

## 『分類語彙表』元版・新版のコードの比較

田 島 毓 堂

### 0. 題意

日本語シソーラス『分類語彙表』は、第二次世界大戦後に設立された国立国語研究所の各種語彙調査の結果の提示法の一つとして、林大氏を中心に編纂されたものである。それ以前の各種語彙表と共に、国立国語研究所の行った語彙調査の結果、或る程度の使用頻度を持つ語を採用して、意味分野別に配列したものである。19世紀半ばに公表され、以後、現在に至るまで各種の形態を取りながら発行し続けられている P. Roget: Thesaurus of English Words and Phrases (London 1851) の意味分類を参考にしつつ、語の品詞的特性を加味し、「体の類」「用の類」「相の類」「その他」に4大分類し、その中を、1. 抽象的關係、2. 人間活動の主体、3. 人間活動——精神及び行為、4. 生産物および用具、5. 自然物および自然現象の5部門に分け、(但し、2. 用の類、3. 相の類には、「人間活動の主体」、「生産物および用具」はない。また、4. その他は若干基準が異なる) それぞれに小数点以下4桁の意味コードを与えて、語を分類配列する。正確には数えられないが、『分類語彙表』初版の前書きに寄れば、およそ32600語が798項目に分けられている。意味の記述等は一切無い。

本書は、その前書きに依れば、1. 表現辞典・詞藻辞典、2. 類似語形の史的比較研究の資料、3. 意味の一覧表として異なった言語語彙対照の物差しとしての役割、4. 基本語彙設定のための基礎データとしての役割を持つものとして、1964年に編纂、公刊された。

私どもはこの三つ目の役割を果たすものとして、本書を活用しているが、本書は、種々の使われ方をしてその後公刊し続けられ、2004年には、その増補改訂版が出版された。

増補新版は、収録語彙が大幅に増加した。前書きには、重複する語を含めて95811語とある。種々の改訂がなされ、大いに利便性を増した。それは言うまでもないが、元版とこの新版を比べると、量的な違いは別として、見逃せない違いがある。意味分野の設定の違いと、所属語彙の違いがある。従って、元版を元にしたある語彙の意味分野別構造と、新版の意味コードを与えた場合の意味分野別構造には差が出てくる。どちらがいいとか悪いとか言う問題ではないが、とにかく違ってくる。両者を単純には比較できなくなる。

本稿は、その両者の意味コードを一々比較して、一体どういう差があるのかを示し、同時に、調査の過程で判明した若干の事柄を示そうとするものである。

## 1. 『分類語彙表』元版と新版の比較

1964年の元版に対して2004年の増補改訂版はまさに大増補版である。問題は、その意味コードが相当変わっていることである。その比較について以前、ほんの一部、或いは半分近くを対象にした結果を示したことがあるが、今回全体の比較作業を終え、その結果を報告出来ることになった。

両者を比較すると、まず、増補版では語数が大幅に増補されている。それは、利便性向上という点で大いに歓迎すべきことである。それと共に、同一語にかなり多くのコードが付いている（異なりで79516、出現数95811。付け加えて言えば、この出現数を前書きには「延べ」と書いてある。そのため随分この「延べ」という用語が紛らわしくなっている）、つまり、文脈によって一つのコードでは表せないものに別のコードを付けて区別したということである。16000以上の重複がある。一語平均ほぼ1.2回出現している（こんな平均値は殆ど無意味だが）。これはプラスに評価できるが、難しい問題もある。必ずしもいいことばかりではない。一体、どのコードが当該の文脈で適当なのか、判断が非常に難しいことが沢山ある。

掲載語の単位として、複合語が採用されたことはかなり大きな違いである。元版に複合語がなかったわけではないが、新版にはかなり大量に、新語と共に増補されている。中で、複合サ変動詞の増補が顕著である。元版では、一字漢字のサ変動詞は掲載されていたが、語幹が漢字2字以上の長いものはなかったのに対し、増補版には大量の漢語サ変動詞が掲載されている。大雑把に言って、サ変動詞関係の語は10000語を超える。

その漢語サ変動詞のコード化は、原則としてその語幹（名詞が多い）のコードの整数部分を2に変えたものである。私共はこの増補版の出来る遥か以前から複合サ変動詞にも形容動詞と共にコードを付けることを実践していた。その「コード付けの基準」を最終的にはバージョン5まで示した（「コード付けの基準—単語コードと語素コード・比較語彙論のために（その5）—」『名古屋大学文学部研究論集』2001年）。その中にあるとおり『分類語彙表』増補改訂版が取った方法と同じことを実施していた。どちらが真似したということではなく、偶々一致したということであるが、このコード付けについては、合理的なこととして自信が持てる。

その他、増補版には慣用句もかなり採用されている。

## 2. 両者のコードの比較

従来、元版によって、私共の作成した種々の語彙表に、単語コードと語素コードを付けていた。平成15年（2002）度から18年（2006）度にかけて実施した科研のプロジェクト「語彙論の基盤となる語素コード・単語コードの研究及びその電子辞書開発」において使用したのは、増補改訂版2004の公開前だったので、当然元版に依っている。成果の一つとして関係者に配

布した電子辞書案も元版のコードである。

元版は、現在、もう入手が困難なので、これからはどうしても増補改訂版に依らざるを得ない。一体、両者のコードはどう違うのか、同じなのか。その比較をしてみよう。これは、『分類語彙表』新版の改訂の為に必要になるであろう作業である。

比較の方法としては、『分類語彙表』元版のフロッピー版の索引(1992)をベースにして、それと、新版の索引を逐一比較した(以下の記述で「ファイル」と称するのは、このフロッピー版を言う。「索引」は書冊形式の『分類語彙表』の索引である)。電子的に、その比較をこころみたが、一対一で完全に一致する語の他は、全然信頼できなかつた。結局、一々見比べなければ確実なことは分からないことが分かつた。これは、ひとえに、日本語の表記法による。つまり、現在の日本語の表記法は、漢字だけ、または、仮名だけである語をある語として固定することが、とても困難なのである。尤も、この表記法が別に大きな表現上の特徴を持っていることも事実だから、一概に、日本語の表記法を良くないと退けることは出来ないことは言うまでもない。その恩恵は大きく、ただ、コンピュータにとっては不都合だと言うだけである。

元版の索引も同時に対照資料として用いたが、今まで使ってきて思いもしなかつたような種々の問題が出てきた。索引に落ちているものが、相当ある。最初偶然かと思っていたが、次々索引にないものが出てきたのには、少々驚きを感じた。当時のことをご存じの方にこのことを言ったら、編者は大して気にもしていなかつたということだった。これは、新版の改良には関係ない。新版の索引は、コンピュータによるものだろうから、こういう心配はないと思うが、本当のことは分からない(新版は全部索引によって対照した、もし、索引に問題があれば、そのまま引き継ぐことになる)。これは今回は確かめていない。

元版『分類語彙表』のフロッピー版は平成5年12月に出ている。その時、元版『分類語彙表』は29版を重ねていたと書かれている。前述のように、対照のために用いている昭和40年4月10日の7版の索引とその索引ファイルとは、かなり違いがある。決して、版による違いではない。最終的にほぼ3.78%の落ちがあつた。尤も、この数字はなかなか正確を期することが出来ない。つまり、実際に比較してみれば分かるが、果たしてそれが何に当たるのか、索引にないとは言っても、別の読み方で出ていたりする。それをどうするのか。抽象的に言っても分かりにくいので、実例を一つ掲げる。

- ・「私立」、ファイルには、「わたくしりつ」に相当する位置に掲出されている。しかし、元版の索引にも、新版の索引にもその「わたくしりつ」の位置にはない。しかし、「しりつ」の位置にはある。こういうのは、後述する「元版索引に無し」という扱いはしていない。
- ・「若盛り わかざかり」、3.166は元版索引にはない。新版にはある。但し、コードは1.1652と体の類になっている。ところがファイルを少し進んでいくと「若盛り」というのがあるが、「若者・我が物」の次「我が家」のまえにある。「わかもり」とでも読んだのか。こんなのはファイル入力者のミスだろう。元版索引に無しとすべきか迷う。とにかく索引とし

ては意味を成さない。

次のようなのも、どうしていいか、難しいところである。元版索引に「碌3.120 3.131」とある。ところが新版にはこういう項目はない。元版にはもう少し後に「ろくに3.120」がある(「ろくな」はない)。新版は「ろくな3.1331」「ろくに3.1391」という項があり、更に「ろくでなし」「ろくでもない」という項もある。元本文は3.120には「ろくに」とあり、3.131には「ろくな」とある。こういう小さい字の部分はどうか。これをどう対照させるか。中々厄介である。索引ファイルは「ろくな」「ろくに」を立項する。元版索引で見えていく限り「ろくな」はない。逆に索引の「碌」は一体どうすべきか。「碌」という本文はない。本文は「ろくに」とある、この小字を無視するということなのか。これは、元版索引にはあり、本文に無いとすべきものなのかも知れない。なかなか簡単に説明できない。

この種のことを言いだすと、これだけでどれだけでも紙数を手取る。索引はあくまで索引、本文によって比較すべきなのだろう。

このフロッピー版の索引ファイルには色々問題があるが、見ていく限り、本文を直接入力しているようで、「索引」部分を入力したものではないようである。だから、これを全面的に信用して進めたいのであるが、それも駄目である。と言うのは、ファイルに脱落している項目が、正確ではないが少なくとも160項目以上有る(中に一つのコードに属する大半の語が脱落している場合もある)。0.44%に当たる。逆に重複している項目も相当ある。最初からきちんと数えなかったので、大雑把だが、60項目以上ある。もっとある筈である。

だから、何を頼りにしたらいいか分からなくなる。自分で、最初から、入力していけばいいのだが、今更それはしていない。結局、随分時間を費やしたが、所詮、このいろいろな数値は、完全どころか、色々問題を含むものである。仮に、自分で最初からやればと言ったが、そうしたところで、ある項目を一つにすべきか、二つに分割すべきか、迷うこともかなりあり、最初に言った、表記法上の問題はかなり大きな問題を含む。何処まで行っても、確実だと言うところへは到達できそうにない。私がしっかりしていないからだと思われる方があったら、そうかも知れないが、是非、ご自分で試して頂きたい。

一寸、話が逸れてしまった。

この索引にあってファイルにないもの、逆にファイルにあるが索引にはないもの、コードに違いがあるもの等、区々である。また、『分類語彙表』の本文には有っても、索引にないものがかなり沢山ある。このことは、私は、今まで知らずにいた。ファイルにおける読み方にも色々問題がある。ただ、これは、このフロッピー版が最初から、意図したところもある。つまり、同じ言葉を色々な読み方をしたということで、32600語という『分類語彙表』の語数に対して、ファイルは36800以上有ることでも分かる。この項目数は、重複があったり、脱落があったりして、正確な数がはっきりしない。一応、今回の調査での最終の数は36917項になった。昨年9月の中間報告(「『分類語彙表』元版と新版のコードの比較(中間報告)」『語彙研究

12』2015)では、ほぼ36850項目だった(このように、調査を進めるにつれて、不要な項目が出てきたり、必要なものがなくて補ったりしなければならず、項目が増減してしまうのである)。この全体の比較の結果を示す。今までの報告同様に示す。

1. 大分類の差があるもの、整数部分の違い→類差
2. 部門の差があるもの、小数第一桁の違い→部門差
3. 中項目の差があるもの、小数二桁目の違い→中項目差
4. 小数三桁目以下の違い→微差
5. その他

に分ける。

完全に一致するものは当然ながら多数有る。以下、( )内はサ行(48%)までを調査した数値、[ ]はア行12%調査の数値(最終的なものだけでいいかもしれないが、それが一部の場合とどんな関係かを示した)。

①完全に一致するもの

- イ 1対1で一致するもの16695語—45.22% (47.08%) [42.26%]
- ロ 1対1対応以外に対応するコードがあるもの2892語—7.83% (7.96%) [8.43%]
- 計19587語—53.06% (55.04%) [50.72%]

②微差

- イ 1対1対応 8158語—22.10% (21.50%) [22.57%]
- ロ 1対1対応以外に対応があるもの2376語—6.43% (5.80%) [6.85%]
- 計10534語—28.53% (27.30%) [29.45%]

③中項目の差

- イ 1対1対応 3510語—9.51% (9.33%) [9.72%]
- ロ 1対1対応以外に対応があるもの694語—1.88% (1.65%) [2.05%]
- 計4204語—11.39% (10.98%) [11.77%]

④部門の差

- イ 1対1対応 755語—2.04% (1.96%) [1.96%]
- ロ 1対1対応以外に対応があるもの131語—0.35% (0.26%) [0.36%]
- 計886語—2.40% (2.23%) [2.32%]

⑤類の差

- イ 1対1対応403語—1.10% (0.96%) [1.40%]
- ロ 1対1対応以外に対応があるもの72語—0.20% (0.18%) [0.18%]
- 計475語—1.28% (1.14%) [1.58%]

⑥その他 1232語—3.33% (3.32%) [4.15%]

新版に無し 1232語 (内29項元版も無し)

①②③の34324語—92.98% (93.32%) [91.94%] については大きな問題は無い。ただし、新版では、1対1の一致、微差、中項目差があつたりする(イ)ほか、同形語にさらに微差・中項目・部門・類の差があるもの(ごく稀に一致する場合もある)が対応する場合(ロ)がある。新版では同形語にはほかの意味コードの付いているのが、①②③の34324語の中にそれぞれ「ロ」とした5962語あり、①②③全体の17.37% (15.41%) [18.81%] に当たる。①～⑤では35685項目中6165項目に複数の語が対応している(①～⑤全体の17.28%)。その様子を以下に記す。

①ロ2892語の中に、

- 「一致」するもの12項目(12語)、
- 「微差」のあるのが838項目(789語の中に)、
- 「中項目」の差があるのが1070項目(972語の中に)、
- 「部門」の差があるのが1444項目(1248語の中に)、
- 「類」差のあるのが208項目(178語の中に)ある。

以上の内、「一致」「微差」「中項目」の差はともかくとして、部門差、類差のあるものが沢山含まれる(合計が、2892を超えるのは、一語の中に複数項目(1項目から6項目)が含まれるからである。この「一致」「微差」等は、元版のコードに対していう)。この内、部門差のあるのは、同じ語の意味の認定に関わる。主として、基本義と比喩的意味とでも言うべきものに部門差が現れる。又、類差が生じているのは同一語を名詞と捉えるか、副詞と捉えるか、形容動詞の語幹と捉えるかというような品詞の差に関わる語が大半である。そのほか、微差についても、小数第4桁が僅かに違うものから、小数第3桁以下が全く違うものまであり、更に検討の必要がある。また、類差にしても、整数部分だけの違うものから、小数点以下全体すっかり変わるものまであり、一括して類差として処理するには些か抵抗がある。これも、微差の問題と共にさらなる追求の必要がある。具体的な例は、後述する。

②ロ2376語の中に、

- 「微差」ありが含まれるのが649項目(599語中)、
- 「中項目」差ありが989項目(869語中)、
- 「部門」差ありが1012項目(866語中)、
- 「類差」ありが416項目(345語中)ある。

1対1対応の他に2376語の中に3066項目が対応している。前述のように、1語に複数の項目(1項から9項)が対応しているからである。1対1対応の8158項以外では、対応する語があるものは、1項目から、最大は9項目に及ぶ。

③ロ694語の中に、

- 「微差」5項目(5語中)、
- 「中項目」差ありが236項目(213語中)、

「部門」差ありが416項目(350語中)、  
「類」差ありが207項目(172語中)ある。

④ロに131語あり、この内

「部門」差のあるのが111項目(95語)、  
「中項目」差のあるのが5項目(5語)、  
「類」差のあるのが38項目(33語)ある。

この部門差があるものが、イ、ロ合わせて886語ある。これは意味の把握の仕方が変わったことを示しているものが含まれる。更に元版一項目に対して新版で複数項目が対応するもの、つまり、同形異義語が新版で複数掲載されている中に(上で、「ロ」とした項目)も部門差のあるものが①②③を通じて2872項目(2464語中)あった。

⑤類差が生じているのが474語、イが402語、ロ72語、この内、「微差」1以外全て類差を持つもの82項目(72語)である。

類差があるということは、『分類語彙表』の意味コードで整数部分に差があるということ、つまり、品詞的認定が異なる語である。かなりの多さと言うべきである。元版で形容動詞を語幹だけで示す場合や、「に」を付けて副詞的に用いる語を「に」を付けたり付けなかったりすることに依る場合がある。元版で「副詞」と認定している語を名詞扱いするものも幾つかある。新版でも、この語の認定はかなり揺れている。類差があると認められたものについては、前述したが、単に品詞的認定の差があるだけものと、意味的に差が認められるものがある。この点については、更に検討する必要がある。

以上をもう一度集計すると、

「一致」は $19587 + 12 = 19599$  53.09% (55.05%)、  
「微差」があるのは $10534 + 838 + 649 + 5 + 1 = 12026$  32.58% (31.43%)、  
「中項目」差があるのは $4203 + 1070 + 989 + 236 + 5 = 6503$  17.62% (17.17%)、  
「部門」差があるのは $886 + 1444 + 1012 + 416 + 111 = 3869$  10.48% (9.73%)、  
「類」差があるのは $474 + 208 + 416 + 207 + 38 + 82 = 1425$  3.86% (3.05%)

となる。索引ファイルの36917語に対して、43423項目が対応している。

⑥その他

新版にない項目が1232語ある。その内、元版にもないのが29語ある(今元版を基準に比較しているのに、元版にないとはどういうことかと言えば、ファイルにはあるが、元版の索引にない、或いは本文もないというものである、読み方の問題その他いろいろの原因によってこういうことが起こっている。もともと、上に言っている36917語自体、元版『分類語彙表』の掲載語数より4000語以上多くなっている。つまり、一語を色々な読み方で掲載しているからである。前述した日本語表記法上の問題がかかわっている)。

その他注意しておきたいこととして、元版索引にない項目、索引ファイルのみにある項目、逆に索引ファイルにはない項目など色々の場合がある。この調査だけで、元版の索引にはないというのが、完全に正確とは言えないが、1396語(3.78%)ある。これはかなりの数と言わざるを得ない。逆に、ファイルにはなくて索引にあるので加えたのが163語、他に重複している項目もかなりあった。これは、正確には言えないが、60項目以上ある。これは途中で気がついたので、最初から、正確に数えてなかった、それで、こんな曖昧な言い方なのである(これを再度、調べるほどの意義も感じない)。索引ファイルのことを言いだすと、色々問題があるが、本筋ではないので、問題があることだけを指摘しておく。

従って、このファイルだけを頼りにして作業した結果は信頼できない。第一に、『分類語彙表』本文の一語が、一項目になっているとは限らないからである。先に指摘したように、ファイルの項目数は、『分類語彙表』の公称、32600語より、4300項目以上多かった。同じ語を色々な読み方で入力したと、付属文書に記されているとおりでである。こういうことはいずれも『分類語彙表』の改訂には直接結びつかない。とにかく比較の途上気づいたので付言した。索引にないことは脱落だし、ファイルにのみ有るのは、ファイル作成者の読み間違いの可能性も皆無では無かろうが、色々な読み方から引けるためという親切の結果と受け止めておく。これは、前述のとおり『分類語彙表』が、漢字表記のみまたは仮名表記のみで語が掲載されているため、日本語表記の一特徴を示していることにもなる。語をきちんと同定するためには、どうしても、「読み方」「漢字表記(またはそれに類する語形表記)」の両者が必要になるのである。

なお、今回の調査は、新版は索引を信頼して使用している。それに元版索引のような問題があるかどうかはまだ、確かめていない。

### 3. 調査上の問題等

以上、かなり事細かに数量的実態を指摘した。その実態については、中間報告に述べたこととは、出来るだけ重複は避け、問題とすべき事柄について述べる。その前に、大まかに、調査項目について、記しておく。

#### 3.1 ①「一致」

前項で、一致としたものについては、特段の問題はない。特に、1対1で対応する①のイは、全くの一致ということで、問題ない。ただ、これは半数弱である。

一応、同じ意味コードで対応はするが、同時にその語に、他の意味コードも対応している場合=①のロ、これも、一致することに着目して、一致の中に入れてあるが、同時に対応している項目を見ると、様々である。これは、「一致」の場合に限らない。②以下も「ロ」項目は同様である。こういう、別に対応するものがあるものを含めて「一致」としたものが53%、以前の12%調査や、48%調査の場合と、若干の数値の差はあるが、ほぼ同様である。( )や



[ ] 内にそれを示した。

### 3.2 ②「微差」

29%弱が小数2桁まで同じで、3桁目以下に差異があるものである。これには、実は、本当に、最後の1桁だけに僅かに差のあるもの（例えば、「愛飲」は元版1.3332、新版1.3331）から、小数点2桁だけは一緒だけれども、3桁以下は全く違うもの（例えば、「あくどい」元版3.3410、新版3.3422）があるが、これを今回は「微差」として一括している。この違いについては、更に検討の必要がある。

②口については、①口と同様である。

### 3.3 ③「中項目」の差

小数点2桁目以降に差が出るものである。11%強ある。これも、「微差」の場合同様の問題があるが、更に、新版では元版と作成方針が変わり、意味分野の枠組みが変化したと思えるところがある。3.30……と、3.34……、及び、1.5の植物・動物に関する部分は、一括して中項目の差がある。これは実は、「微差」と称したものの内にも、.19……の項目については、元版・新版の間に顕著な差があった。

### 3.4 ④「部門」の差

この部門差についても、中項目の場合同様、改訂方針に若干、差がある。身体の動作にかかわる言葉、例えば「起きあがる」元版は2.1513ですが、新版は2.3391となっている。この類例は沢山ある。また、寒暖に関して、元版が、.5〔自然物及び自然現象〕に収めていたものが、.19に変わっている。例えば「温暖」元版3.5150が新版3.1915になるといったものである。これは、編者の考えによるもので、とやかくは言えないのだが、釈然としないものが残る。2.4%ある。

### 3.5 ⑤「類差」

意味コードで、整数部分に違いがあるものである。こういう類差が生じているものが、1.29%にあたる475語ある。これらは元版と新版で品詞的認識が異なっている語である。同時に意味的認識に差があるものもある。単に、品詞的認定の異なるものもかなり多いが、意味的にもまるで違うものもかなりある。この場合、掲出語のあり方にも問題がある。形容動詞語幹を掲げて、3.……としたり、それを1.……としたりすることがあるからである。一語に、3.……と1.……と両方のコードを掲げるものもかなりある。ここに、意味分類辞典としての『分類語彙表』の編集方針についても考えるべき問題がある。つまり、意味コードの内、整数部分は品詞的範疇を示すもの、即ち文法的認定であり、意味的観点からは、むしろこれは後方におくべきものである。これらは、整数部分を取り払えば、わざわざ別にコード立てる必要はないものである。若し、品詞的区別をしたいならば、別の手段、コードで言えば、最初の整数部分でなく、末尾に示せば問題無かろう。

そういうものと、意味的に全然別のコードになるものもある。正に、語の意味の認定が変

わったのである。部門差の場合にも同じようなことがあった。比喩的意味と直接的意味、両者必要な場合も勿論あるが、両者の掲載される場合と、一方だけの場合がある。この、類差のある項目は、『分類語彙表』編成自体を考える切っ掛けになりそうだし、語の表記をもっと突き詰めて考える契機にすべきものようであり、さらなる検討の必要を感じる。

### 3.6 ⑥新版で掲載されない語 その他

新版でなくなっている項目の多さに驚く。1232語(3.34%)ある。この中身を見なければならぬ。色々な場合がある。

#### 3.6.1 元版にも無い 読み方の問題

元版を元に比較作業をしているのに、「元版にない」とは何事?と思われるかも知れない。前述のとおり、フロッピー版『分類語彙表』の索引ファイルを元にしてしているからである。

どんなものか、

- ・「あがた」1.255 これは、本文の1.255の欄を見ると、「県」とある。これをファイル作成者が「あがた」と読んで入力したのだろう。索引としては、色々な読み方があれば、便利なことは確かだから、これを一概に否定することはない。
- ・「あわい間」1.176 これは、1.176の欄を見ると、「間(あいだ・かん)」とある。「あわい」はない。「間」を「あわい」と読めないことはないが、索引だけ見れば、「あわい」という語があるように思ってしまう。それでいいかどうか。
- ・「いい飯」1.431、「いお庵」1.441、「いもと妹」1.214、「おとと弟」1.214、「おみ臣」1.231、「おもて面」1.175・1.457、「おろがむ拝」2.3390・2.3360、「おんじゅ飲酒」2.3332など、確かにそういう読み方はあり得る。ただ、『分類語彙表』本文で「臣」には(しん)とあり、1.175「面」には(めん)と読み方が示してあるのは「間」の場合と同様である。ただ、そういう読み方からもこれらの語に到達できるというだけに留めておくべきか。
- ・「かあさま」1.212は、その欄には「かあ(おかあさま・さん)」とあり、「かあさま」という語形はない。あってもいいけれども、現実的にはこうである。
- ・「加持祈祷」1.356は、1.356の欄にはない。ただ、「加持」と「祈祷」が並んで別々にある。これの一つにしてしまったのである。同様に「善悪」3.341は「善」と「悪」が別々なのを一つにしている。
- ・「あめ天」1.52 この欄に「天(てん)」と「天」がある。読み仮名のない方が、「あめ」に当たるのだろうか。「てん」と読み仮名を付け、「あめ」は付けないということも考えにくい。「天」は何と読むのだろうか。行末にあるのだが、次の行は一時下げで「空」とある。ところが「空(そら)」は別にある。本当は「天空」と続けて一語ではなからうか。こんなことまで思いめぐらす。こう思って、索引を見ると「てんくう 天空」がある。見逃された本文のミスプリントである。

こんな細かなこと、これは、余り意味のないことかもしれない。索引だけ、或いは、ファイ

ルだけを頼りにすれば、間違ってしまうことになるし、本文のミスプリントもある。注意が必要である。元版ナシとしたものは、一応29項目あるが、新版の索引にもない。実際には、同一語を読み方を変えたりしている場合も、厳密には、此処に列挙すべきかも知れないので、実際には、このような例はもっともっとあるということである。

「いまし汝」「うから族」「おちかた遠方」「おっけん越権」「おつねん越年」「おみな女」「か  
いりよく怪力」「たかぞら高空」「たてぬき経緯」なども、読み方の問題と密接にかかわる。

以上は、索引ファイルにあり、そのコードを頼りに本文を見れば、それとおぼしきものがある。読み方によって生じた問題であり、索引としては、或る意味で許されることである。「新版ナシ」としたもののなかにも「読み方」がかかわっているものがある。「あいじゃく愛着」「いんきインキ」「さくき数奇」「しんゆう辛酉」「どうけい憧憬」などは、この読みでは新版索引にない。「あいちゃく」「いんく」「すうき」「かのととり」「しょうけい」として出ている。

### 3.6.2 差別語や侮蔑語

これはかなり多数の語が関係している。差別語、侮蔑語と一括できるようなものが、かなり削除されている。国語辞典とすれば問題のある処置であるが、一時、こういう差別語等に対してかなりナーバスになり、問題になることに対して、増補改訂に際して神経質になったのではないかと思われるフシもある。国語辞典としては、当然包含すべき語彙であるが、改訂にかかわった関係者の一人からは「表現辞典」のつもりで作成したとの証言も得ている。しかし、苟も、日本語シソーラスとしてはこういう語も実際に存在する語として登録すべものとする（なお、『分類語彙表』は「表現辞典」としてのみ発刊されたものでないことは、元版前書きにはっきり記されている）。数え方により、違いはあるが、少なくとも、差別語・侮蔑語の範囲かと思われる語が120語内外、この中には、とても、それにはあたらないと思われる語も含まれている。「身体的」不具を表すような語もかなりある。身体的特徴から、外国人を呼ぶ名称も含まれる。性的な語、男女差別にかかわる言葉、病気に関する語、人を罵る言葉、下品な言葉なども含まれる。

こういう語が排除されたことは、一面理解は出来るが、必ずしも賛成は出来ない。例えば、「農夫」「郵便屋」「百姓」「馬てい」までも排除されることには、表現の自由という観点からは、「表現辞典」たろうとしていることにも抵触するように思われるが如何か。

なお、どういう語かということ、以上の他にも掲げるべきかも知れないが、差し控えた。

### 3.6.3 商品名

関係する語は大して多くはない。「味の素」「コンサイス」など、一般的な言葉として、元版に登録されたのだろう。「高野（こうや）」まで削ってしまう神経には、些か異常なものを感じる。

### 3.6.4 複数掲載語

元版一項目に対して、新版では何項目かが対応することが多いことは、既に言った。6000

項目以上、つまり、約6分の1の語に対して、複数の項目が対応している。この逆もある。元版で複数項目あって、新版がその一部にしか対応しないという場合である。「新版ナシ」の中に、こういうのが、762項目あり、「新版ナシ」としたものの6割以上になる。一々については述べる必要はないと思うが、若干の例を掲示する。

「とらわれる2.3390 [全身的動作] 2.360 [支配・統治]」(新は2.3613 [捕縛・釈放] だけ)、「にちにち 日々3.162 [一度に・再びなど] 1.1613 [毎日・平生]」(新は1.1612 [毎日・毎度] だけ)、「犯人1.234 [人物] 1.2450 [臨時的地位]」(新は1.2450 [その他の仕手] だけ)……、こういうものが762項目ある。その内152項目は、元版索引にもない。ファイルにあることによって判明する。こういうものに何か傾向があるかどうかについては、特に気づいていない。

### 3.6.5 その他

#### ①意味の捉え方

種々の要因により、語の意味をどう捉えるかという問題が起こる。一語が一つの意味しかない、ということは稀である。幾つも考えられる意味のどこに着目して分類するか、色々な因子が関係してくる。その中で、その事柄そのものを重視して分類する場合と、その比喩的意味を重視して分類する場合が目立つ。それは、部門差として現れる。

グループとして纏まって、元版と異なるのが、2.15……として元版にあった特に身体の動きを示す言葉(「受け流す」2.1523、「起きあがる」2.1513、「手渡す」2.1521等)が、新版では2.3……(それぞれ2.3132、2.3931、2.3770)と手足の動作等として示されることが多い。それと、寒暖に関して、元版が3.5……(「暑い・暖かい・寒い・涼しい」等3.515)としていたものが、新版では3.1915(因みに、中項目としての3.19には「量」という名前が付いている)として示されることがかなり顕著な意味の捉え方の違いである。

部門差ではなく、中項目の差として現れるのが、意味の捉え方の違いと言うよりは、新版の編纂方針としての枠組みの変更であるが、動物・植物・身体・病気など.5……に属する項目が、殆ど全て、中項目の差として現れている(例えば、元版1.56は動物、新版動物は1.55)。また、3.3……人間の精神及び行為に関する語の中で、一括して、元版・新版で移動が見られる。

これも、意味の捉え方と言うより、編纂の枠組みにかかわることのようだが、.19……(量)に関係する諸項目で、小数第3桁以降に差が出ている項目が多数ある。

#### ②品詞認定に関して

前述した、類差のある語に関することである。体の語1と相の語3にかかわるものが大半、他に、その他4に関するものもある。475項目ある中で、新版3……が39項目、4……が13項目、1……は423項目ある。圧倒的に、元版で相の語が、新版で体の語に変更されている語が多い。尤も、元版が相の語一項目だったのが、新版では、体の語と相の語の2項目になっているものがかなりある。

これに関しては、語形認定或いは、掲載語形そのものについてももう少し細かく考える必要が

ある。

### ③語の認定自体、語形の問題

前項に関わり、既に述べたことだが、具体的に言えば、ある語に「に」や「な」を付けて副詞或いは形容動詞としていたり、語幹だけを示したり、必ずしも統一的ではない。掲載語形を「形容動詞」と認めるかどうかという文法的論議も出てくるかも知れないが、その立場をきちんとしておく必要もあろう。特に、意味分類体辞書の役割を考えるならば、文法論的に形容動詞を認めるかどうかは別としても、形容語になりうる語とならない語とを区別することは有用なことである。私自身は、その意味で、形容動詞を積極的に認める立場を取っている。

### ④意味の認定、語の対象の消滅、時代性

その時代には通用していた語が、時の推移と共に意味変化することはごく自然なことである。『分類語彙表』の元版と新版は時間として、40年を隔てている。一世代以上の差である。この間の社会状況の変化はかなり大きかった。

時代的に不要になってしまった語がある。約半世紀を近く経て、使われなくなったような語に、野球の新聞用語「安」とか「犠」、国名に関して「西独」「日ソ」「蘇側」「中華民国」、政党として「右社」「左社」、社会関係の「ニコヨン」「為替(交通・運輸)」、廢語に近い「シミーズ」「アルマイト」「セルロイド」なども、新版にない。少し違うが「うっすり」3.501などという語、有りそうに思えても、普通の国語辞典にはない。広辞苑には「うっすらと同じ」とある。話が逸れるが、広辞苑にないような語も元版には少々見られる。「航機」1.467、「中数」1.1912、「胴接ぎ」1.156と1.3841、「日長」1.1624、「まどらか」3.134(これは『分類語彙表』新版にもある)、「としばい年配」(これはこの索引ファイルだけの読み方か)、こんな言葉が広辞苑に見られない。もう一つ「徹見」1.3091(新版もナシ)という言葉がないのは、何かの間違いでないかと、何回も見直した。幾つかの国語辞典を見たがやはりない。『大漢和辞典』にもない。ありそうな言葉だけれども、元版はどこから取ってきたのだろうか。「中数」も『広辞苑』はじめ国語辞典には見あたらない。『大漢和辞典』にはあるが、「中気の数」とか、「真ん中の数」という意味が書いてあり、1.1912には当てはまらない。元版の趣旨がよく理解できない。

### ⑤ファイルに起因する問題

既に言及したが、この索引ファイルは索引の役割としては、一語を色々な読み方から引けるようにしており、親切な処置である。こうした項目が4000項目以上になる。だから、再三言っているが、今回の数量的な調査結果は、そういう重複を含んだものである。ただ、脱落も少なくとも163個はある。これは索引としては拙い。また、上に述べた「年配」を「としばい」と読むことなどは如何か。読み方についての問題は既述の中にもあった。

何はともあれ、こういう索引ファイルの実際を知っていなければ、間違った結論を出しかねない。

## ⑥不明

なぜ、新版に採用されなかったのか、その理由の分からないものが、80項目ほどある。中には差別語と考えて採用しなかったのかも知れない「四つ足」などというのがある。しかし、そうでないかも知れない。「打ち延ばす」がなぜないのか、「かんしょ甘藷」がなぜないのか、「漁夫」は差別語と誤認したのだろうか、「首切り」という語はあるのに、「くびきる」はない。一つ一つ見ていけば、大変興味深いことがあるが、きりがないので、この辺りでやめる。

## 4. 付言

単に、『分類語彙表』コードの新旧比較というだけのことであるが、実に種々の問題がある。ただ、以前、「あ」「い」だけについて調べて、語彙研究会の例会で発表したことと、ア行（ほぼ12%）について分析してシソーラス研究会で話したこと、語彙研究会大会で発表した約半数についての調査結果で述べたことのほかに特別に新しい問題は出て来なかった。数量的割合については、若干数値上の違いはあるが特段のことはない。出てくる問題も、ほぼ同様であった。元版については論ずべきところが沢山あるが、新版も、改善すべき所がある。

半世紀近くの時間が、言葉の意味の違いも発生させていることが感じられる。

語の認定に関しては、両版ともに問題がある。語形の認定、掲載方法に関わることであるから、きちんとすべきことである。特に、形容動詞語幹か、名詞かは微妙である。どちらとして採用するか方針が揺れているように感じられる。この点は、今後増補版を改訂していく上で考える必要がある。

新版で掲載されなかった語、元版で認められた意味で掲載されなかった語は、理由が分かるのと、不明なものがある。語自体が掲載されないのは、身体的差別語・女性差別語や、侮辱的な用語、時代の変化によって使用されなくなった語などある。表現辞典としての役割を強く持たせたからという編纂にかかわった方の証言があるが、これから文章を書く際の表現辞典としてなら敢えてそういう差別的な語は必要ないかも知れないが、或る意味で辞書的存在としての『分類語彙表』に無くていいものかどうかは考えてみる必要がある。

元版索引は不備がかなり目立つ。今更、言っても仕方ないことだが、フロッピー版のファイルによってそのことを知ることができた。一方、そのファイルの読みには、同一語を種々のよみで掲載するという方針があったことは方針として記されているが、問題が無いわけではない。これは、元々、本書が、仮名だけ、或いは、漢字だけで語を掲出していることに依るものであるが、仮名・漢字表記を併用することは書冊形式では困難かも知れないが、電子的には特に問題はない。日本語をきちんと示すにはどうしても、仮名漢字の両方が必要になってくることははっきり分かる。

全体を見て、一致する項目、特に1対1で対応するもの、これが約半数であるが、これは問

題なしと見て差し支えない。つまり、新版は元版と同じ意味分類をしたということである。微差のある項目については特に問題を指摘しなかったが、同じ微差でも、かなりの幅がある。それを分かるようにすべきかどうかも含めて考えていかなければならない。微差ありとしたのは、小数3桁目以下が違うものであるが、3桁目以下が大幅に違うものと、4桁目が僅かに違うものを、同列に扱っていいものかどうか迷う。ただ、これを進めていく意味があるかどうか、重苦しい思いである。ただそれを言えば、実は、この『分類語彙表』の比較自体、精査する必要があるかどうか、本当のところは分からない。しかし、やり始めてみれば、やはり最後まで調べておこうとしたものである。数量的にかなり整えた筈であるが、実際には、索引ファイルを使った都合で、実数よりかなり多くの項目がある。同じ語が重複掲載されているわけである。そういう基礎的データに依っていることを了解頂きたい。このことが、少しでも、『分類語彙表』改善に資することが出来れば幸いである。

もう一言付け加えれば、比較作業の実際は、個々の問題になると、どの語をどの語に対応させればいいのか、その同定には随分神経を使う場合が多い。実例を示すべきかと思うが、余りに細かく成るので、結果だけを示した。

(本稿は、2015年3月8日台湾大学での語彙研究会特別大会の講演原稿に補訂を加えたものである)

2015年8月1日

キーワード：『分類語彙表』、意味コード、語彙分析

**Abstract**

A Comparison of “BUNRUIGOIHYO” id est Thesaurus of Japanese,  
between the first edition and the revised edition

Ikudo TAJIMA

“BUNRUIGOIHYO” id est Thesaurus of Japanese was published at 1964, and the revised edition at 2004. There are differences between the first edition and the revised edition in meaning code given each word. Now we inquired into those which differ in the meaning code. We will contribute the next revision of “BUNRUIGOIHYO”, by this study.

Keywords: Thesaurus of Japanese, Meaning code, Analysis of vocabulary