

万葉集の文字の使用頻度調査

大駒誠一
永坂弘子

1. はじめに

インターネット経由で入手した万葉集全文の中から、その歌に用いられた文字の使用回数などを数量的に調査したところ、いくつか興味ある結果が得られたので報告する。

2. 万葉集

万葉集は、言うまでもなく、仁徳天皇時代の磐の姫の歌（5世紀半ば）から天平宝時三年正月一日（西暦759年）の犬伴家持の歌まで約300年間という長い期間の歌4516首を収めたわが国最古の和歌集である。万葉集に収められている歌には、長歌、短歌、旋頭歌、仏足石歌といった歌体に分類され、全20巻で構成されているが、その成立時期や編纂者についての記録は現存せず、はっきりとはしていない。また万葉集の原本は残っておらず、写本のみが伝わる。万葉集は、まだ平仮名、片仮名が作られる以前の和歌集で、すべて漢字で書かれている。これらの漢字は正訓字、訓仮名・音仮名、義訓・戯書に分類される。

3. 吉村のテキスト原文

山口大学教育学部吉村誠教授のインターネット上のホームページには、万葉

集原文テキスト (<http://yoshi01.kokugo.edu.yamaguchi-u.ac.jp/manyou/manyou.html>) が公開されており、誰でも利用できるようになっている。このホームページに公開されている原文は西本願寺本を底本とし、その他の写本による校訂をも取り入れたもので、他にも独自の訓読を施した訓読文をもとにした仮名書き文のファイルなどがある。本稿ではその原文テキストのみを利用した。

万葉集には JIS 規格の漢字コードにない漢字がたくさん使われているが、それに対して、このテキストでは、外字を作ることはせず、「い」とか「R」などと漢字以外の文字で代用している。代用した漢字については、別のファイルにおいて、1 字をいくつかの JIS 規格で表示可能な文字で表して説明している。いろいろな種類のコンピュータ、OS、ソフトウェアが氾濫している状況では止むを得ない賢明な選択であったといえよう。

4. 文字数

表 1 は万葉集に使用されている文字の各巻毎の使用回数である。これを見ると、万葉集の総文字数は約 165,000 字で、巻の中で最少は第 1 巻の約 4,900 字、最多は第 12 巻の約 12,400 字であり、平均は約 8,200 字である。

万葉集は次に示す例のように、歌、題詞、挿入注、左注とよばれるもので構成されている。

幸_二于紀伊國_一時川嶋皇子御作歌 或云山上臣憶良作（題詞）

34 白浪乃 濱松之枝乃 手向草 幾世左右二賀 年乃経経去良武（歌）

一云（挿入注） 年者経経尔計武（歌）

日本紀曰 朱鳥四年庚寅秋九月天皇幸_二紀伊國_一也（左注）

表 1 はこれを別々に集計しその合計を示したものである。吉村の原文テキストは、上記のように「一云」以降の漢字の書き換えの部分も歌として分類している。本稿もこれにしたがい、この書き換え部分も文字の集計に加算している。また、本稿においては、万葉集で使用されている漢字を文字と呼んで統一して

表1 万葉集 巻別・種類別文字数

種類 巻	挿入注	題詞	左注	歌	巻毎の計
1	25	890	989	2,978	4,882
2	137	1,613	691	5,830	8,271
3	31	1,954	639	6,966	9,590
4	20	2,002	591	6,693	9,306
5	238	3,472	283	5,489	9,482
6	4	1,620	979	4,946	7,549
7	11	148	243	6,650	7,052
8	36	1,981	651	5,472	8,140
9	16	834	380	4,730	5,960
10	55	207	181	10,309	10,752
11	52	25	232	8,750	9,059
12	46	29	194	7,285	7,554
13	23	176	336	6,126	6,661
14	37	19	263	7,540	7,859
15	14	439	369	7,396	8,218
16	37	794	1,465	3,179	5,475
17	81	1,757	1,122	7,751	10,711
18	45	1,312	716	5,431	7,504
19	51	1,471	1,012	5,747	8,281
20	18	1,114	2,780	8,470	12,382
種類別 の合計	977	21,857	14,116	127,738	164,688

いる。以下の、文字に関する考察は、すべて、歌、題詞、挿入注、左注の内の歌の部分についてだけのものである。

5. 頻度順

表2は漢字の使用頻度の多い順に40位まで並べたものである。最も多く使われている文字「之」で4,922回である。この表2には現れていないが、文字の種類総数は1,925字。そのうち、万葉集全20巻中、使用回数がたった1回だけの文字は414字、2回だけの文字203字、3回の文字は137字である。

表2 文字の巻毎の使用回数 (40位まで)

巻 漢字	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	計
之	102	318	263	253	75	193	350	217	190	491	338	338	301	98	226	126	275	198	267	303	4922
尔	87	189	237	183	179	144	173	225	144	277	174	194	136	278	275	120	303	191	277	303	4089
者	54	170	197	215	10	174	172	152	129	322	216	244	187	3	12	62	10	8	67	2	2406
乃	113	147	210	118	67	171	74	150	107	134	109	89	101	180	48	77	134	81	71	211	2392
乎	48	103	115	148	49	60	98	78	65	135	132	138	93	183	145	46	136	83	106	168	2129
毛	34	70	84	109	45	40	58	85	55	90	90	135	55	162	229	40	119	86	83	107	1776
可	11	16	44	39	142	14	31	53	15	63	40	47	35	292	223	23	178	138	103	195	1702
波	14	24	42	32	180	19	23	32	25	15	21	11	15	267	171	29	185	147	92	350	1694
奈	16	11	27	31	150	12	10	43	26	22	9	11	16	275	196	13	245	149	87	270	1619
能	27	18	36	9	192	17	11	23	10	22	13	11	13	230	266	4	196	190	87	180	1555
不	23	82	107	145	5	57	101	62	69	155	212	181	104	2	0	26	3	8	31	6	1379
久	17	33	31	32	129	14	19	26	24	28	25	17	23	172	152	18	170	116	82	214	1342
見	49	89	111	96	8	81	107	72	61	113	131	110	53	22	42	26	52	39	52	24	1338
良	35	26	24	23	148	20	24	35	25	50	15	11	15	172	180	30	152	106	69	155	1315
多	9	20	6	25	137	18	10	12	12	14	19	19	12	162	199	9	224	139	58	207	1311
等	19	35	18	7	145	17	25	8	19	30	20	15	24	106	176	19	145	134	95	211	1268
麻	10	17	13	9	129	6	12	11	11	5	3	6	13	209	198	13	185	108	60	228	1246
而	33	96	107	104	3	59	82	75	86	133	105	134	108	4	2	42	5	2	60	1	1241
吾	27	75	66	149	3	43	92	74	52	133	177	128	85	3	0	44	7	2	38	1	1199
母	16	25	21	17	120	13	16	24	17	21	19	28	36	162	132	11	141	101	63	173	1156
有	33	77	106	120	7	58	70	78	52	121	109	95	66	0	0	31	0	1	23	1	1048
都	10	13	24	10	91	11	6	24	6	12	8	3	15	132	133	8	138	110	93	193	1040
流	16	34	29	34	73	21	16	30	17	18	19	20	12	81	144	18	133	97	44	149	1005
安	8	12	9	9	19	7	5	9	3	4	4	15	9	182	235	5	165	124	43	127	994
我	15	14	25	20	46	12	24	23	14	31	51	30	21	143	121	25	99	68	48	137	967
美	11	8	12	5	72	10	4	12	4	7	5	6	3	121	126	4	151	102	53	230	946
伎	3	8	6	0	64	3	3	5	4	1	0	0	5	125	196	7	137	98	55	217	937
伊	11	15	28	9	95	4	6	15	4	8	7	5	17	149	170	8	114	76	44	150	935
夜	9	17	19	41	52	12	29	33	17	85	48	43	28	99	51	18	84	73	25	98	881
比	5	7	15	12	104	5	4	7	9	12	10	17	6	115	123	9	128	90	39	143	860
山	49	52	84	38	2	59	83	48	50	102	59	47	64	0	21	18	14	5	29	7	831
将	3	40	75	86	1	42	76	43	41	96	83	106	40	0	0	37	0	0	21	0	790
知	19	46	24	23	58	34	21	7	12	21	37	29	34	44	66	8	99	47	36	97	762
礼	12	16	15	14	56	9	4	23	12	27	5	9	8	71	130	11	104	70	50	75	721
来	18	39	35	46	2	33	49	44	43	130	73	46	67	1	0	39	6	3	26	1	701
戀	6	28	18	86	0	11	25	31	21	106	161	132	41	0	1	8	3	0	10	2	690
思	16	20	31	37	4	18	25	29	13	37	51	38	52	144	60	7	57	13	13	10	675
許	9	9	6	26	52	5	14	23	8	13	12	11	6	104	79	8	105	80	34	66	670
佐	12	14	15	8	81	10	12	15	5	9	2	5	2	91	44	14	111	65	13	121	649
為	17	33	54	53	7	27	67	20	27	57	76	68	44	12	7	42	8	5	15	9	648

6. 巻毎の使用回数

表2の各文字を巻毎に見ていくと、その使用回数にはばらつきがあり、かつ、そのばらつきにはだいたい三つのパターンがある。このパターンをよく表している三つの文字「之」と「波」と「者」の各巻毎の使用回数をグラフにしたのが、図1、図2、図3である。この図2と図3を見ると、明らかにポジティブとネガティブの関係にある。この関係を瀬古の分類⁽⁵⁾——第1、2、3、4、6、7、8、9、10、11、12、13、16、19巻を正訓中心の用字の巻、第5、14、15、17、18、20巻を借音中心の用字の巻とする——にあてはめると「波」(図

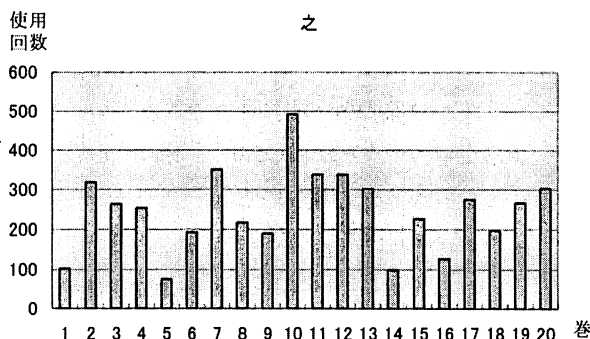


図1 「之」の巻毎の使用回数

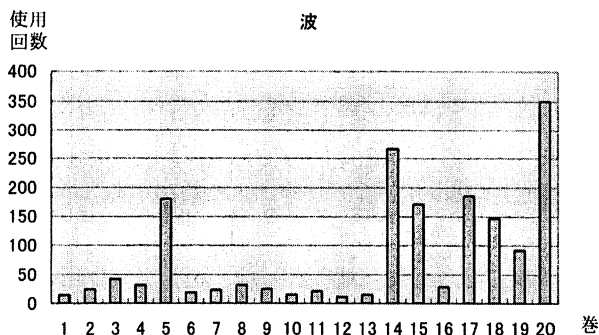


図2 「波」の巻毎の使用回数

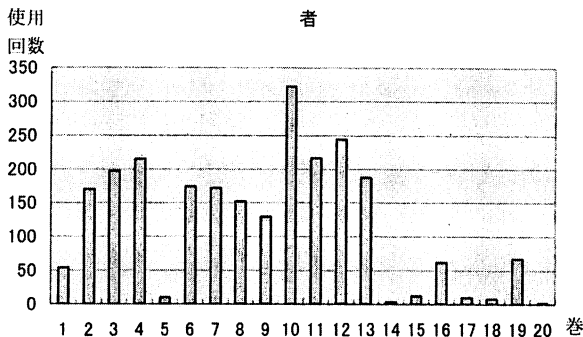


図3 「者」の巻毎の使用回数

2) は借音中心の用字の巻で使用されており、「者」(図3)は正訓中心の用字の巻で使用されていることがわかる。ところが、「之」(図1)は図2や図3に見られるような際立った特徴はなく、ほぼ全巻通して使用されている。このことから万葉集の文字は、その巻の用字法に影響されるものとされないものに大別され、さらに、影響される文字については、「正訓中心の用字の巻」か「借音中心の用字の巻」どちらかに分かれている。

7. 巻毎の文字種

各巻でそれぞれ何種類の文字が使われているか、その文字種数をグラフにしたのが図4である。文字種数の最多は第11巻の943種、最少は第14巻の220種であるが、巻毎に見ると、正訓中心の用字の巻(以降、正訓中心巻)ではいずれも500種以上の文字が使われ、借音中心の用字の巻(以降、借音中心巻)では500種以下であることがわかる。1音1字とすれば、いろは48字に、濁音20字(万葉集では半濁音は使っていない)を加えた約70字あれば仮名として十分なわけであるが、第5巻などは451種もの字を使っている。これは、一つの音に対して、例えば、「き」として支、岐、吉、企、枳、棄、伎などとたくさんの文字が使われているからである。

表3は万葉集が最初に成立したのが第1と第2巻であることを前提に、第

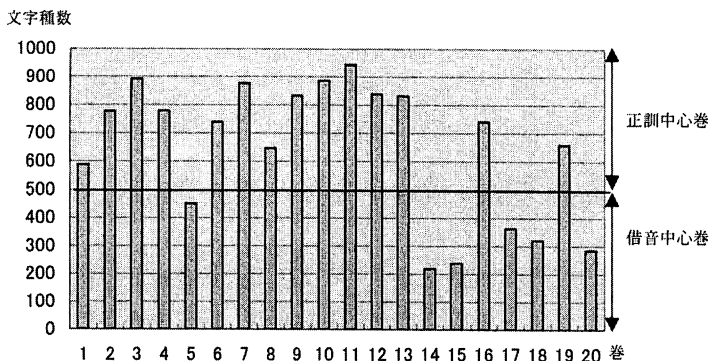


図4 巻毎の文字種数

1・2巻で使用された文字が第3巻以降にどの程度使用されているかを調べたものである。ここでは各巻の総文字種数に対し、第1・2巻にある文字種が何%

巻	1巻+2巻の文字種 (%)	新出の文字種 (%)	中心巻
1	100.0	100.0	正訓
2	100.0	40.4	正訓
3	73.7	26.4	正訓
4	77.0	13.9	正訓
5	74.9	15.1	借音
6	80.0	7.7	正訓
7	71.1	13.2	正訓
8	80.7	6.2	正訓
9	74.6	8.9	正訓
10	70.9	6.8	正訓
11	70.0	8.1	正訓
12	72.9	6.1	正訓
13	73.3	5.4	正訓
14	86.4	1.4	借音
15	88.3	1.3	借音
16	71.8	7.7	正訓
17	87.3	0.6	借音
18	85.7	1.2	借音
19	79.2	3.0	正訓
20	84.2	1.5	借音

か、新出文字が何%か調べたものである。これによれば、第3巻以降、どの巻も第1・2巻で使用された文字によってほぼ構成されていることがわかる。正訓中心巻においては第1・2巻の文字種が70.7～80.7%を占め、新出の文字種は第3巻の26.4%から第19巻の3.0%へと減少している。これに対し、借音中心巻では、第5巻は第1・2巻の文字種の使用は74.9%、新出文字種は15.1%で、正訓中心巻との違いはあまり見られないが、その他の借音中心巻第14・15・17・18・20巻は第1・2巻の文字種数の使用率が84.2%～88.3%と正訓中心巻に比べて高く、新出の文字種数は1%前後と少ない。これは、第5巻を除く借

音中心巻では、第1・2巻で使用された文字種によって大部分が書かれていることを示している。

8. 正訓中心巻の文字と借音中心巻の文字の比較

次に、主に正訓中心巻で使用されている文字と、借音中心巻で使用されている文字について調べてみた。ある文字が正訓中心巻と借音中心巻において平均何回使われているか数え、その比をとって見る。その比の値をその文字の訓音比と呼ぶことにする(表4)。

訓音比 = 正訓中心巻の平均使用回数 / 借音中心巻の平均使用回数

この比が大きければ、その文字は主として正訓字中心巻で使われ、小さければ主として借音中心巻で使われていることになる。その中間の値の文字は正訓字中心巻、借音中心巻の両方で使われていることになる。

図5は全20巻で使用回数が100回以上の文字の訓音比の頻度分布である。横軸は訓音比の対数で、縦軸はその訓音比をとる文字の個数である。ただし、右端は、借音中心巻では1回も使用されず、訓音比の計算ができなかったものである。すなわち、全体では100回以上使用されているが借音中心巻では1度も使われなかった文字が9字あった。

この図5を見ると見事に山が二つあり、図の左方が主として借音中心で使われている文字、右方が主に正訓中心巻で使われている文字である。中央部は両

表4 訓音比の例

	波	者	之
正訓中心巻14巻の使用回数、a	394	2361	3817
正訓中心巻14巻の平均使用回数、 $b = a / 14$	28.1	168.6	272.6
借音中心巻14巻の使用回数、c	1300	45	1175
借音中心巻6巻の平均出現回数、 $d = c / 6$	216.7	7.5	195.8
借音比、 b / d	0.13	22.5	1.37
借音比の対数、 $\log_{10}(b / d)$	-0.89	1.35	0.14

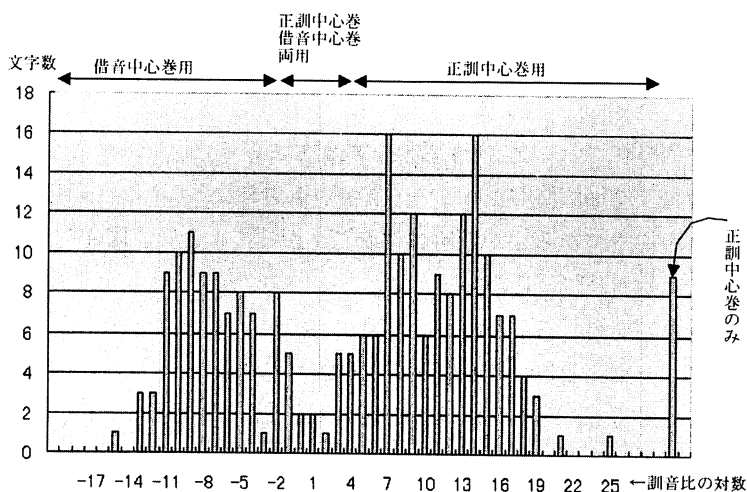


図5 訓音比の分布

表5 用途別文字

- (a) 主として正訓中心巻用に使われている文字、137字

者不而吾有山将来戀為人念妹君跡所相子去無手立二花名玉鴨聞雖鳴今物八時白木秋
鳥如神春日三此海國出待在言津公居朝丹雲間方乍行裁浪水常従草衣成小裳風田師高
道渡真隠事往云上四石落葉零座開嶋耳益屋足心川原下生十莫引雪瀬取情奥明過浦
芽根長荒鹿音越置雨打藻千宮妻巻通社令袖經た戸露繁咲依

- (b) 主として借音中心巻用に使われている文字、62字

可波奈能久良多等麻母都流安我美伎伊比礼許佐里須布弓登牟和婆於保由奴加宇祢氣
利倍志古呂米敝己阿太受豆具藝杼理余刀蘇追賀斯努C吡

(CはJ I S漢字コードにない字)

- (c) 正訓中心巻にも借音中心巻にも使われている文字、50字

之尔乃乎毛見夜知思日曾家々左吉野大天香世月欲武何也河故与邊兒其御遠勢敷船年
夫宿未末留度自代射梅得是弥

方、すなわち、全般的に使われている文字である。

ここで、万葉集で使われている文字を、正訓中心巻用、借音中心巻用、両方の3種類に分類した。その境目をどこにとるかは問題であるが、ここでは一応、訓音比が3以上を主として正訓中心巻用、1/3以下を主として借音中心巻用とみなし、その中間の値をとる文字を正訓中心巻、借音中心巻両用とした。全20巻で100回以上使われている文字249種について分類したのが表5である。左上より右下へ使用回数の多い順に並べてある。

9. むすび

万葉集で使われている文字を数量的に検討してみると、いろいろなことがわかった。特に、巻毎に使用されている文字の種類を数えると、正訓中心巻であるか、借音中心巻であるかが明瞭に浮かび上がってくるし（図4）、文字毎にどの巻で何回使われているかを数えると、その文字が主として正訓中心巻用であるか借音中心巻用であるかわかる（図5、表5）、などなどである。

こうして、万葉集で使われている文字を調べてみると、いろいろ不思議に思われる点もでてくる。その一つは、正訓中心巻と借音中心巻とが混在していることである。もし、万葉集の編纂に長い時間がかかり、その途中で万葉仮名の使用が始まったとすれば、それ以降はすべて、借音中心巻となりそうなものである。だいたい13巻以前が正訓中心巻、14巻以降が借音中心巻であるが、借音中心巻の5巻だけ正訓中心巻に混じって前の方に入り込み、正訓中心巻の16巻と19巻が後ろの借音中心巻のグループに混じっている。筆写を繰り返しているうちに巻の順序が乱れてしまったなどということはないのだろうか。

最後に、誠に残念ながら、JIS 漢字コードの不足により、万葉集のすべての漢字をパソコンであつかうことができない。このためにかなり不便を感じた。それこそ、日本のすべての古典や固有名詞で使われている漢字がどのパソコンでも利用できる環境が望まれる。

10. 謝辞

万葉集の原文をコンピュータで処理できる形にして、インターネット上に公開している山口大学の吉村誠教授に感謝する。

11. 参考文献

- (1) 伊藤博、万葉集釋注原文篇、集英社、2000年5月。
- (2) 森淳司、万葉集研究入門ハンドブック、雄山閣、1988。
- (3) 古屋彰、万葉集の表記と文字、和泉書院、1998。
- (4) 古橋信孝、万葉集一歌のはじまり、ちくま書房。
- (5) 瀬古確、万葉集に於ける表現の研究、風間書房、1966。