

モノと情報班

博物館資料データベースのためのシソーラス構築の試論 - ラオス北部・中国雲南省における棒締頭絡の調査から -

小島摩文（鹿児島純心女子大学）

キーワード：博物館資料 データベース シソーラス 棒締頭絡 馬具

Thesaurus construction for Museum Database
: A Case Study of Wooden Hackamores in Northern Thailand and Northern Laos

KOJIMA Mabumi (Associate professor, Kagoshima Immaculate Heart University)

Keywords: Museum Database, Wooden hackamore, Research agenda, Thesaurus for research

要旨

本報告は、本プロジェクト「モノと情報」班におけるデータベース構築に向けた作業のうち、検索に関わるシソーラス部分について、個人的な提言をおこなうものである。特に博物館資料、すなわち物資文化資料に関するシソーラスについて、具体的に「棒締頭絡」という馬具を例にとりながら提言をおこなう。また、データベース、シソーラスともにその理論的な面には言及せず、もっぱら利用者からみた便宜性と整合性について論じたい。

はじめに

データベースについても、シソーラスについてもこれまでに膨大な論文が存在している。とくにシソーラスはコンピュータ化、データの電子化以前からの歴史がある。報告者は原野農芸博物館（鹿児島県旧住用村 [現奄美市]）に勤務していた 1994 年から 1997 年にかけて博物館資料の台帳作りの必要性からデータベースに、そしてやがてシソーラスに興味をもった。しかし、博物館学の分野では扱う資料が多岐にわたるためであろうが、データベース理論に関する議論は活発でなく、その多くを図書館学を援用するかたちでおこなわれていた。また、シソーラスについては、図書館学での蓄積にくらべると博物館学のそれは貧弱であった。

また、文化人類学の分野では、HRAF (Human Relations Area Files) というデータベースがある。HRAF はエール大学 (Yale University) の G. P. マードックらによって 1949 年に開発・研究された「データベース」で、世界の様々な民族の社会や文化について書かれた文献（単行本、論文等）のすべてのページの内容を専門家が独自の分類方法を使って分析し、インデックスを付したファイル資料である。もともとは印刷（コピー）資料で、インデックスをもとに、カードを検索すると、そこに源資料のページのコピーかタイプがあり、原文に当たることができるシステムである。現在では電子化され eHRAF としてウェブ上でも検索できるようになっている。

HRAF では、OWC (Outline of World Cultures) という地域による分類、OCM (Outline of Cultural Materials) という項目別の分類の二つの分類インデックスによる検索が可能である。これに加えて、新しいウェブから検索できる eHRAF ではキーワードによる検索が可能になっている。

本報告との関連で言えば、馬については OCM の 230 の Animal husbandry からたどっていくことになる。

OWC も OCM も一種のシソーラスといえバシソーラスだが、分類に重点が置かれており、やはりインデックスである。インデックスが対応する本文のキーワードと完全に一致するわけではない、どのインデックスを振るかは実は単語と言うよりもパラグラフごとの主題によって決められている。HRAF の最大の特徴だと思われるが、チェックする人間の判断によってインデックス付けをしていることの利点であろう。したがって厳密な意味での

シソーラスを必要としない理由でもある。OCMの中から関連するいくつかの項目についてインデックス番号を控え、ファイルを一覧しながらこの番号を探していけばいいのである。このことはまた同時に欠点にもなる。探している対象そのものに行き着くのに時間がかかるという点である。しかし、インデックス化する時間と、インデックス化されていない文献からその対象を探し出すことを考えるとHRAFがいかに優れたシステム化が理解できる。

本報告では、文献の中から必要なデータを探し出すシステムとしてのHRAFを出発点としながら、さらに推し進めて、文化人類学・物質文化の調査・研究のヒントとなるシソーラスの仕組みを考えてみたい。

一般に、シソーラスを構築する際には「等価関係」「階層関係」「連想関係」などの概念を使って用語を関係付けている。これらを「基本関係」と呼んでいる。

等価関係には「同義語」「準同義語」などの概念がある。同義語には、例えば「異形綴り (Variant Spelling)」がある。日本語では「馬」の異形綴りとして「ウマ」などがあげられる。「準同義語」は例えば「馬」に対して「牛」などのように同義語ではないが検索する際に網にかかった方が良い語が選ばれる。俗に言う反対語なども準同義語になる。階層関係には上位概念と下位概念があり、上位概念は類あるいは全体を表し、下位概念は一要素、あるいは一部分を表す。例えば上位概念「家畜」に対して下位概念「馬」などである。連想関係の定義は難しいが、乱暴に言えば等価関係にも階層関係にも分類しないが関連のある語として収録すべき語、ということになる。

こうした一般的なシソーラスの構造は上に示したような基本関係によって構成され、選ばれた語はそれぞれの関係の中で整合性があるか、注意深く選ばれている。つまり対象となっている語を分析し、その事物、事象と関連のある語だけが選び出され関連づけられているのである。

しかし、こうしたやり方には避けたい欠点がある。それは現在わかっていることしか関連づけられないということである。いわば帰納的に関係が見いだされていて、わかっていることしか分類できないという限界があるということである。

これに対して報告者は、演繹的なシソーラスを提案したい。ここで演繹的とか帰納的という用語を使うことは誤解を招くかも知れない。しかし他によい言葉も見つからないので、とりあえず、この場では帰納的方法を現実の事物や事象を分析して語の関連をみつける方法と定義し、演繹的方法とは現実に関連はないように見えるが理論的には想定しうる関連も対象語彙と関連のある語として採用していく方法と定義しておく。

関連しない語も関連語として採用する演繹的なシソーラスの試み

演繹的なシソーラスとは、現在の知識では関連してないと考えられる語も関連語として取り入れているシソーラスである。単に検索だけのためのシソーラスであれば関連していない語は単なるノイズでしかない。しかし、シソーラスをある事物・事象をみる際の手がかりとして利用しようとした場合、演繹的なシソーラスが有効である。

具体例を「馬」の関連語で見えていきたい。まず、試みとして、「馬」に関連する語を拾い集めてみた(図1)。例えば「制御具」の項目を見てほしい。「頭絡 八ミ 首輪 鼻輪(鼻環) 胴輪」と並んでいる。このうち現在の常識でいえば馬と関連するのは頭絡と八ミである。首輪は犬や猫、山羊などの制御具、鼻輪は牛やラクダ、熊などの制御具、胴輪は犬やブタなどの制御具である。

演繹的シソーラスでは「馬」と「制御具」を関連付けた段階で「頭絡 八ミ 首輪 鼻輪(鼻環) 胴輪」すべてが関連づけられる。一般的なシソーラスでは、ここで統制して「頭絡 八ミ」だけに整理してしまう。そうしないと検索の精度が上がらないからだ。

しかし、制御具として「頭絡 八ミ 首輪 鼻輪(鼻環) 胴輪」があげてあると、調査する際に、馬の鼻輪はないのだろうか、あるいは何故ないのかを質問項目としてあげることができる。私は沖縄の調査の際になぜ沖縄の馬は棒締頭絡なのかという問いと同時になぜ馬は鼻輪では制御できないのかについても質問してみた。多くの答えは昔からそうしているからというものであったが、数名から「馬の鼻は切れやすいから鼻輪ではもたない」と聞かされた。馬の鼻は切れやすいという事実(畜産の研究者も同じ事を言っていた)にも驚いたが、馬を扱う人々がそうしたことを知識として知っているということにも驚いた。

そうして見てくると今度は、牛の鼻は切れないのか、切れて場合はどうするのかという疑問が出てくる。質問してみると牛の鼻はなかなか切れないという。それでも切れることがある。すると多くの所では頭絡をつける、場所によっては棒締頭絡を付けると言うところもある。中国雲南省のある村では鼻が切れて鼻輪が付けられなくなった牛は殺すという。制御の聞かなくなった牛はどんな牛でもいらないというわけである。

家畜に関しては常に制御具と関連するようになっていけば、制御具を用いない象の場合はなぜ制御具がないのかと考えることもできる。あまりに大きく力が強いので制御具が意味をなさないというとは推測できるが、他の家畜と比較しながら象の制御方法を考えることができる。

演繹的シソーラスでは、上位概念である「家畜」と「制御具」が関連づけられたら、その下位概念でもそれぞれを関連づけてしまうのである。繰り返したが、一般のシソーラスでは統制して下位概念の個々の家畜とマッチする制御具だけが選択され関連語としてシソーラスに採用される。

生産方法 (入手経路)
自家生産
自村生産
購入
市場
他村

「制御具」と「制御具」の下位概念の個々の制御具をセットに（オブジェクト化）して、その統制しない形の「制御具」グループをそのまま「家畜」の下位概念の個々の動物に関連づける。演繹的シソーラスはデータベースとしてはオブジェクト指向プログラミングの”オブジェクト化”をイメージしている。オブジェクトの中からプロパティを取り出して利用するというイメージである。

左の図のように家畜の生産（入手経路）についてのオブジェクトがあり、それを馬にも関連づけ、牛にも関連づけという形で利用していく。

終わりに

このようなシソーラスを着想したのは、以前、総合研究大学院大学時代に、狩猟・漁労の方法とその道具について簡単な分類表を作ったときに、最初、網を使った漁・猟の中に鳥を捕まえる霞網と捕虫網形式の網が抜けていたことがきっかけであった。

以来、ある技術、道具を分析する際には、それが他の場合だったらどんなものがあるかを考えるようにし、調査や研究の場面での抜けが少なくなるように心がけている。

今回、モノと情報班で博物館資料のデータベース化の話が出たときに、まずこの仕組みをなんとか行かせないだろうかと考えた。具体的に実現させていくノウハウを私自身はもっていないので、まずは図1のようなチャートとして表現してみた。

データベース全体のシソーラスをこのような演繹的なシソーラスにしていくのは大変な作業なので、まずモデルケースとして「馬」を起点に演繹的シソーラスを構築してみたいと考えている。

Summary: The purpose of this essay is to suggest a new way of constructing thesaurus for museum database. The thesaurus suggested here is analogous to object-oriented programming. Semantically affiliated index and terms are mapped in a chart. This would provide a deductive information retrieval system and enable to find an unexpected relation(s) among the words in the chart. That should be heuristic for future research agendas.

