

# JapanKnowledge Lib.を活用した研究

島根大学学術情報機構附属図書館長 田 籠 博

## 1 研究計画

定年を目前にして一つの計画を立てた。ある文献資料を電子化し、その用語に関する研究成果を発表して退職を迎えようと思った。対象にしたのは室町時代の言語資料として有名な桃源瑞仙の『史記抄』（文明9年1477成）。この活字本『史記桃源抄の研究 本文篇』6冊（計3,000頁）をOCRで取り込み、パソコンで自在に利用できる形に整え、それによって用語の性格を明らかにしようと考えた。

3月から作業を始め、休日も趣味を諦めながら多大な時間を費やして、8月末ようやく第一段階を終えた。しかし、用語の性格を検討する方向で中々適切なアイデアが得られず、しばらく索引作成などでお茶を濁す時期もあった。

そんな時、附属図書館のHPから使用できるデータベースの中に「JapanKnowledge Lib.」があり、そのコンテンツに筆者が日頃利用する『日本国語大辞典』（小学館）が含まれていることを「発見」した。

以下の記事は、それを利用した筆者の経験にもとづく。

## 2 『日本国語大辞典』の用例探し

日頃使用する辞典は大型本で、当然のことながら重い。棚から取り出し、頁を繰って目的の項目を探し、語釈（語の意味・用法の説明）を読み、そこに引かれている用例を見る。

一方、データベース化された辞典では、冊子体では思いも寄らない調査方法を取ることができる。『日本国語大辞典』（『日国』）全体から、『史記抄』の用例がいくつ、どの語項目にあるかを瞬時に探せるのである。

筆者は、『日国』における『史記抄』用例の在り方を調査することによっ

JapanKnowledge Lib

メニュー | 購入情報 | お問い合わせ | よくある質問 | ENGLISH | 最新データベース | ログアウト

基本検索 | 詳細(個別)検索 | 本棚

目録体系的検索  
 範囲:  条件:  検索  
 見出し  前向き一致  
 かつ(AND)  見出し  前向き一致  
 かつ(AND)  見出し  前向き一致 クリア

項目種別

親見出し項目 (2988)  
 自立語・付属語: 単語などの単語の頭  
 子見出し (117)  
 漢字・ローマ字の頭  
 字音要素 (1)  
 漢字とローマ字の字音要素

品詞

名詞 (284)	代名詞 (28)
動詞 (55)	形容詞 (78)
形容動詞 (68)	副詞 (185)
連体詞 (7)	接続詞 (14)
感動詞 (37)	助詞 (8)
助動詞 (33)	助数詞 (30)
接尾語 (3)	連語要素 (8)
漢字読本 (1)	漢語 (13)
慣用語 (8)	

ジャンル

人名 (1)	地名 (3)
姓 (1)	番号・作品名 (3)
日本の地域 (8)	神社 (1)
動物名 (8)	植物名 (8)

品類

図辞書 (1)

日本国語大辞典

3098 件

- あい【あひ】【合】  
 項——七「この裏の目のあひの、例の裏の目のあひよりも近くて、目のせもかはりたけければ」(史記抄)  
 【1477】一四・扇語會公列伝「扇小ならば指のあひをい」...
- あいしらい【あひらうり】  
 愛す答ふこと。あえしらい。「い人に對して應對すること。待遇すること。また、もてなし」。(史記抄)147  
 7三・周本紀「魯武公の來朝、たる時、あひしらいに、不 ...」
- あいしるし【あひ-】【合】  
 【名】(1)戦場で敵味方の区別をはっきりさせるために、笠、袴その他、武装の一部に付ける一定の目印。\*  
 (史記抄)1477六・項羽本紀「一軍のあひしるしにぞ」\*星見 ...
- あいさい【あひ-】【指】  
 是ねても音のからは、此質屋も指時(あイタイ)と思はる(4)対等であること。対等で事をなすこと。(史記  
 抄)1477九・考武本紀「野・衛」と云は積 ...
- あいたがい【あひたがい】【指】  
 九・元寶兵衛入道軍事「其の外命を輕んじ、書を置んじて、衛にて勝負を決せん」と程互に互戦ひける」(史記  
 抄)1477四・素本紀「あひたがいのことでさるほどにあたへられ ...」
- あいでてひさ  
 【名】形容詞(あいでてひさ)の語幹に、接尾語(ひさ)の付いたもの(あいでてひさ)のこと。また、その原音。\*  
 (史記抄)1477一五・羽扇明伝「我が子を食へばあいまり ...」
- あいてくみ【あいて-】【指】  
 【名】他の者を交えなくて、二人以上で対すること。また、その組み合わせ。\*相手舞。(史記抄)1477  
 六・素本紀「あいてくみも手紙の書をあいてくみしてやれた ...」
- あいま【あひ-】【合】  
 \*申楽談(1430)神事委任の事「神事を本(ほん)」して、そのあひまの身上動かしたため、上下なり」

図 『史記抄』 検索結果

て、『史記抄』の用語の性格が測れるのではないかと予想した。日本語の代表的な辞典である『日国』における利用状況は、間接的ではあるが、『史記抄』の用語の性格を何らかの形で反映しているに違いない。

実際の手順は次の通り。附属図書館のHPから「電子リソース」を選び、「分野別データベース」の下にある「事典・辞書」の「JapanKnowledge Lib」を選択する。タイトル下の「詳細(個別)検索」にカーソルを当てると「日本国語大辞典」が現れるから、それをクリックする。

検索欄に「史記抄」と打ち、範囲を「全文(見出し+本文)」として実行すれば、たちまち検索結果が表示される。「史記抄」を含む見出し項目数、品詞などが左に示され、「史記抄」前後の記事が右欄に五十音順で表示される。親見出し項目が2,980、子見出しが117、字音要素が1である。

作業が簡単に行くと欲が出る。同じ室町時代の文献資料についても同様に調査して比較しようと考えた。同類の抄物、キリシタン資料、狂言資料について試みた。これも簡単に終えたので、他の時代の著名な文学作品についても同じ作業を繰り返した。

至って順調で楽しくなる。せっかく新しい調査方法を採るのだから、楽に行えるのが望ましい。全ての結果は表のようになる。

表 『日国』所見文献名の項目数（親見出しと子見出しの合計）

史記抄	3,068	万葉集	8,736
毛詩抄	805	源氏物語	11,598
玉塵抄	2,311	枕草子	3,454
中華若木詩抄	1,192	今昔物語	4,447
虎明本狂言	3,686	平家物語	7,668
天草本平家物語	903	徒然草	2,433
天草本伊曾保	1,024	太平記	9,794

表から、『万葉集』『源氏物語』『平家物語』などには及ばないものの、『日国』が『史記抄』から数多くの用例を引いていることが分かる。その事實は、『史記抄』が室町時代語の資料として重要なものであることを示している。

### 3 用例検索の陥し穴

しかし、前掲の表には重要な問題点がある。文献名で検索した場合、目的のものだけでなく、類似する書名の文献までもが検索結果に含まれていることが判明したからである。

『徒然草』は江戸時代初めに大流行し、類似の書名をもつ作品が数多く作られた。表の数字には、例えば、『徒然草講談之事』『徒然草野槌』なども含まれている。『太平記』はさらに多く、『難太平記』『娘太平記』『化物太平記』など「～太平記」という書名、および『太平記大全』『太平記聞書』といった「太平記～」という作品も多く、全部で20作品以上もあった。

こうした不要な項目を排除するにはどうするのか。予め類名書が分かっているならば、それを別に検索し、その数を全体の項目数から除けばよい。しかし、類名書の全てを想定することはとてもできない。

実際には、大変素朴な話だが、検索した項目を一つ一つ見て、目的の書名と異なるものを数え上げ、それを差し引くことになる。『太平記』で試してみると、510項目が本来の『太平記』とは異なる書からの引用であった。1,000を超える項目を丹念に見ていくのは、相当の時間と根気を要する。

## 4 検索結果の問題

『日国』の検索ではもう一つ問題がある。検索結果として出力されるのが見出しの項目数であって、用例の数ではないことである。これは筆者などの調査では重要な相違を生み出す場合がある。極端な例だが、『徒然草』を一項目中に20例も引いている場合がある（助詞「に」の項）。従って、正確な用例数を知りたければ、やはり一項目ずつ記事を見ていかなければならない。

もう一つ、解説文中に現れる書名も検索対象になるため、用例が存在しない項目も含まれることがある。これも正確な結果を得るためには除外する必要がある。

三つめに、検索自体の問題ではないが、項目の表示数が系統的に1,000を超えたものは表示できないという制限がある。逆順にすれば2,000までは何とか可能だが、それ以上になると色々と工夫する必要がある。一般の利用であれば問題にならないことだが、研究的な利用では困ったことで、何とか改善して欲しい所である。

## 5 複合検索の効用

筆者の調査内容は、『史記抄』の用語が『日国』で単独例（『史記抄』の用例しかない項目）や初出例（その項目で最も古い例）、または最初例（語釈の各項目で最も古い例）かを調べることであった。その結果、全体の6割が何らかの意味で初出または最古の例であることが判明したが（因みに、『徒然草』は2割）、それについては今は述べない。

初出例や最初例かどうかは、語釈（語の意味・用法の説明）の後、用例の最初に『史記抄』があるかどうかで判断したが、そこで不可解な事実があることに気づいた。室町時代の辞書『文明本節用集』と『史記抄』との先後関係が、『日国』の中で定まっていなかったのである。

『史記抄』と『文明本節用集』との二つを鍵にして検索すると、『史記抄』を先に置くのが60項目、逆に『文明本』を先にするのが47項目あって矛盾し、初出例・最初例の判定に混乱をきたすのである。

わが国で最も信頼される『日国』に、こうした不統一があることを見出したのは興味深かった。冊子体の辞書では、個別の事例には気づいても、二つ

の書名を複合検索すれば直ちに結果を出すデータベースには到底及ばない。これは予想していなかった利用法である。

JapanKnowledge Lib.では三つの鍵語までの複合検索が可能だから、『万葉集』と『源氏物語』に用いられて夏目漱石の作品でも用いられている言葉を探す、などが簡単にできる。検索方法（AND,OR,NOT）を工夫すれば、筆者も気づかなかった利用法がまだ隠れているかも知れない。

研究発表の資料を昔はガリ版刷りで作成していたと言っても、最早だれも理解できない。資料作成にはワープロが当たり前になり、OHPも過去の道具となった。だとすれば、ここで紹介したような研究方法も一つの手段として認めてもらえるのかも知れない。

最後に、今後の利用拡大を考えると、本学でのユーザー数が2でしかないのは是正したい所である。

