独立した検索手段にはなっていないが、これはコンピュータなら可能で必要性の問題にすぎないだろう。

### Ⅲ. シソーラスの課題

それでもMeSHや医中誌 Webのシソーラスが、やっぱりNDCと違うように感じるとしたら、その理由は、シソーラスは中身の言葉が難しいからだと思う。シソーラスに出てくる言葉の専門的なことと言ったら、NDCの比ではない。NDCのごく一部の分野を対象にあれだけ詳細につくっているから、当然のことである。結局、検索にあたっては検索技術だけでなく、対象となる専門知識に踏み込む必要があり、これは専門図書館員がもつ課題であろう。

他にもシソーラスがたくさん言葉を細かく管理することにより生じている問題点がある。

例えばJSTの「シソーラス用語インデックス」で「果実飲料」で検索すると、「トマトジュース」はヒットしない<sup>3)</sup>。トマトジュースは「野菜製品」の下の「野菜ジュース」の下位に位置づけられているからである。「トマトジュース」を調べたい場合はそれにあたるディスクリプタを確認するだろうが、「果実飲料」を調べていて、そこからトマトジュースが排除されているということまで気が回るかどうか。上でも述べたとおり、一つの言葉に付随する多くの情報を、検索者はやはり汲み取れないといけないうし、システム全体に目を配る必要もあるということが言える。

それから、よく言われるのが、新語に対応しにくいという欠点である。シソーラスが詳細であればあるほど、改訂が大変な作業になる。

シソーラスのもつ階層構造についても問題点が指摘されている<sup>4)</sup>。AとB、2つの概念で、Aの中にBが含まれるが、Bが必ずしもAでない場合がある。こういうBがAの下位語になると、かけあわせた場合の検索結果に影響を及ぼすという。これは検索漏れを防ごうとするとノイズが増大するというジレンマの問題で、やはり詳細な検索システムにとって避けられないだろう。

参考文献

- 1) 大柴忠彦. 主題目録法. In: 日本図書館協会図書館ハンドブック編集委員会編. 図書館ハンドブック 第6版 補訂版. 東京: 日本図書館協会; 2010. p. 318-33.
- 2) 丸山昭二郎ほか編. 情報アクセスのすべて 増補改訂版. 東京: 日本図書館協会; 1992. p. 88.
- 3) 情報サロン味岡. 「シソーラス」と、あなたの「観点や解釈」との「ズレ」を修正しながら検索する. [引用 2012-06-09].  
<http://homepage3.nifty.com/salon-ajioka/column78.htm>
- 4) 武田宣之: シソーラスの階層構造における問題点 — 情報検索の結果に及ぼすその影響 —. オンライン検索. 1994; 15(4): 183-6.

(文責: 増田 徹/藍野大学中央図書館)