

# 語彙資源活用シンポジウム

## 『分類語彙表』の特徴と問題点

山崎 誠(国立国語研究所)

# 概要

- 1. 『分類語彙表』とは
- 2. 分類体系
- 3. 収録語(数, 単位, 多義語)
- 4. 問題点
- 5. まとめ

# 1. 『分類語彙表』とは

- 現代日本語を対象としたシソーラス。
- 1964年初版刊行。約32,600語収録。
- 2004年増補改訂版刊行。約95,800語収録。
- 初版から一貫して林大(1913-2004)が編纂を主導した。



『分類語彙表』初版

## 1.1. 『分類語彙表』の歴史

- 「分類語彙表」が世に出たのは、『婦人雑誌の用語』（1953）に、「意味論上の試み」として、「分類語彙表」（約4300語）という章が掲載されたのが最初。
- 1953年の時点で分類の大枠は出来上がっていた。

I.3  
心身。心。精神。魂。根気 意識」  
印象。感じ。感覚。興奮 疲 疲労 痛み。苦痛。  
美味しさ」  
気。気分。気持。心持。気立 心情 機嫌 御機嫌  
心境 情。感情。感動 感激 感傷 落着 無我  
驚き 喜び。楽しみ。嬉しさ 不快 苦しみ。苦  
悩 悩み。哀れ。悲しみ。悲しさ 淋しさ。恐怖  
恐れ 恐しさ 怒り 疍癩。憤り。心配。不安。  
満足 不満。誇り 好み。愛。愛情。恋。恋愛。  
嫉妬。遠慮 反感 恨み 憧れ 友情 好意 感  
謝。迷惑。ユーモア 謙讓 何気」

『婦人雑誌の用語』 p.267

## 1.2. なぜ『分類語彙表』と言う名前か

- 日本語としては「語彙分類表」というほうが自然。
- 『婦人雑誌の用語』には、ほかに五十音順の語彙表、使用率順の語彙表があり、それと並ぶ語彙調査の成果として、(意味)分類(による)語彙表という位置付けにした。
- 『総合雑誌の用語(後編)』(1958:45-77)にも「分類語彙表」があり、こちらは同語彙調査で出現した16,000語を分類している。
- 『分類語彙表』(1964)は、それに先立つ現代雑誌90種の用字用語調査の成果だけでなく、坂本一郎『教育基本語彙』(1958)等から増補し、より汎用性を持たせたものの。

## 1.3. 意味分類体辞書の目的

- 木村(1993:31)による分類体辞書の編集目的の類型
  - ①作文の際、思っていることを適切に表現する語句を捜すためのもの、すなわち言葉探し辞典。
  - ②類似の表現がいくつかあるとき、どれが最も適切であるかを判断するためのもの、すなわち使い分け辞典。
  - ③語彙の質的構造を見るためのもの。すなわち言葉で表される意味の世界を分析し、組織立て、特定の語彙（現代日本語、古代英語、教育基本語彙など）を採り上げた場合に、どの部分にどれだけの厚みがあるかを見渡すための語彙表。

## 1.4. 『分類語彙表』の役割

- 分類語彙表(1964)の「まえがき」より
  - 1. 表現辞典, 詞藻辞典
  - 2. 方言の分布や命名の変遷を知る手がかり
  - 3. ある個人, またはある社会の言語体系もしくは言語作品について, 表現上の特色を見る物差し
  - 4. 基本語彙選定のための基礎データ
    - <前略> 類義語間のつりあいを, 広い見渡しの上で, 狭くも広くも見ることができ, したがって適切な語の選択ができる<中略>ある特定の観点から語を広く拾いあげるには, 五十音順の辞書からでは何としても偏りを免れない<後略>
  - 5. 辞書編集における解説作業に寄与
  - 6. 翻訳や情報処理たとえば文献要約などの機械化の一つの方法に, 基礎的な作業となる

## 1.4. 『分類語彙表』の役割

- 分類語彙表(1964)の「まえがき」より

- 1. 表現辞典, 詞藻辞典

- 2. 方言の分布や命名の変遷を知る手がかり

- 3. ある個人, またはある社会の言語体系もしくは言語作品について, 表現上の特色を見る物差し

- 4. 基本語彙選定のための基礎データ

- <前略>類義語間のつりあいを, 広い見渡しの上で, 狭くも広くも見ることができ, したがって適切な語の選択ができる<中略>ある特定の観点から語を広く拾いあげるには, 五十音順の辞書からでは何としても偏りを免れない<後略>

- 5. 辞書編集における解説作業に寄与

- 6. 翻訳や情報処理たとえば文献要約などの機械化の一つの方法に, 基礎的な作業となる

日本語学

言語処理



## 2. 分類体系

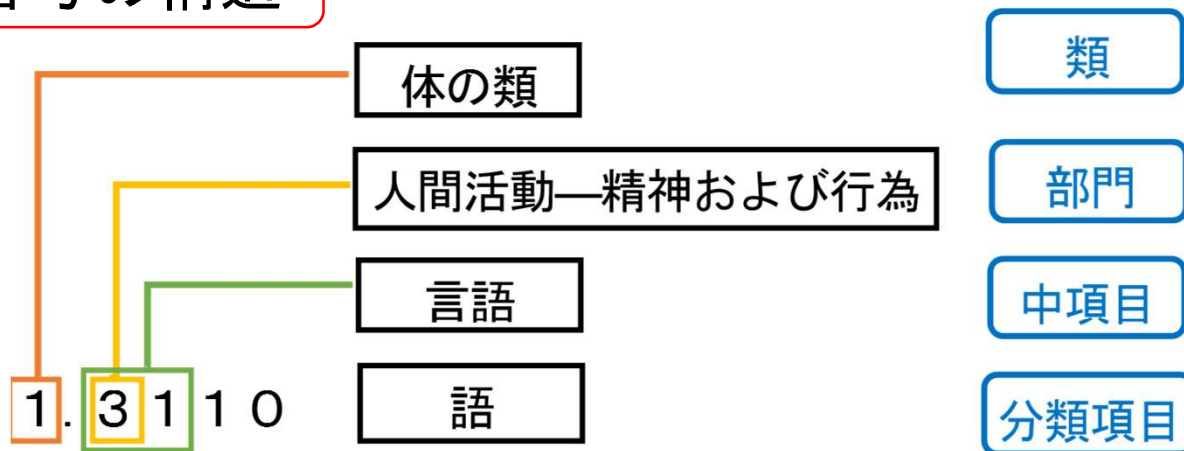
- 分類の階層：4階層

1.3110 語

01 単語 言葉 語(ご) 品詞 詞 辞(じ)  
自立語 付属語

『分類語彙表増補改訂版』 p.128

### 分類番号の構造



## 2.1. 類

- おおよそ品詞のグループに対応する。分類番号の上位1けた。
  - 体：名詞
  - 用：動詞
  - 相：形容詞，形容動詞，副詞，連体詞
  - その他：一部の副詞，接続詞，感動詞
- 
- 体・用・相の区別は，仏教用語に由来する。
    - この概念がどこから出てきたか<中略>これはお経の勉強をさせられている時に出た来たのです。<中略>何でも，「大乘起信論」にあるらしい。（林ほか 2004:41-42）

## 2.2 部門

- 類の直下にある意味の大きなまとまり。分類番号の2けた目。
  - (.1) 抽象的關係
  - (.2) 人間活動の主体
  - (.3) 人間活動—精神および行為
  - (.4) 生産物および用具
  - (.5) 自然物および自然現象
- 体の類にはすべての部門がそろっているが、それ以外の類には「人間活動の主体」と「生産物および用具」を欠く。その他の類は部門との対応は無い。

## 2.3. 中項目・分類項目

- 中項目：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たもの。49個。
- 分類項目：『分語彙表』における最小の意味のまとまり。895個

### 3. 収録語

- 語数はほぼ小型国語辞典程度。
- 接辞・造語成分から慣用句まで収録。
- 多義語はそれぞれの意味に対応する分類項目に配置。

## 3.1. 語数

- 小型国語辞典程度の語彙を収録。

類	延べ	異なり
体	64,457	56,130
用	21,605	16,704
相	8,879	7,357
その他	870	789
全体	95,811	79,516

『分類語彙表増補改訂版』 p.3

## 3.2. 分類項目と語数

- 所属語数が多い分類項目

語数	分類番号	項目名
1,046	1.2340	人物
621	1.3374	スポーツ
605	1.5721	病気・体調
521	1.1962	助数接辞
485	1.2410	専門的・技術的職業
464	1.5402	草本
443	1.4310	料理
426	1.2450	その他の仕手
423	1.3074	学問・学科
407	1.3160	文献・図書
402	1.2390	固有人名
362	1.3230	音楽
360	1.4360	薬剤・薬品

### 3.3. 言語単位

- 言語単位としては、接辞・造語成分から慣用句まで幅広く収録されている。

3.1931 過不足

01 余計 余分 - 余り  
超 - スーパー - ウルトラ -

2.1400 力

01 力を入れる 力を出す 馬力(ぼりき)を掛ける  
力を合わせる 力を付ける  
力を貸す 合力(ごうりょく)する エンジンが掛かる  
拍車を掛ける／が掛かる  
力を加える／が加わる 力に任せる  
圧する 圧迫する 強圧する 勢いに乗る  
調子に乗る  
押す 波に乗る 余勢を駆る



## 3.4. 収録されていない語

- 助詞・助動詞

- ただし、「れる・られる・せる・させる・たい・だけ・しか・くらい」などは収録されている。

- 固有名

- 人名・地名は代表的なもの（歴史や地理の教科書に載るようなもの）のみ収録。

- 専門用語

- 日常語を対象とするため、専門用語は収録していない。

- 生産性の高い派生語，自明な複合語

- 「食べ始める」「法律的」「縞模様」

- 差別用語

- 初版には掲載されていたが，増補改訂版で削除。

## 3.4. 多義語

- それぞれの意味に対応した分類項目に配置。
- 「甘い」は以下の7つの分類項目に現れる。
  - 3.1341 弛緩・粗密・繁簡：(ex.ねじが甘い)
  - 3.1346 難易・安危：(ex.甘い球)
  - 3.3002 感動・興奮：(ex.甘い旋律)
  - 3.3068 詳細・正確・不思議：(ex.甘い判断)
  - 3.3680 待遇・礼など：(ex.甘い対応)
  - 3.5030 音：(ex.甘い声)
  - 3.5050 味：(ex.甘い菓子)

## 3.5. 語義数の分布

- 分類語彙表の語義数

語義の数	語数	割合
1	67,602	83.20
2	11,227	13.81
3	1,839	2.26
4	439	0.54
5	82	0.10
6	36	0.04
7	18	0.02
8	5	0.01
10	1	0.00
11	1	0.00
合計	81,250	99.98

- 約17%弱の見出し語が多義。

## 4. 問題点

- 言語資源としての連携性, 拡張性を中心に問題点を指摘する。

## 4.1. 語のIDがない

- 語(語彙素)を特定する情報が厳密でない。
- 見出し語だけでは不十分。「人気」(にんき)と「人気」(ひとけ)が同じになってしまう。
- 見出し語と読みでも多義か別語かの区別が付かないものがある。
- 「パス」1.1524(通過), 1.3321(合格), 1.3374(スポーツ), 1.4040(定期券) → 多義
- 「バス」1.4470(風呂), 1.4650(乗り物), 1.5030(低音) → 別語
- 現状では表記が違えば別語と見なされる。
- UniDic等, 他の語彙資源との対応を取る際に不便。

## 4.2. 表記のバリエーションが考慮されない

- 原則として1つの代表的な表記しか認めていない。
- 「空揚げ」「玉ねぎ」「バイオリン」
- 「々」を使っていない。
- UniDic等の他の言語資源との連携に不便。

## 4.4. 増補の際の位置

- 刊行から13年たっているため、増補の必要がある。
- たとえば、「スマホ」「ツイッター」をどこに入れるか。分類項目より下の特定が難しい。

1.4620電気器具・部品

07 電信機 電話機 電話 テレホン 無線電話  
有線電話  
加入電話 固定電話 公衆電話 赤電話  
プッシュホン 留守番電話 留守電  
携帯電話 携帯 ← PHS  
ポケットベル ポケベル  
ファクシミリ ファックス テレックス  
受信機 親機 子機

スマホ

1.3122通信

15 パソコン通信  
電子メール Eメール[e-mail] メール  
ISDN ADSL ブロードバンド  
ネットワーク インターネット イントラ  
ネット LAN[ラン]

16 SNS  
ツイッター LINE  
フェイスブック

## 4.5. 過去の版との継続性

- 初版と増補改訂版との間には対応表がない。
- 今後, さらに改訂や増補を行う際には, 互換性を考慮する必要がある。



## 4.6. 分類体系の問題

- 体の類と相の類との区別をどうやって決めるか。
- 増補改訂版では1,301語が体の類と相の類とに現れる。
- 例えば,「よこしま」は相の類にしかないが, 名・形動とする辞書もあり。
  
- 体の類と用の類との重なりが多い(サ変動詞)。
- 体の類に「する」を付けて用の類になっているものが約7,900語ある。
  
- 品詞による分類を行っていることで, 語数が(見かけ上)増えている。

## 4.7. 研究で利用する場合の問題点

- 日本語学では、データの意味的な特徴をマクロにとらえる目的で利用される場合が多いが、複数の分類項目に対応する場合、そのうちどれを採るかが分からない。文脈に沿った適切な意味を1つ割り当てればよいが、手間がかかる。

## 4.8. 実用性・利便性からの問題点

- 語釈・例文がない。

3.1551 統一・組み合わせ

02 集中的 翕然(きゅうぜん)

- 「翕然」の意味が分からなければ使えない。
- 文体, 語の新旧等の位相情報がない。

3.1332 良不良・適不適

01 よい いい  
よろし・よろしい  
結構 グッド グー よさそう 頂ける  
よく よう よろしゅう  
よしなに 体(てい)よく  
体(てい)のよい 体裁のよい 格好のよい  
かっこいい  
不体裁 不格好


- 利用のための適切なツールがない。

## 5. まとめ

- 『分類語彙表』自体の拡張性, 他の語彙資源との連携, また, 実用的な部分でも改良の余地が多い。

## 参考文献

- 柏野和佳子(2006)『分類語彙表』の特徴と位置付け, 「日本語科学」19, 143-160.
- 木村睦子(1993) 意味分類体辞書の系譜, 「日本語学」12(5), 31-39.
- 林大・宮島達夫・鈴木泰・松本泰丈(2004) この人に聞く 日本語研究と私, 『国文学解釈と鑑賞』69(1), 28-45.

- 
- ご清聴ありがとうございました。