

史書の本文の研究を支援するツールについて

1H-5

星野 聡 村尾義和

京都大学大型計算機センター

筆者らは、日本古代の正史である日本書紀とそれに続く続日本紀の全文を一つのデータベースとして文献情報検索システムPAIRSに格納し、さらに、これらを同時に検索できる一字索引も作成した。日本歴史の分野では、歴史書テキストのオンライン検索、複数の史書の一字索引作成は今迄に見当たらない。しかし、続日本紀の研究には日本書紀の参照も必要であるなど、複数の史書にまたがる考察が重要である。この他、頭注の処理、二字索引作成などについても述べる。種々の型式の索引があれば、研究の手掛りを見出すのに有効である。

(1) 歴史書本文のデータベース化

日本書紀は神代より持統天皇(西暦697年)続日本紀は文武天皇即位から桓武天皇(697~791)の間をカバーしており、この時代の歴史、文学などの研究に最も重要な史書である。これらの分野の研究レベルが高くなっている現在、一字索引は研究上で必須の道具になっている。しかるに、日本書紀には国学院大学で手作業による一字索引があるが、⁽¹⁾続日本紀には未だにない。これは膨大な量になるので手作業では限界がある為である。

筆者らは、先に京都大学人文科学研究所荒井健教授、同研究所附属東洋学文献センター勝村哲也助教授その他の方々の御援助を得て、漢籍本文のデータベース及び一字索引の作成を行ったが、そのソフトウェアに若干手を加えて、今回日本書紀と続日本紀の本文のデータベース化を行った。⁽²⁾

使用したテキストは入手の便を考慮して国史大系本を使用し、異体字は差し支えない場合には字体を統一化したが、慣用の字は索引の見出しに旧字(正字)と並べて出力

している。このため異体字に対し、置き換える文字と処置を記した表を用いている。また、日本書紀を続日本紀と同一形式とする必要上、各干支に日付の記入を行った。続日本紀の巻数は日本書紀の続きに一連番号とした。続日本紀の本文には123個の文字種がJEPになく、文字パターンを作成した。続日本紀本文の文字数(316222個)がほぼ同じ文字数(308062個)を有する樊南文集で714個の外字パターンを作成したのに較べるとカバー率は遙かに良いが、これは日本の書でしかも文学書でないからであろう。データベース検索例を図1に示す。項目A700の内容は指定された句が存在する巻(2桁)・ページ(3桁)・行(2桁)である。巻の情報は史書の成立過程を研究する際に有効である。

(2) 一字索引の作成

検索システムに格納されたデータベースの内容をファイルに出力し、これを分類・編集して一字索引を出力する。この一部を図2に示す。ラインプリンタ用紙一枚に3列、この用紙二枚分を1ページに格納するとして、続日本紀で1291ページ日本書紀との合本で1851ページとなる(但し本文のみの場合)。○印は見出しの漢字が句中にある位置をしめす。

排列に用いる音は、樊南文集の索引で行われたように索引作成者により指定できる。

```
RS> AND A300 EQ @賑給@
      マスタファイル検索を行います。
      40 みつかりました。
      RS> OUTPUT
```

```
#1
BANGO      02309750
A010      続日本紀巻第卅二
A300      賑給之
A700      6240410
```

図1 検索例

(3) 頭注の処理

日本書紀・続日本紀の国史大系本では注釈は本文の上部に漢文体で記されている。その内容は現本文の元になったテキスト(底本)からの差異Dか、別のテキストにある当該箇所(の)の記述、あるいは筆者の説Eから成る。Dを用いると元のテキストへの修正ができる。Eは単なる注釈である。これらを扱うため記入用紙を作製し、元テキストへの復元を行った。同じ用紙は現テキストの誤りを修正するためにも用いられる。この用紙に修正箇所の開始・終了位置、修正・注釈の区別、根拠、注釈の内容などを記入できる。

4) 本文の表示と二字索引の作成

日本書紀と続日本紀など(六国史)は漢文体で記されているが、国史大系本では限られた箇所だけに空白があるだけである。従って年、月、日により先頭の開始位置を順にずらして印字して判りやすくした。

次に、二字索引は本文中の連続したすべての二文字に対し、出現する位置を示す索引である。熟語は二字のものが多く(例えば日付は二字の干支で示される)ので、二字索引は熟語その他の全巻に於ける分布を一覧するのに便利である。この例を図3に示す。横方向は続日本紀の巻数1~40に対応し、記された数字nはその巻での使用率が全巻の $10n+1 \sim 10(n+1)$ であることを示す。ただし全く用いられていない巻には*印を表示した。巻の途中から使用され、又は使用されなくなる場合は編集者の違いや制度の変化を意味することがあり、研究の手掛りを与えるので有用である。三字索引も同様にして作成できる。他の六国史への適用や類聚索引作成などは今後の課題である。

続日本紀のデータベース化は昭和60年度

データベース作成費、総合研究「東洋学研究支援データベースの研究」で行われた。謝辞 文字パターンの準備作業を行って頂いた本センター桶谷氏に感謝します。

参考文献

- (1) 日本書紀総索引、全4巻(中村啓信編 国学院大学日本文化研究所) 角川書店(1964~1968)
- (2) 星野聰、村尾義和: 続日本紀のデータベース化、研究発表報告集 第1号 京都大学大型計算機センター(1986)
- (3) 日本書紀、前後篇及び続日本紀、前後篇(新訂増補国史大系) 吉川弘文館(1984)

【域】(イ)		
膽望○	中國	03-120-01
我生○	倭國	05-160-02
悉往○	倭國	06-186-05
遠絶○	倭國	06-192-08
養之○	陵	08-232-01
煙氣○	不陵	11-296-01
因以○	逃起	13-341-05
縦削○	武内	17-437-04
普天○	下莫	18-459-05
壁鏡○	區	18-459-07
日赴○	新羅	19-486-02
則外○	外積	22-578-04
其外○	方七	25-653-02
其外○	方五	25-653-04
吾昔○	往西	25-653-05
吾從○	西自	31-005-16
其奉○	君使	31-006-03
以奉○	使絶	33-018-04
並庭○	以奉	33-021-14
朝河○	庭異	33-028-13
山諸○	又諸	37-070-09
汝奉○	汝奉	40-111-13
		40-122-07
		64-428-05

図2 一字索引の例

(日本書紀と続日本紀を同時に検索)

図3 二字索引の例

(左側の数字は出現した最初と最後の頁及び回数で左端の*印は300頁以内の範囲でのみ(2回以上)用いられている場合)

409	409	1	賑飢	*****9*****
2	554	141	賑給	0***0*****110***10***1*934211202710101*3
* 28	132	7	賑貧	***4*49***4*9*****
13	550	60	賑恤	*093111200031*002030*0***10***0***000013
524	524	1	賑贈	*****9**
* 13	35	3	給稻	*9*4*****
462	462	1	給桶	*****9*****
* 15	52	14	給藥	*14933*****
* 2	45	5	給醫	*9*4*****
117	117	1	給錢	*****9*****
14	14	1	給鑑	*9*****