

国立国語研究所学術情報リポジトリ

分類語彙表の特徴と位置付け

メタデータ	言語: Japanese 出版者: 公開日: 2020-06-29 キーワード (Ja): キーワード (En): 作成者: 柏野, 和佳子 メールアドレス: 所属:
URL	https://doi.org/10.15084/00002959

分類語彙表の特徴と位置付け

柏野和佳子(国立国語研究所 研究開発部門第一領域)

1. はじめに

一般の国語辞典では語句がアイウエオ順に掲載されている。そのため、種類の違う国語辞典を見比べた場合、見出し語の収録に多少の差異はあるものの、語句の並びは同じように感じられる。一方、シソーラスや類語辞典と呼ばれるものの多くは語句が意味によって分類・配列されている。そのため、見出し語の収録の差以上に、分類・配列の違いによって、語句の並びは非常に大きく異なった印象を受けるものである。『分類語彙表』をはじめとし、近年、様々なシソーラスや類語辞典が続々刊行されている。そこで、『分類語彙表』の分類体系についてくわしく述べてその特徴を明らかにし、他のシソーラスや類語辞典と比較することでその位置付けについて考察する。

2. 『分類語彙表』の特徴

2.1 『分類語彙表』とは

『分類語彙表』は1964年に国立国語研究所資料集第6として公刊された、現代日本語を対象とした最初のシソーラスである。1999年までに31版を重ねた一方、1981年より増補改訂作業がはじまった。1996年3月のモニター用公開(中野1996a; 1996b)を経て、2004年1月には現在の『分類語彙表』の増補改訂版が刊行され、同時に電子化データベースも公開¹された(山崎2004)。『分類語彙表』とは何か、元版の「まえがき」には次のように述べられている。

ここに分類語彙表というのは、一般に一つの言語体系の中で、その語彙を構成する一つ一つの単語が、それぞれどのような意味で用いられるかを一覧できるように、単語が表わし得る意味の世界を分類して、その分類の各項にそれぞれの単語を配当したものである。

そして、主に次の二つの役割を担うものであると述べられている。

- ① 言葉や概念を手がかりに、適切な言葉を見つけるもの
- ② 語彙の分布や偏りを見るための「物差し」となるもの

その結果、表現辞典として、また、言語の研究資料として数多く利用されてきている(宮島・小沼1994; 中野1995)。例えば、宮島・小沼(1994)には『分類語彙表』を言語研究に利用した論文136例(1965年～1994年)が取り上げられている(表1)。電子化されたFD版(国立国語研究所編1994)が市販された後は、工学分野における言語処理研究における利用も一気に広がった。近年は医学や建築学の利用もあり、『分類語彙表』の研究利用のすそのはさらに広がっている。

2.2 収録語

元版の延べ語数はFD版によれば36,780語である。この数がそう多くはなかったのは、語は意味体系の例示のためにあげる、という方針があったためである。特に、表2に示すような特徴があった。

¹ <http://www.kokken.go.jp/katsudo/kanko/data/index.html> を参照。

表1 『分類語彙表』を用いた言語研究の論文の内訳(宮島・小沼(1994)による)
(件)

分類	論文数
1.語彙体系	21
2.作品の用語調査	41
3.文法(文中の単語結合や、合成語における要素結合の意味構造など)	22
4.方言	3
5.日本語史	5
6.教育・発達	11
7.言語情報処理(語と語の類似度の計算など)	7
8.類語群	14
9.意味分類	12
合計	136

表2 元版の収録語数の特徴

- ①サ変動詞のうち、1字漢語に「する」がついた「愛する」「信ずる」の類は収録していたが、2字以上の漢語に「する」がついた「愛好する」「信用する」などは収録せず、「する」なしの漢語のみ収録していた。
- ②基本語の多くは多義的であるが、代表的な語義に限定して収録していた。
- ③類推が自明なものは収録していなかった。例えば、〈真北、真西〉はあったが、〈真南、真東〉はなかった。
- ④短い単位で収録することが多く、複合語や慣用句はあまり収録していなかった。

しかしながら、表現辞典としても言語の研究資料としても、上記4点を補う増補改訂作業の必要性が高まり、これに新語なども加え、現代の日常社会で普通に用いられる語を中心に数多く増補された。その結果、現在の増補改訂版の収録語数は、延べで95,811語、異なりで79,516語である。この数は、一般の小型国語辞典に相当する数である。なお、元版も増補改訂版も収録語の選定に際しては、国立国語研究所がそれまでに、雑誌、新聞、教科書、テレビを対象に行った各語彙調査結果(国立国語研究所1953; 1957; 1962; 1970; 1983; 1986; 1987; 1995)を参照している。各語彙調査において使用率の高かった一般的な語は、ほぼ網羅的に収録されている。

2.3 分類体系

2.3.1 分類の特徴

以下、特に断らない限り、増補版に沿って説明する。『分類語彙表』の分類体系は固定された階層構造になっており、各語の体系的な位置付けは分類番号によって示されている。木の形は明示されていないが、その体系は木構造であるといえる。図1に例を示して説明する。分類の各項目、例えば〈話・談話〉には〈1.3131〉のように、〈類〉を整数位に置いて小数点以下4けたの「分類番号」が付されている。この数字が全体の中に占める個々の分類項目の位置付けを示している。分類番号によって表される意味的範疇は、より広い概念から順に、〈類〉〈部門〉〈中項目〉〈分類項目〉となっている。

2.3.2 <類>:1階層目:品詞論的な4分類

大分類として、はじめに、品詞論的に4分類されている。この4分類の番号が分類項目の整数位の番号になっている。

1. 体の類…名詞の仲間—《何、何ごと、何もの、どれ、だれ、いつ、どこ、いくつ》等の概念を表す語と、それらを問いとしたときの答えとなるべき語。
2. 用の類…動詞の仲間—《ある》に関するもののほか、《どうする、どうなる》等の答えとなるべき語。
3. 相の類…形容詞の仲間—《ない》に関するもののほか、《どう、どうだ、どんな、どんなに》等の答えとなるべき語。いわゆる形容動詞、連体詞、ある種の副詞を含む。
4. その他の類…その他の仲間—いわゆる接続詞、感動詞、ある種の副詞の類等。概念間の関係付け、叙述間の関係付け、感動、呼びかけ応答、判断・期待・仮定などの叙述態度の予告、待遇表現などを表す語。

しかしながら、品詞分類が優先されると、次のような語は意味的に近くとも離れてしまうことになる。

(1)名詞と複合サ変動詞語幹

「変化」<体の類> 「変化する」<用の類>

(2)名詞と形容詞・形容動詞語幹

「美」<体の類> 「美しい」<相の類>

(3)動詞の派生名詞と動詞

「動き」<体の類> 「動く」<用の類>

(4)形容詞・形容動詞の派生名詞と形容詞・形容動詞²

「高さ」<体の類> 「高い」<相の類>

このような分離に対処するため、大分類以下を細分する際はなるべく平行に行い、細分の番号をある程度一致させることが試みられている。例えば、「光」に関する語は次のとおり分類されている。

- 例: <1.5010 光> 名詞「光、輝き」など
<2.5010 光> 動詞「光る、輝く」など
<3.5010 光> 形容詞や副詞「明るい、くっきり、きらきら」など

つまり品詞をまたいで関連する意味ごとの語を見比べたい場合は、分類番号がその手がかりになるよう便がとられている。

2.3.3 <部門>:2階層目:意味範囲の5部門

大きな意味的まとまりとして、<抽象的關係>、<人間活動の主体>、<人間活動—精神および行為>、<生産物および用具>、<自然物および自然現象>という5つの部門を設けた。ただし、<人間活動の主体>、<生産物および用具>は体の類のみである。分類項目の小数点以下1桁目の数字がこの部門を示す。

² 形容詞・形容動詞の派生名詞は網羅的には収録されていない。例えば「長たらしさ」は、<相の類>で取り上げられた「長たらしい」の一種の活用形とみなされ、<体の類>には収録されていない。

2.3.4 <中項目>と<分類項目>:3・4階層目:意味範囲の5部門の細分類

分類項目の小数点以下, 1桁目と2桁目を合わせた部分が中項目を示し, 3桁目と4桁目を合わせた部分で一つの分類項目を示す。中項目は部門をより具体的に細分したもので, その下位に位置する分類項目をまとめるものである。分類項目は体・用・相それぞれの類の中にほぼ同様に設け, 同様の配列をとった。先に述べたように, 三類相互の参照の必要から, 原則として三類共通の項目名と項目番号が付けられている。そのため, 類によっては番号が飛んでいる場合がある。また, 分類番号は元版の番号ができるだけ引き継がれたため, 必ずしも連続していない。

しかしながら, 3桁目と4桁目が10番台ごとに飛んでいたり, 4桁目が「0」であるものがあつたりなかったりと, 特徴的な飛び方をしている。実はこれは元版の頃より行われている, さらに細分類を示すものである。4階層目がある程度大きく, さらに細分類されると考えられるものについて10番台ごとに細分類をした結果であり, その細分類をまとめるような一群が認められればそれが4桁目「0」の位置にすえられていることが多い。

2.3.5 分類項目以下の細分類

各項目には段落番号が付与されている。そして, 一つの段落は場合によって行単位に改行されている。さらに, 意味のまとまりの区切りを示すために, 複数の段落を区切る「*」が付与されている場合がある。段落番号については, 増補版の「まえがき」に「あくまで検索の便宜のためであって, 分類自体の小分けを意味するものではない」と断り書きがあるが, 分類項目以下に, 「*」, 段落, 行, という最大3つの細分類を示す階層が存在していることになっている。

2.3.6 項目と語の分類と配列

(1)項目の配列

基本的には検索の便宜のために, 互いに関連する項目は接続して配列されている。また, 次の例のように, 一般的総括的内容を持つ項目は, 部分的な内容を持つ項目より先にあげられている。

例: 身体に関する部分の総記 <1.5600 身体>
身体各部位に関する記述 <1.5601 頭・目鼻・顔>, <1.5602 胸・背・腹>, <1.5603 手足・指> … <1.5608 卵>

(2)項目の見出し

分類項目や中項目の見出しは, それらの内容がなるべくよく表せるように, 項目の全体を示すような代表単語のほか, 代表的な単語の列挙, 説明的な語句などによってつけられている。

例: 代表単語 <1.2510 家>
代表的な単語の列挙 <1.2650 店・旅館・病院・劇場など>
説明的な語句 <1.2750 国際機構>

(3)項目の大きさ

各項目に収載する語の数は限定されていない。よって, 項目によって収載されている語の数は異

なり、項目の大きさには大小さまざまある。

例： 小さいもの

<2.52 天地>には<2.5220 天象>があるのみでこの項目の総語数は13語。

<2.14 力>には<2.1440 力>があるのみ。かつこの項目の段落数2は最小。

大きいもの

<1.2340 人物>の段落数77は最大。

(4)項目間の概念の重なり

山崎(2004)は階層の重なりや概念の重なりがあることを指摘し、見直しの必要があることを述べている。前者の例としては、「移動」は<1.1510 動き>と<1.1521 移動・発着>の2つの項目に分類されているが、<1.1510 動き>は<1.1521 移動・発着>の上位概念に当たる。後者の例としては、<2.3393 口・鼻・目の動作>と<2.5710 生理>とでは22語が共通している。また、<1.1330 性質>と<1.3420 人柄>とでは17語が共通している。

(5)語の多重分類

第一に多義語はその意味ごとに分類されている。第二に語のもつ意味の異なる部分に着目し、それに応じて分類されている場合がある。以下、それぞれ例をあげる。

例： 【多義】 顔 <1.5601 頭・目鼻・顔>，<1.3030 表情・態度>，
<1.3041 自信・誇り・恥・反省>，<1.3142 評判>

【着目の差異】 言い直す <2.1500 作用・変化>（「書き直す，読み直す」と共にあり）
<2.3071 論理・証明・偽り・誤り・訂正など>
（「書き直す」と共にあり，「読み直す」はなし）
<2.3100 言語活動>（「書き直す，読み直す」はなし）

(6)一項目に収めた語の性質と配列

一つの項目は、同義や類義の関係でまとめられている。中には、対義の関係までまとめて扱った方がわかりやすいと判断されたものは同一の項目や、時には同一の段落に含まれている。なお、元版の「まえがき」で述べられていたとおり、自由連想による語群をとらえることは語彙論上意味のあることであると認めつつも、一つの項目が自由連想による語群になることは極力避けられている。例えば、「ビール」についての連想による語群として、「酒，ウイスキー，飲む，酔う，一杯，あわ，ジョッキ，コップ，ほろにが，ホップ，赤ら顔，ビヤホール」などがあげられる。しかし、『分類語彙表』では、「ビール」は、<1.4350 飲料・たばこ>の下に、嗜好品の一つとして扱われ、上記のうちの、「飲む」や「ビヤホール」などとの関係は断たれている。

配列については増補版の「まえがき」に「段落および段落内の語の順序は、なるべく意味・用法の広いほうから狭いほうへ配列しているが、必ずしも厳密ではない。」と説明されているとおりである。

2.3.7 体系のまとめ

以上、述べたとおり、はっきりと分類をうたっている階層は<類><部門><中項目><分類項目>の4階層である。しかしながら、形式上階層が認められるものまであげれば、4階層を表す分類項目下2桁目を10番台ごとに区切る5階層、「*」の6階層、段落の7階層、行の8階層がある。一番深い場合に語はその8階層目に分類される。一番浅い場合では、段落が5階層目にあたり、そこに語が分類される。つまり、語は5～8階層目に分類されていると言える。体系のまとめを表3に示す。そして、分類項目数と語数の内訳を表4と表5に示す。

表3 体系のまとめ

収録語数	延べ 95,811 語
分類の特徴	品詞(4),部門(5),中項目(93),分類項目(893)の4～5階層。 さらに、「*」(任意),段落,行の区別があり,それを入れると5～8階層になる。

表4 <類>と<部門>別の分類項目数の内訳

	体	用	相	その他
抽象的關係	141	77	59	8
人間活動の主体	55			
人間活動—精神・行為	173	148	37	14
生産物・用具	78			
自然物・自然現象	62	24	16	1
計	509	249	112	23

表5 <類>別の収録延べ語数の内訳

	体	用	相	その他	合計
抽象的關係	64,457	21,605	8,879	870	95,811

3. 『分類語彙表』の位置付け

3.1 主なシソーラス・類語辞典とその特徴

『分類語彙表』以外の日本語の主なシソーラス・類語辞典とその特徴について以下に、簡単に記述する。

(1)『角川類語新辞典』(1981)

収録語数:約 60,000 語

分類の特徴:大分類(10),中分類(100),小分類(1,000),細分類(約 3,000)の4階層。品詞の区別なし。大分類は、「自然」「人事」「文化」という枠の中で設けられたもの。

その他特徴:語釈,用例あり。位相が付与されている。

(2)『日本語語彙体系』(1997)

収録語数:約 300,000 語

分類の特徴:1階層目が品詞。2階層～12階層目に2,876個の意味属性を木構造で配置。

その他特徴:さらに6000語の用言には日英の対訳文型14,000パターンを付す。機械翻訳を主目的にして開発されたため、固有名詞も多く収録されている。

(3)『類語大辞典』(2002)

収録語数:約 79,000 語

分類の特徴:カテゴリー(100), 小分類(916), 品詞(11), 小見出しの4階層。カテゴリーは用言に基づいて作成し, 人に近いところから遠いところへという原則にしたがって配置。

その他特徴:語釈, 用例あり。位相が付与されている。

(4)『日本語大シソーラス』(2003)

収録語数:延べ約 320,000 語, 異なり約 200,000 語

分類の特徴:カテゴリー(1,044), 小語群(14,000), セミコロン(任意)の3段階。品詞の区別なし。

カテゴリー部分が先に2~4階層に分類されているため, あわせると一番深い場合で6階層。品詞の区別なし。語釈なし。文法的カテゴリーなどの一部に用例あり。

その他特徴:故事成句, 人名やオノマトペが多く収録されている。

(5)『類語例解辞典』新装版(2003)

収録語数:約 25,000 語

分類の特徴:大分類(10), 中分類(200), グループ(約 6,000)の3段階。助詞・助動詞は別枠で78に下位分類。品詞の区別なし。

その他特徴:語釈, 用例のほか, 使い分けの解説, 対比表などがある。

(6)『類語新辞典』(2005)

収録語数:約 50,000 語

分類の特徴:柱(3), ジャンル(18), 分野, 領域の3段階。

品詞の区別なし

その他特徴:語釈, 用例のほか, 位相, 使い分けの解説, 図解などがある。また, 類語のニュアンスを解説するコラムが 86 ある。オノマトペも積極的に収録されている。

木村(1993)によれば, このようなシソーラスや類語辞典の編集目的には次の3つがあると言う。

- ① 適切な言葉探しか ②語彙の分布や偏りを見ることか ③使い分けを知るためか

そして, 体裁面から区別するための指標には次の3点が考えられる。

- A) 語釈や用例などがあるか。
B) 分類体系は品詞別であるか。
C) どのような語をどれくらいの数, 収録対象にしているか。

このうち, 今回とりあげたシソーラス, 類語辞典は, A)に該当するか否かで, 『分類語彙表』『日本語語彙体系』『日本語大シソーラス』と, それ以外のものと大きく区別される。この A)は目的③をもつかもたないか, ということに強く関わってくる。A)に該当しない『分類語彙表』『日本語語彙体系』『日本語大シソーラス』の主な目的は①や②にある。そして, それ以外の A)に該当するものは目的

③も強くもつと考えられる。「シソーラス」の簡単な定義は「語句を意味によって分類・配列したもの」であるが、この定義に該当するものうち、目的①や②の強いものが「シソーラス」と呼ばれ、目的③の強いものは「シソーラス」ではなく「類語辞典」と呼ばれて区別されていると言えるだろう³。さらに今回リストにあげた「類語辞典」と呼ばれるものの中では、目的③を強く打ち出し、収録語を限定しているという点において、『類語例解辞典』が他三つの類語辞典『角川類語新辞典』、『類語大辞典』、『類語新辞典』と大きく区別される。

3.2 比較

3.2.1 分類体系は品詞別であるか

『分類語彙表』と、同じシソーラスである『日本語大シソーラス』とを主に比較し、続けて三つの類語辞典『角川類語新辞典』、『類語大辞典』、『類語新辞典』とも比較してみる。同じシソーラスである先の二つは、先にあげた「B)分類体系は品詞別であるか」で大きく区別される。『分類語彙表』は一番はじめに品詞で分類される。『日本語大シソーラス』は品詞の区別はされていない。ちなみに、シソーラスの先駆けと言われるロジェのシソーラス(Roget 1852)は、はじめに意味で分類され、下位で品詞分類が行われている。品詞別の分類体系をとるかとらないかは、シソーラスの設計において一つの大きな分かれ目であろう。そして、この品詞分類が望ましいかは利用目的によるであろう。『分類語彙表』は、同じ品詞どうしでの語の比較のしやすさや、品詞別にどのような意味分布があるかの見通しの良さを優先させ、先に品詞論的に分類されている。

3.2.2 人名はどう分類されるか

『分類語彙表』と『日本語大シソーラス』はどちらも人名が多く取り上げられている。これは、他の類語辞典には見られない面白い特徴である。『分類語彙表』では、人名をシソーラスで扱うとどのような分類があり得るかという試みとして<1.2390 人名>が設けられ、段落 01 の、日本人に多いといわれている「佐藤 鈴木 田中」の分類にはじまり、国内外の歴史上の著名人がいくつか分類され、最後は「浦島太郎 シンデレラ ピノキオ」など、物語上の人物分類で締めくくられている。「浦島太郎」と「シンデレラ」だけを取り上げて比べると、少し距離がありそうにも思えるが、ここでは物語上の人物は一段落でまとめて扱われている。『日本語大シソーラス』では、「序」に「人名は人物典型を表現する際の最初の手掛かりになるものである」と明言してあるとおり、人名の扱いには積極的である。「シンデレラ」は<架空人物>、<美女>、<ななめ中>(継子関係)、<いじめる>など、「シンデレラ」が取り上げられ得るところで多重に分類されている。

3.2.3 属性分類か主題分類か

池田(1993)は、「我々が個物を分類できるのは、我々の目からみて、重要な性質を選んでいるからなのである」と述べている。言葉をどう分類するかは、言葉のどの性質や観点に着目するかによる、と言えるだろう。『分類語彙表』や『日本語大シソーラス』をはじめ、『類語大辞典』以外の類語辞典は、主に「関係」「人間」「自然」を分類の出発点にし、その下に細かな分類を展開している。よって、

³ アイウエオ順になっている類語辞典もある。これらは「シソーラス」の定義からは外れるが、目的が言葉探しにあるものもある。たとえば、『早引き類語連想辞典』(2001)(野元菊雄監修、米谷春彦編集、ぎょうせい)は語釈や用例がなく、言葉探しという目的に絞られているようである。

どれも似たような体系なのかと思うと、部分的に似たところもあるとはいえ、実はどれ一つ同じ体系にはなっていないのが事実である。荻野(1993a)は「概念化のレベルも観点もさまざまな多くの語彙があるとき、どのように「共通概念」を括り出すのか」と述べている。この「共通概念」の括り出し方が多様であり得るため体系に相違が生まれるのであろう。

たとえば、木村(1993)は、『分類語彙表』と『角川類語新辞典』とを比較し、後者には主題分類が含まれると指摘している。具体例として、後者は 073「発病」という小分類の下に c「治療」という細目も含まれるが、前者では、「発病」は<1.5 自然現象>の下に、「治療」は<1.3 人間活動>の下に収められ、大分類からして別になっている、という点があげられている。

木村(1993)に主題分類の定義はないが、主題分類とは、語が本来持っている事柄の性質に沿って分類することより、ある主題のもと関連する語を集めて分類することを優先するものと定義されよう。逆に、語が本来もつ性質によって分類することを優先するものを、ここでは属性分類と言うことにする。

実は、『分類語彙表』と『角川類語新辞典』で言われた同じ指摘が、『分類語彙表』と『日本語大シソーラス』の間にも言えそうである。『分類語彙表』の一番の目的は「語彙の分布や偏りを見る」ことであるために属性分類を行い、その点が、「言葉探し」を一番の目的に主題分類を行う『日本語大シソーラス』や類語辞典との分類の違いを生んでいると思われる。例えば、『日本語大シソーラス』では、「医者」は「医術」などと共に<医学>のところに、「易者」は「占術」などと共に<当たり外れ>に、それぞれ離れて分類されている。『類語大辞典』も同様に、<治す>と<見込む>に分かれている。『類語大辞典』はそもそも用言の分類から主発しているため、もっとも主題分類が徹底していると言えそうである。一方、『分類語彙表』では「医者」も「易者」も同じく<1.2410 専門的・技術的職業>に分類されている。この二つ上の分類<1.2 人間活動の主体>という括りで語を眺めると、「人間」に関し、性別、年齢、親族、対人関係、社会的地位など、さまざまな場面でのどのような語が用いられるのか、一覧することが可能になっている。類語辞典のうち、『角川類語新辞典』や『類語新辞典』も職業に関する語は一つの括りになっており、「医者」と「易者」などはその括りの中で一覧できるようになっている。しかし、親族や代名詞に関する語は別分類になっており、『分類語彙表』ほど「人間」に関するすべての語が一覧できるというような体系にはなっていない。表現のバリエーションの一覧性の高さは『分類語彙表』が属性分類を行っているための特徴であろう。

さて、「言葉探し」の目的には主題分類になっている方が使いやすいそうである。しかしながら、どのような主題分類がより「言葉探し」に適切であるかは、また難しい問題である。学校や教育という主題のもとでは直感的に近い関係にあると思われる「先生」「担任」「学生」を例に、これらがどのように分類されているかを調べた結果を表6に示す。『分類語彙表』では主題分類をしていないことにより、「先生」「担任」と「学生」とでは離れた分類になっている。なお、人物という属性が共通するのでそう大きくは離れていない。『角川類語新辞典』は「担任」が少しずれて分類されていた。『類語新辞典』と『日本語シソーラス』では同じところに分類されていた。もっとも主題分類に徹していると思われる『類語大辞典』は、別の主題が優先されていることによって、三者は全く別々に分類されていた。

このように、結局は、属性分類であろうと、主題分類であろうと、先に述べたとおり、言葉のどの性質や観点に着目するかによって分類、配列は異なるものである。「一覧」や「言葉探し」のためにどのような分類、配列が適切であるかはその時々によって違ってくるものであり、単純に優劣を論じられるものではないであろう。

表6 「先生」「担任」「学生」の分類比較

	先生	担任	学生
『分類語彙表』	<1,2410 専門的・技術的職業>		<1,2419 学徒>
『角川類語新辞典』	<572 教育者>	<552 担当者>	<572 教育者>
『類語新辞典』	<J3 職業>		
『日本語大シソーラス』	<0380 学事>		
『類語大辞典』	<2102 教える>	<5307 になう>	<1806 学ぶ>

3.2.4 科学的分類であるべきか

荻野(1993a)は、科学的分類は日常の感覚とは相当に異なっていることを、魚、虫、野菜の例を用いて説明している。太田(1993)もまた、食べ物を対象に、「クジラはけものかさかなか」「トマトは野菜か果実か」などを例に、科学的な分類と文化的な分類とは違うことを説明している。言葉の分類に厳密な科学的な分類は必要ない、ということはおおよそ共通に認識されていると思われる。しかし、『分類語彙表』と『日本語大シソーラス』、その他類語辞典で、魚、虫、野菜、果物のいくつかを引き比べてみると、科学的分類をどれだけ行うかという方針はそれぞれで異なることが浮かび上がる。例えば、「メロン」「すいか」「いちご」は分類学上は野菜に分けられ、青果市場では果物として扱われているものである。これらを果物とは区別して、野菜や植物として扱っているのは『分類語彙表』、『角川類語新辞典』である。そして、果物として扱っているのは『日本語大シソーラス』、『類語大辞典』、『類語新辞典』である。『分類語彙表』や『角川類語新辞典』は、魚、虫、野菜、果物など、全般的に、ある程度科学的分類に即した分類を行っている。一方、『日本語大シソーラス』は、魚、虫、野菜、果物などはすべてアイウエオ順に並べ、細かな分類はしていない。それよりも網羅的に季語として用いられる際の季節を表示し、言葉の分類であるという態度を徹底させているようである。『類語大辞典』と『類語新辞典』も言葉の分類であることを優先しているように見える。ちなみに、『類語大辞典』で「野菜」は<食べる>に、「果物」は<味わう>に区別して分類されている。『類語新辞典』では「野菜」は<植物>に、「果物」は<産物・製品>の<菓子>の下位に区別して分類されている。それぞれ独特の分類観点を持ち込んでいるようである。

4. 『分類語彙表』の今後の課題

最後に、柏野・他(2003)で一度取り上げた点を含め、今後の課題を述べる。

(1)意味関係の明示

田中・仁科(1987)は、計算機でシソーラスを扱う立場から、『分類語彙表』などの「従来のシソーラスは、階層化されたもの相互が意味的にどのような関係にあるかが不明確で曖昧なことが多い。例えば階層化されたもの相互が上位／下位関係にあるのか、それとも部分／全体関係にあるのかがはっきりしない」と指摘し、それらが自然言語意味処理には不十分であり、階層関係を明確にしたシソーラスの作成が重要であると主張した。

人が使う場合は意味関係を推察することが可能であるため、『分類語彙表』など多くのシソーラス、類語辞典は意味関係が明示化されていない。『分類語彙表』で、例えば部分／全体関係のものを見てみると、「日本」と「東日本」は同一項目内の別段落にあるが、「東日本」と「東北」は同一段落にある。同じ段落に「西日本」や「奥羽」「奥州」「三陸」なども分類されており、どれとどれとが部分／

全体関係にあるのか、確かに分かりにくい。

そこで、意味関係の明示化試みはじめられている。例えば、荻野綱男の現代日本語名詞シソーラス(荻野 1993b)では、上位／下位関係と、部分／全体関係とが区別して明示されている。

また、英語のシソーラスに WordNet がある(Miller, G.A. 1985)。これは、語と語の間、また、語義と語義との間にある様々な意味関係をポイントとしてネットワーク状に張り巡らしたものである。ここでは意味関係を詳細にとらえることが試みられている。上位／下位関係の下位にあたるものは「部分」「属性」「機能」に分けてとらえられ、また、部分／全体関係の部分にあたるものは、「一部」「要素」「材料」に分けてとらえられ、階層化されている。また、「同義」だけでなく、「反義」もあわせて扱っている。

(2) 視点の明示

長尾真編(1996)では、「語の分類をする際には語に対する視点の問題を考慮する必要がある。すなわち言葉には別の視点からみれば別の意味の側面に焦点があたるという性質がある。」という指摘がされ、次の例が示されている。

道具という視点:	たわし, 洗濯機 ほうき, 掃除機
動力という視点:	たわし, ほうき 洗濯機, 掃除機

つまり、シソーラスが固定されていると、別の視点による意味関係がとらえられないということである。そのため、シソーラスは、理想的には視点別に動的に生成できることが望ましいということになる。そのようなシソーラスをめざすものとして、川村・他(1994)の「語を種々の観点から分類した多次元シソーラス」がある。これは、ひとつの語に観点を複数与え、観点ごとに動的にシソーラスを構築する方法を提案するものである。例えば、「鳥」と「飛行機」について、次のような記述が提案されている。

動物＝生物{動的属性:(動く), 静的属性:(性別), …}
鳥＝動物{動的属性:(動く(空中)), 利用・用途(食物, ペット), 全体／部分(翼), …}

乗りもの＝人工物{動的属性:(動く), 利用・用途(移動), …}
飛行機＝乗りもの{動的属性:(動く(空中)), 全体／部分(翼), …}

『分類語彙表』など固定化された従来のシソーラスでは、「鳥」と「飛行機」は、それぞれ「生物」と「人工物」として分類されるため、離れてしまう。しかしながら、上記のように複数の観点によって記述してあれば、両者から「空中を動く」「翼を持つ」という共通点をとりだし、その観点でシソーラスを生成すれば、両者を近くに並べてとらえることが可能になる。

このように、川村・他(1994)の提案する、各語について視点や観点を記述するという方法は、意味関係の明示や意味素性の情報明示にもつながり、上記(1)で取り上げた問題も同時に解決しようとするものといえる。

筆者も観点を付与するシソーラスの作成に参加した経験をもつ(情報処理振興事業協会 2001)。

視点や観点の付与はたいへん面白く、語の分類・配列に有用であると思われたが、試作できたのはわずかに 5 カテゴリーであった。このような作業を何百、何千とやることは厳しいとの印象がある。視点や観点を付与することの有効性は荻野(1993b)でも述べられているが、やはり試したのは数十項目と、その数は少ない。

(3)位相の明示

多くの類語辞典ではすでに試みられているが、語が列挙されているとき、それぞれの語を把握するために、位相情報は有用であろう。

(4)語釈と用例の付与

語釈や用例がなく、単語が列挙されているだけの方が語の一覧性は格段に良い。すでに類語辞典がいくつも存在し、また、国語辞典とうまく相互参照しさえすれば、『分類語彙表』そのものに語釈や用例の付与は必要はないのかもしれない。しかし、言葉の並びをみていると、意味や用例の助けがほしくなる場合が多いことは確かである。特に多義語の場合、どの意味で分類されているかを特定させた方がわかりやすい場合がある。先に述べたとおり、視点や観点の記述が進めば解決されることではあるが、それに先立って、簡単な語釈の付与により多義を明示することや、用例の付与により意味の特定を容易にすることを検討したいと考えている。

5. おわりに

『分類語彙表』の特徴は、何よりも語彙の分布や偏りを見るための「物差し」である、という点が第一である。そのための分類体系になっており、そのために固定化されており、そのための収録語が選定されているのである。この点において、他のシソーラス、類語辞典とは線引きされるといえる。

筆者は『分類語彙表』の増補改訂作業に1998年から加わった。おおもとの体系はすでにあつたとはいえ、個々の増補対象語を分類、配列する作業には迷いがつきものであつた。いくとおりも分類、配列が考えられることが多々あつた。どのような語を増補すべきかの選択も常に問題であつた。今回、『分類語彙表』の特徴をまとめ直し、また、他のシソーラスや類語辞典と比較をしたが、おおもとの体系の設計が違えば、なおさら、語の分類、配列や収録語の選定は、似ていてかなり異なるという事実を改めて強く実感した。

一般には、シソーラスや類語辞典は「言葉探し」を目的に使うことが多いであろう。分類、配列、収録語が個々に違うため、『分類語彙表』を含め、その時々を探しやすいものは違ってくるように思われる。いろいろ引き比べてみるのが言葉探しをより確実に達成できるコツであるのかもしれない。

参考文献

Miller, G.A. 1985 'Wordnet: A Dictionary Browser' in Information in Data, Proceedings of the First Conference of the UW Centre for the New Oxford Dictionary. Waterloo, Canada: University of Waterloo.

Peter Mark Roget 1852 "Roget's Thesaurus" First edition, Longman.

NTTコミュニケーション科学研究所監修 1998 『日本語語彙体系』岩波書店。

池田清彦 1993 「分類とは何か」日本語学 Vol.12, No.5, pp.4-10.

- 太田泰弘 1993 「食べ物の分類」日本語学 Vol.12, No.5, pp.58-64.
- 大野晋・浜西正人 1981 『角川類語新辞典』角川書店.
- 荻野綱男 1993a 「シソーラスのための語彙の意味分類をめぐって—「焼き魚は魚か」—」日本語学 Vol.12, No.5, pp.18-30.
- 荻野綱男 1993b 『現代日本語名詞シソーラスから見た語彙の意味分類』平成4年度科研費補助金研究成果報告書.
- 柏野和佳子・中野洋・石井正彦 2003 「情報処理研究とターミノロジーから見た『分類語彙表』—分類の体系と専門語の扱い—」情報知識学会誌 vol.9, No.4, pp.12-28.
- 川村和美・片桐恭弘・宮崎正弘 1994 「語を種々の観点から分類した多次元シソーラス」信学技報 NLC-94-48, PP.33-40.
- 木村睦子 1993 「意味分類体辞書の系譜」日本語学 Vol.12, No.5, pp.31-39.
- 国立国語研究所 1953 『婦人雑誌の用語—現代語の語彙調査』秀英出版.
- 国立国語研究所 1957 『総合雑誌の用語 前編—現代語の語彙調査』秀英出版.
- 国立国語研究所 1962 『現代雑誌九十種の用語用字 第1分冊—総記および語彙表』秀英出版.
- 国立国語研究所 1964 国立国語研究所資料集 6『分類語彙表』秀英出版.
- 国立国語研究所 1970 『電子計算機による新聞の語彙調査』秀英出版.
- 国立国語研究所 1983 『高校教科書の語彙調査 I』秀英出版.
- 国立国語研究所 1986 『中学校教科書の語彙調査 I』秀英出版.
- 国立国語研究所 1987 『雑誌用語の変遷』秀英出版.
- 国立国語研究所 1994 『分類語彙表 フロッピー版』秀英出版.
- 国立国語研究所 1995 『テレビ放送の語彙調査 1—方法・標本一覧・分析』秀英出版.
- 柴田武・山田進編 2002 『類語大辞典』講談社.
- 小学館辞典編集部 2003 『類語例解辞典』新装版, 小学館.
- 情報処理振興事業協会 2001 『IPAL(SURFACE/DEEP)の研究—新世代の辞書記述を目指して—』情報処理振興事業協会.
- 田中穂積・仁科喜久子 1987 「上位／下位関係シソーラス IMIMAPI の作成(I)」情処技法 Vol.87, No.84.
- 長尾真編 1996 岩波講座ソフトウェア科学(15)『自然言語処理』岩波書店.
- 中野洋 1995 「分類語彙表の増補とその利利用」言語処理学会第1回年次大会発表論文集, PP.141-144.
- 中野洋 1996a 『「分類語彙表」形式による語彙分類表(増補版)第1分冊<本表>』国立国語研究所.
- 中野洋 1996b 『「分類語彙表」形式による語彙分類表(増補版)第2分冊<索引>』国立国語研究所.
- 中村明・芳賀綏・森田良行 2005 『類語新辞典』三省堂.
- 宮島達夫・小沼悦 1994 「言語研究におけるシソーラスの利用」『語彙論研究』pp.539-568, むぎ書房.
- 山口翼編 2003 『日本語大シソーラス』大修館書店.
- 山崎 2004 「『分類語彙表—増補改訂版—』の分類の特徴について」日本語文学第20輯, pp.73-86, 韓国日本語文学会.