

『計量国語学』アーカイブ

ID	KK300503
種別	書評
タイトル	宮島達夫・鈴木泰・石井久雄・安部清哉編(2014) 『日本古典対照分類語彙表』笠間書院 宮島達夫編(2015) 『万葉集巻別対照分類語彙表』笠間書院
Title	MIYAZIMA Tatuō; SUZUKI Tai; ISII Hisao; ABE Seiya. (eds.) (2014). <i>Frequency Dictionary of Old Japanese with Semantic Indexes</i> . Tokyo: Kasama Shoin. MIYAZIMA Tatuō. (ed.) (2015). <i>Frequency Dictionary of Man'yōshū by Volume with Semantic Indexes</i> . Tokyo: Kasama Shoin.
著者	北川 和秀
Author	KITAGAWA Kazuhide
掲載号	30巻5号
発行日	2016年6月20日
開始ページ	286
終了ページ	291
著作権者	計量国語学会

書評

宮島達夫・鈴木泰・石井久雄・安部清哉編 (2014)

『日本古典対照分類語彙表』 笠間書院

宮島達夫編 (2015)

『万葉集巻別対照分類語彙表』 笠間書院

北川 和秀 (群馬県立女子大学)

1. はじめに

宮島達夫・鈴木泰・石井久雄・安部清哉編『日本古典対照分類語彙表』(2014年6月笠間書院刊)は、上代から中世に至る17の古典文学作品において、どの単語がどの作品に何回使われているかを、作品間で対照できる表にまとめたものである。

また、宮島達夫編『万葉集巻別対照分類語彙表』(2015年1月笠間書院刊)は、『万葉集』の各巻において、どの単語がどの巻に何回使われているかを、巻別に対照できる表にまとめたものである。

そして、両書ともそれぞれの語には意味分類を示している。

このように両書はよく似た内容であり、密接な関係にある。

『日本古典対照分類語彙表』は宮島達夫編『古典対照語い表』(1971年笠間書院刊)の改訂増補版にあたる。改訂の1つに、『万葉集』について依拠した索引を差し替えたといふことがある。すなわち、『古典対照語い表』では正宗敦夫『万葉集総索引』(初版は1929年)に依ったのに対し、『日本古典対照分類語彙表』では古典索引刊行会『万葉集索引』(2003年塙書房刊)に依っている。この差し替えのために新たに『万葉集』の語彙データを作成する過程で、巻別に集計した成果が『万葉集巻別対照分類語彙表』ということになる。いわば、『万葉集巻別対照分類語彙表』は、『日本古典対照分類語彙表』の『万葉集』についての巻別内訳表という位置づけになる。

そのような次第で両語彙表の基本的な構造は重なり合う。そこで、次節では両語彙表に共通する点について述べ、第3節では『日本古典対照分類語彙表』について、第4節では『万葉集巻別対照分類語彙表』について述べる。

2. 両語彙表を通して

『古典対照語い表』と『日本古典対照分類語彙表』との大きな相違点は、書名の相違にも現れているように、後者には国立国語研究所編『分類語彙表』(増補改訂版2004年大日本図書)の分類番号が付加されていることである。これによって、今までは困難だった方向からの語彙の分析が可能になった。これは意義深くありがたい増補である。

ありがたいことといえば、語彙表のエクセルデータがCD-ROMに入って附属していることが第1に挙げられよう。『古典対照語い表』も最初は書籍版のみであったが、1989年

にフロッピー版が発売されたことで飛躍的に便利になった。こういったデータは、印刷された五十音順の一覧表のみでは十分に力を発揮することができない。今回増補された分類番号も、番号による並べ替えや抽出ができて初めて威力を発揮する。

たとえば、17 作品の中で地名の比率が多いのはどの作品か、人名の比率が多いのはどの作品か、種々の植物名の作品別の使用度数、などの調査は、今回増補された分類番号によって容易になったが、それも電子データが提供されていればこそこの話である。五十音順に印刷された紙媒体では、折角の分類番号も十分に活用することはできない。

『日本古典対照分類語彙表』の語彙表データは、1 語を 1 行とし、項目（表計算ソフトの「列」）は「見出し（ひらがな）」「順（見出しが同形である語の順位付け）」「漢字」「語種」「品詞」「注記」「作品（用例のある作品数）」「合計（17 作品全体の延べ語数）」（以下「徒然」～「万葉」の 17 作品の各項目）「意味分類」という構造になっている。

一方の『万葉集巻別対照分類語彙表』も 1 語 1 行で、項目は「見出し」「順」「漢字」「語種」「品詞」「注記」「巻数（用例のある巻の数）」「合計（万葉集全体の延べ語数）」（以下「巻 1」～「巻 20」の 20 巻の各項目）「意味分類」である。

『日本古典対照分類語彙表』で 17 作品の作品別データが示されている部分が、『万葉集巻別対照分類語彙表』では各巻別のデータになっているという点のみが両者の相違である。

『日本古典対照分類語彙表』所収の各作品のデータは既存の索引に基づいている。それぞれの索引は編者も異なるので、見出し語の立て方などにも相違がある。17 作品毎に区々では不都合であるから、本書では独自の基準を設けて作品間の調整がなされており、それは凡例の中に詳細に記述されている。たとえば、自立語だけを対象とする、複合動詞・複合形容詞は全体として 1 語とする、名詞についた動詞・形容詞・形容動詞は原則として切り離す（例外あり）など。同語と別語の認定も具体的な例を挙げて基準が示されている。その中で、同音の人名は別人であっても合併している（例示しているのは、「すけただ（資忠～輔尹）」「たかちか（隆親～挙周）」）が、この処置には異論がありそうに思える。枕詞は積極的には立てない方針であるため、「くさまくら」が名詞として扱われ、意味分類が「14270（寝具）」となっている点には違和感がある。

意味分類に際して、多義語には複数の分類が示されている。意味数の最も多いのは下二段動詞の「たつ（立）」（「閉つ」を含む）であるという。内訳は、21211（発生・復活）/21503（終了・中止・停止）/21513（固定・傾き・転倒など）/21521（移動・発着）/21532（入り・入れ）/21553（開閉・封）/23042（欲望・期待・失望）/23142（評判）/23630（人事）/23823（建築）/25030（音）/25155（波・潮）で、意味数は 12 に及ぶ。

多義語を対象にする場合、注意すべき点は、意味ごとの用例数が示されているわけではなく、作品毎の内容も不明だという点である。たとえば、形容詞の「あかし（明・赤）」は 15 作品に登場し、延べ語数は 152 例である。意味分類には「33420（人柄）/35010（光）/35020（色）」という 3 つが挙がっている。そして作品毎の用例数は、徒然 2、平家 5、宇治 28、方丈 1、新古 0、大鏡 11、更級 15、紫 3、源氏 38、枕 33、蜻蛉 8、後撰 2、土左 2、古今 1、伊勢 1、竹取 0、万葉 2、という数字が示されているわけであるが、伊勢や古今にはそれぞれ 1 例ずつしか用例がないわけだから、3 つの意味分類のうちの 1 つの意味でしか用いられていないことになるが、それがどの意味かは分からない。38 例ある源氏においても、3 つの意味分類の全ての用例があるのかどうか、内訳はどうなっている

のかも、実際に 38 例に当たってみないと分からない。

『日本古典対照分類語彙表』所収の「古典語の統計と意味」の 13 ページにも「14400 住居」「14410 家屋・建物」などで大鏡がたかひのは、「みや (228)」・「との (174)」などのせいであるが、じつは、これらは建物をさす例と人をさす例とをふくんでいるので、厳密には文脈ごとの検討を要する。」という記述がある。

どういうことかという、と、「との (殿)」の意味分類は「12110 (夫婦) / 12330 (社会階層) / 14410 (家屋・建物)」という 3 つで、大鏡にはこの語の用例が 174 例ある。「14410 家屋・建物」で検索すれば、大鏡にこの語が 174 例あることが分かるが、174 例全てが家屋・建物の例とは限らないということである。「みや (宮)」も同様で、この語の意味分類は「12320 (君主) / 12630 (社寺・学校) / 14400 (住居)」の 3 つであり、大鏡にはこの語が 228 例ある。「14400 (住居)」で検索すれば、大鏡にこの語が 228 例あることが分かるが、その全てが住居の例とは限らないということである。

大変便利な語彙表ではあるが、個々の作品の性格を考えず、個々の用例を吟味せずに数字だけを操作していると、思わぬ失考をすることがあろう。注意すべきことである。

3. 『日本古典対照分類語彙表』について

旧版に当たる『古典対照語彙表』では、徒然草・方丈記・大鏡・更級日記・紫式部日記・源氏物語・枕草子・蜻蛉日記・後撰和歌集・土左日記・古今和歌集・伊勢物語・竹取物語・万葉集の 14 作品を対象としていたが、新版の『日本古典対照分類語彙表』では、これに平家物語・宇治拾遺物語・新古今和歌集の 3 作品を加えて 17 作品を対象としている。旧版では中古が中心で中世は随筆が 2 作品のみであったところ、これに軍記と説話と和歌とが加わって、比較の幅が広がり、内容が充実した。延べ語数の多い『平家物語』と『宇治拾遺物語』とが加わった意義も大きい。

そして、前節でも述べたように、旧版に対して、『万葉集』の依拠索引が差し替えられたこと、『分類語彙表』の分類番号が付されたことが主な増補改訂となる。

別冊が付き、そこには宮島達夫「古典語の統計と意味」、小木曾智信「Excel による『日本古典対照分類語彙表』データの活用」が収録されている。前者は『日本古典対照分類語彙表』の具体的な活用例であり、後者は『日本古典対照分類語彙表』のエクセル・ファイルの使用方法を解説したもので、エクセルの使い方に習熟していない使用者にとって親切な解説になっている。

付録の CD-ROM には語彙表のエクセルデータのみならず、これらの解説も収録されており、さらに「展開」というフォルダには、書籍には収められていない、出現頻度配列語彙表・意味分類配列語彙表や、『分類語彙表』の意味分類項目の一覧も収録されている。

以下、別冊所収の「古典語の統計と意味」を読んで興味を惹かれたことを 2 点述べる。

6 ページに「6 作品間の用語類似度」という項目があり、17 作品のうちの 2 作品の組み合わせ全てについての類似度が示してある。ここから類似度の高い組合せを高い順に 10 位まで示せば、①古今－後撰、②後撰－新古今集、③古今－新古今集、④蜻蛉－源氏、⑤蜻蛉－枕、⑥枕－源氏、⑦源氏－紫、⑧蜻蛉－更級、⑨枕－紫、⑩枕－宇治となる。逆に、類似度の低い組合せを低い順に 10 位まで示せば、①万葉－紫、①万葉－大鏡、③大鏡－新古今集、④古今－大鏡、⑤万葉－方丈、⑥古今－平家、⑥新古今集－平家、⑧万葉

－平家，⑨後撰－大鏡，⑩後撰－平家となる。古今・後撰・新古今集という3つの和歌集相互の類似度の高さが目立つ。ただ、和歌集でも万葉とこれら3歌集との類似度は高くない。これは時代差によるものであろう。また、蜻蛉・源氏・枕・紫・更級などの中古の作品は、物語・日記・随筆というジャンルにかかわらず類似度が高いが、同じ中古の作品でも大鏡のような歴史物語との類似度は高くない。歌集と大鏡・平家との類似度は低い。和歌と歴史・軍記とでは描かれる世界が大きく異なるのであろう。万葉は上代の作品である故か、多くの作品との類似度の低さが目立つ。

8 ページにある「8 各作品の特徴語上位 20 語」も興味深かった。特徴度の計算法は宮島達夫・近藤明日子「古典作品の特徴度」(『計量国語学』28-3, 94-105 2011年)に拠る由である。この結果を見ると、万葉の上位語は「いも」「あが」「恋ふ」「きみ」「わが」、竹取では「かくやひめ」「おきな」「竹取」「つばくらめ」「たつ(竜)」, 伊勢では「をとこ」「をんな」「むかし」「よむ」「返し」、枕では「をかし」「いみじ」「言ふ」「いと」「わらふ」などという結果になっている。単純に作品毎に用例の多い語を並べた結果では、多くの作品で「あり」「こと」「す」「いふ」「ひと」などが上位に来て、作品毎の特徴があまり見えないのに対し、「特徴度」の計算結果にはそれぞれの作品の特徴がよく出ている。

4. 『万葉集巻別対照分類語彙表』について

本書には「集計表」というページがあって、ここで『万葉集巻別対照分類語彙表』の具体的な使用例が示されている。また、CD-ROMには小木曾智信監修・笠間書院編集部編「Excelによる『万葉集巻別対照分類語彙表』データの活用」が収録されており、エクセルの使用法の解説になっている。

以下、「集計表」を読んで気付いたことをいくつか述べる。

248 ページに、「表L 語数が多い意味分類」という一覧が載っている。これは、今回の増補改訂で加わった『分類語彙表』の要素に拠って抽出したデータである。『万葉集』においてどういう分野の語が多く用いられているのかが分かる。その上位 10 位を示す。

①固有地名 805 語, ②木本 142 語, ③草本 137 語, ④山野 133 語, ⑤枝・葉・花・実 87 語, ⑥鳥類 86 語, ⑦往復 83 語, ⑧枕詞 81 語, ⑨生 77 語, ⑩接近・接触・隔離 74 となっている。語数は異なり語数であり、延べ語数で数えればまた違った結果になる。しかし、異なり語数を見ても、地名が圧倒的に多く、植物や鳥がこれに次ぐなど、いかにもと思われる結果が出ている。ちなみに⑦「往復」には「行く」「来」「去る」「通ふ」などが含まれ、⑨「生」には「咲く」「生ふ」「生く」などが含まれる。こういうデータが得られたのは、「意味分類」という項目が新たに加わったことによる大きな成果といえる。

ただ、枕詞は 81 語という数字で第 8 位に入っているが、これは前述したように「草枕」を枕詞とはせず「寝具」とするなど、実際には枕詞として用いられている多くの語を除外した結果であり、実体を現したものにはなっていない。多義語は複数の意味分類を列挙しているのであるから、「草枕」も「寝具」とするのは良いとしても、そこにさらに「枕詞」も並べてくれていれば良いのと思う。

252 ページに「表N 万葉集の巻のあいだの類似度」という表が載っている。『日本古典対照分類語彙表』には、古典文学作品 17 種それぞれの間の類似度が載っていた。これと同様に、こちらは万葉集の各巻相互の類似度を示したものである。類似度が高い順に

10 位まで示せば、①巻 11 - 12, ②巻 4 - 巻 12, ③巻 8 - 巻 10, ④巻 4 - 巻 11, ⑤巻 17 - 19, ⑥巻 4 - 巻 15, ⑦巻 3 - 巻 7, ⑧巻 12 - 巻 15, ⑨巻 3 - 巻 6, ⑩巻 11 - 巻 13 となっている。数値の上では上位 4 組が突出している。その理由について、「集計表」には、「突出した理由についてはいまちいるいとまがない」とあるが、そう難しい話ではあるまい。巻 11・12 は近畿地方の相聞歌を集めた巻で、分量的な理由で 2 巻に分割したものと考えられる。本来は一体のものともみれば、類似度が最も高いのはよく理解できる。巻 4 は相聞のみを集めた巻であるので、その点で巻 11・12 と類似度が高いのであろう。巻 8・10 は四季分類歌巻であるという共通点をもつ。類似度の高さはそれに由来するものであろう。

逆に、類似度が低い順に 10 位まで示せば、①巻 1 - 巻 16, ②巻 1 - 巻 14, ③巻 5 - 巻 14, ④巻 14 - 巻 18, ⑤巻 5 - 巻 16, ⑥巻 15 - 巻 16, ⑦巻 6 - 巻 14, ⑧巻 14 - 巻 16, ⑨巻 8 - 巻 16, ⑩巻 14 - 巻 19 となっている。最も類似度が低いのは巻 1 と巻 16 との間である。巻 16 は雑纂的な巻で、ユーモラスな歌やナンセンスな歌なども含む。最も格調高い歌を集めた巻 1 との類似度が最も低いというのはよく納得できる。巻 14 は東歌の巻なので、これまた巻 1 とは対極にあるのだらう。類似度が低い組合せ 10 位の中に、巻 14 が 6 つ、巻 16 が 5 つ含まれているのも、この 2 巻の特殊性を良く物語っている。下位第 8 位に⑧巻 14 - 巻 16 の組合せが入っている。この 2 巻同士の類似度もまた低いのである。巻 14 は巻 14 で、巻 16 は巻 16 で、それぞれに独自の存在なのである。それは別にこと新しい事柄ではないが、それが語の類似度からも裏付けられたことになる。

独自性の高い巻 14 ではあるが、それでも巻 14 との類似度が相対的に高い巻をあげれば、巻 11, 巻 12, 巻 20, 巻 4 という順になる。巻 14 は大部分が相聞歌であるので、相聞のみを集めた巻 11, 巻 12, 巻 4 とは類似度が高いのであろう。巻 20 には多数の防人歌が含まれるので、巻 20 との類似度の高さはそれに由来するものと思われる。

こうしてみると、1 巻丸ごと相聞歌を取めた巻 4, 巻 11, 巻 12 や、1 巻丸ごと東歌を取めた巻 14 などに特徴が現れやすい傾向が窺える。巻別の集計では、1 巻の中に相聞と挽歌とを含む巻 2 や、雑歌・譬喩歌・挽歌の 3 部立を含む巻 3・巻 7 などでは巻の特徴が現れにくいのではあるまいか。また、相聞歌に特徴的な語を知りたければ、相聞のみからなる巻を集計すればそれが見えてこようが、挽歌に特徴的な語を知りたくても、挽歌のみで編成された巻は存在しないので、そういう集計はできない。それが残念である。

欲が深すぎる注文と言われそうであるが、この語彙表、巻別に留まらず、部立でも加味してもらえたらという思いを抱いている。すなわち、この語彙表のデータは、巻 1～巻 20 の 20 項目の別で示されているが、これを、たとえば巻 1 雑歌、巻 2 相聞、巻 2 挽歌、巻 3 雑歌、巻 3 譬喩歌、巻 3 挽歌、巻 4 相聞、巻 5 雑歌、巻 6 雑歌、巻 7 雑歌、巻 7 譬喩歌、巻 7 挽歌、巻 8 春雑歌、巻 8 春相聞、巻 8 夏雑歌、巻 8 夏相聞、巻 8 秋雑歌、巻 8 秋相聞、巻 8 冬雑歌、巻 8 冬相聞、巻 9 雑歌、巻 9 相聞、巻 9 挽歌、巻 10 春雑歌、巻 10 春相聞、巻 10 夏雑歌、巻 10 夏相聞、巻 10 秋雑歌、巻 10 秋相聞、巻 10 冬雑歌、巻 10 冬相聞、巻 11 相聞、巻 12 相聞、巻 13 雑歌、巻 13 相聞、巻 13 問答、巻 13 譬喩歌、巻 13 挽歌、巻 14 東歌、巻 15 前半、巻 15 後半、巻 16, 巻 17, 巻 18, 巻 19, 巻 20 と分けて貰えなかったものかと思うのである。これで項目数は 46 である。現在の 20 と比べて 2.3 倍になるが、部立の要素を入れることで、より立体的な万葉集の分析が可能になり、語彙表の価値はさ

らに高まったことであろう。そういう望蜀の嘆とでもいうべき思いをもっている。

5. 追記

両書の編者である宮島達夫氏は、平成 25 年 12 月 5 日に 84 歳で逝去された。謹んでご冥福をお祈り申し上げます。

(2016 年 3 月 31 日受付)