

国立国語研究所学術情報リポジトリ

On the frequency ratio of kinds of letters in all sorts of sentence

メタデータ	言語: jpn 出版者: 公開日: 2017-06-13 キーワード (Ja): キーワード (En): 作成者: 佐竹, 秀雄, SATAKE, Hideo メールアドレス: 所属:
URL	https://doi.org/10.15084/00001316

各種文章の字種比率

佐 竹 秀 雄

1. 目的

現在、日本語で表記されている文章の多くは、漢字ひらがな交じりである。その漢字とひらがながどの程度の割合になっているのかを知る一つの指標として、漢字含有率がある。たとえば、雑誌や週刊誌における漢字含有率とか、ある文学作品における漢字含有率とかいった具合にである。これらの場合の漢字含有率は、いったいに、サンプリングした資料をもとに母集団の値を平均値（または平均値の推定区間）によって示すことが多い。これはこれで一つの指標であることに違いはない。しかし、読者が実際に目にするのは一連の文章である。そのときに読者が感じとる文章表記の様相を示すものとして考えると、平均値はどれほどの意味を持つのだろうか。文章は常に平均した漢字含有率で書かれるとは限らない。文章の種類や表現の内容によって漢字含有率はさまざまな値をとると予測される。雑誌や週刊誌などは多種の文章の混合体である。それに対して、平均値で示される漢字含有率は、いわば試験の平均点である。各科目の得点はわからない。受験生の特性を見るには各科目の得点を知ることでもある。同様に文章表記のあり方を調べるにあたって、個々の文章（あるいは一連の文章のブロック）の表記を取り上げることも有意義だと考える。

そこで、文章の種類や表現の内容によって、漢字含有率がどれほど違っているのか、また、さまざまな値をとるとすれば、そこになんらかの法則性、あるいは傾向と呼べるものが存在するのではないか、こうした疑問と推測に対して現実のいくつかのデータ、それもなんらかのまとまった意味を有する、一連の文章のかたまりとしてのデータを用いて、漢字含有率を含む字種

(漢字の異なり種ではなく、文字体系の種類をさす)の比率の種々相を観察し、それに考察を加えるのが、本稿の目的である。

2. 漢字含有率と各字種比率

最初に前提として、漢字含有率の性質についていくつかの確認をしておく。第一は漢字含有率に影響を与えうる要因の問題である。本来、漢字含有率というものは一つの結果である。表記される作品があって初めて存在する。表記される以前から、漢字含有率は何%であるべきだと決められているわけではない。そこで、漢字含有率に対して影響を及ぼすものとして、少なくとも表記される作品、およびその作品を表記する人間のあり方の二つが考えられる。

作品のあり方が与える影響というのは、一般的には次のようなことである。漢字で書くか、仮名で書くかは、たとえば概念的な語は漢字でとか、外来語はカタカナでとかいうように、単語によってある程度決まっている。そして、どういう単語が出現するかは、作品の種類、内容によって決まってくる。つまり、作品のあり方が表現・単語を通じて漢字含有率に影響を与えるというわけである。もちろん、子供向けの童話・絵本などでは、すべて仮名で書くことがあるが、この場合は作品のあり方そのものが漢字含有率を左右していると言える。しかし、このような例は特殊な場合と考えられる。

次に、作品を表記する人間の方であるが、これは言うまでもなく、漢字か仮名かを直接に決定する権利を持つものである。使われる字種が単語によってある程度決まっているとは言え、最終的な決定を下すのはこの表記主体である。漢字が多すぎると読みづらい、あるいは、仮名が続きすぎると読みにくいと考えると、漢字と仮名の割合を加減することもありうる。また表記主体の書きぐせのようなものもある。したがって、作品の表現法や単語のあり方だけで漢字含有率が決まるとは言えず、表記主体のあり方も当然かなりの影響を与えるものと考えられる。

漢字含有率に関する第二の問題は、漢字含有率の意味づけについてであ

る。漢字含有率は確かに文章表記の様相を知るのに便利な指標ではある。しかし、指標として漢字含有率だけでよいかどうかとなると疑問が生じる。漢字含有率は、漢字の量と漢字以外の文字または表記記号の量との相対的な比率によって定まる値である。それゆえ漢字以外の文字または表記記号の量についての考慮がなくてはならないはずである。特に最近の表記では、カタカナ表記の比率の増加が予測される。カタカナの比率がもはや無視できない値をとる可能性がある。2作品において、漢字含有率が同じであっても、カタカナ比率にかなりの差が存する場合も考えられる。このような2作品の文章表記の姿は、当然違ったものになっていると思われる。したがって、文章表記の様相をとらえるには単に漢字含有率だけを問題にするのではなく、漢字以外の字種の比率との関連についても調べる必要があるとしよう。

以上のことから、漢字含有率を問題にする場合、分析の観点としては、少なくとも作品および表記主体のあり方を取り上げる必要があるし、測定すべきものとしては漢字含有率以外の字種比率をも取り上げねばならないと考える。

3. 調査対象・調査方法

調査の対象とすべき文章は、その目的からできるだけ種々雑多なものであることが望ましい。しかし、特殊な表記のものを意図的に取り上げるのでは、かたよりが大きくなる。そこで次のような手順をとった。

- (1) 1979年7～8月に発行された雑誌から、比較的よく読まれているものを選出する（全部で63誌）。
- (2) それらの雑誌に含まれている記事をもとに、雑誌記事の種類として七つの項目（小説、評論・論文、実用・解説、ルポ・報告、インタビュー・座談会、随筆・エッセー、読者投稿）を立てる。
- (3) 7項目にあてはまる文章を各雑誌からすべて抽出する（第1次抽出）。
- (4) 第1次抽出分から、次の条件を満たすように取捨選択を行う（第2次抽出）。

①抽出する文章は、雑誌の見開き2ページ内に納まる程度の長さで、なんらかのまとまりを有する分量であること。

②各項目5万字前後で、全体で35~36万字になることをめやすとする。

上記の(4)の①は、文章としてある程度の意味的なまとまりを有するものを単位とするということである。したがって、短い文章、たとえば読者の投稿やコラムの随筆などは文章全体を抽出することになるし、長い文章、たとえば小説や論文などは、章や節、あるいは数段落が抽出されることになる。このような抽出法は、語彙調査などでは、結果にかたよりが生じる危険度が高くて許されないだろう。しかし、用字調査でしかも字種比率の種々相を調べるといふ目的には、さほど問題がないと思われる。字種比率は、雑誌の見開き2ページ(字数にして1,000~2,000字)分を取り出せば、極端にかたよった値にならないと思われるからである。ただし、第2次抽出における取捨選択が抽出者の任意による点に問題が残るのは否めない。もちろん、抽出者は、結果を故意にゆがめるような意識が入らないように努めたが*。

なお、取り上げた雑誌は最初は63誌であったが、実際に文章を抽出することになったのは55誌であった。これらの雑誌は六つのジャンル(総合誌、小説誌、女性誌、男性誌、実用科学誌、情報誌)に分類できる。このうち、小説誌はさらに純文学系と大衆文学系に分けることができる。その分類と抽出した作品数との関連を表1に示した。

抽出した調査対象は、漢字仮名交じりで磁気テープにパンチ入力される。その際、段落の始まりのスペースはスペースとして扱い、また、会話や段落のために文章が改行されている部分には、改行を示す情報を付加した。ただしルビは無視した。磁気テープをもとに字種のふり分けをするプログラムを通して集計を行った。その他、各種の比率計算やグラフ出力なども電子計算機で処理した。

* 抽出者は国立国語研究所言語計量研究部第二研究室長・野村雅昭と、同室員である筆者の二人である。抽出にあたっては、その雑誌の性格を反映しているような文章を取り上げるように配慮した。

表1

A		B		B'		C	
小説	評論	實用	座談	隨筆	隨筆	隨筆	隨筆
計	計	計	計	計	計	計	計
1	1	2	1	1	1	1	1
2	1	1	1	1	1	1	1
3	1	3	1	1	1	1	1
4	1	1	1	1	1	1	1
5	3	1	1	1	1	1	1
6	1	2	1	1	1	1	1
7	3	1	1	1	1	1	1
8	3	1	1	1	1	1	1
9	3	1	1	2	1	1	1
10	1	1	1	1	1	1	1
11	1	1	1	1	1	1	1
12	1	1	1	1	1	1	1
13	1	1	1	1	1	1	1
14	1	1	1	1	1	1	1
15	1	1	1	1	1	1	1
16	1	1	1	1	1	1	1
17	1	1	1	1	1	1	1
18	1	1	1	1	1	1	1
19	1	1	1	1	1	1	1
20	1	1	1	1	1	1	1
21	1	1	1	1	1	1	1
22	1	1	1	1	1	1	1
23	1	1	1	1	1	1	1
24	1	1	1	1	1	1	1
25	1	1	1	1	1	1	1
26	1	1	1	1	1	1	1
27	1	1	1	1	1	1	1
28	1	1	1	1	1	1	1
29	1	1	1	1	1	1	1
30	1	1	1	1	1	1	1
31	1	1	1	1	1	1	1
32	1	1	1	1	1	1	1
33	1	1	1	1	1	1	1
34	1	1	1	1	1	1	1
35	1	1	1	1	1	1	1
36	1	1	1	1	1	1	1
37	1	1	1	1	1	1	1
38	1	1	1	1	1	1	1
39	1	1	1	1	1	1	1
40	1	1	1	1	1	1	1
41	1	1	1	1	1	1	1
42	1	1	1	1	1	1	1
43	1	1	1	1	1	1	1
44	1	1	1	1	1	1	1
45	1	1	1	1	1	1	1
46	1	1	1	1	1	1	1
47	1	1	1	1	1	1	1
48	1	1	1	1	1	1	1
49	1	1	1	1	1	1	1
50	1	1	1	1	1	1	1
51	1	1	1	1	1	1	1
52	1	1	1	1	1	1	1
53	1	1	1	1	1	1	1
54	1	1	1	1	1	1	1
55	1	1	1	1	1	1	1
56	1	1	1	1	1	1	1
57	1	1	1	1	1	1	1
58	1	1	1	1	1	1	1
59	1	1	1	1	1	1	1
60	1	1	1	1	1	1	1
61	1	1	1	1	1	1	1
62	1	1	1	1	1	1	1
63	1	1	1	1	1	1	1
64	1	1	1	1	1	1	1
65	1	1	1	1	1	1	1
66	1	1	1	1	1	1	1
67	1	1	1	1	1	1	1
68	1	1	1	1	1	1	1
69	1	1	1	1	1	1	1
70	1	1	1	1	1	1	1
71	1	1	1	1	1	1	1
72	1	1	1	1	1	1	1
73	1	1	1	1	1	1	1
74	1	1	1	1	1	1	1
75	1	1	1	1	1	1	1
76	1	1	1	1	1	1	1
77	1	1	1	1	1	1	1
78	1	1	1	1	1	1	1
79	1	1	1	1	1	1	1
80	1	1	1	1	1	1	1
81	1	1	1	1	1	1	1
82	1	1	1	1	1	1	1
83	1	1	1	1	1	1	1
84	1	1	1	1	1	1	1
85	1	1	1	1	1	1	1
86	1	1	1	1	1	1	1
87	1	1	1	1	1	1	1
88	1	1	1	1	1	1	1
89	1	1	1	1	1	1	1
90	1	1	1	1	1	1	1
91	1	1	1	1	1	1	1
92	1	1	1	1	1	1	1
93	1	1	1	1	1	1	1
94	1	1	1	1	1	1	1
95	1	1	1	1	1	1	1
96	1	1	1	1	1	1	1
97	1	1	1	1	1	1	1
98	1	1	1	1	1	1	1
99	1	1	1	1	1	1	1
100	1	1	1	1	1	1	1
101	1	1	1	1	1	1	1
102	1	1	1	1	1	1	1
103	1	1	1	1	1	1	1
104	1	1	1	1	1	1	1
105	1	1	1	1	1	1	1
106	1	1	1	1	1	1	1
107	1	1	1	1	1	1	1
108	1	1	1	1	1	1	1
109	1	1	1	1	1	1	1
110	1	1	1	1	1	1	1
111	1	1	1	1	1	1	1
112	1	1	1	1	1	1	1
113	1	1	1	1	1	1	1
114	1	1	1	1	1	1	1
115	1	1	1	1	1	1	1
116	1	1	1	1	1	1	1
117	1	1	1	1	1	1	1
118	1	1	1	1	1	1	1
119	1	1	1	1	1	1	1
120	1	1	1	1	1	1	1
121	1	1	1	1	1	1	1
122	1	1	1	1	1	1	1
123	1	1	1	1	1	1	1
124	1	1	1	1	1	1	1
125	1	1	1	1	1	1	1
126	1	1	1	1	1	1	1
127	1	1	1	1	1	1	1
128	1	1	1	1	1	1	1
129	1	1	1	1	1	1	1
130	1	1	1	1	1	1	1
131	1	1	1	1	1	1	1
132	1	1	1	1	1	1	1
133	1	1	1	1	1	1	1
134	1	1	1	1	1	1	1
135	1	1	1	1	1	1	1
136	1	1	1	1	1	1	1
137	1	1	1	1	1	1	1
138	1	1	1	1	1	1	1
139	1	1	1	1	1	1	1
140	1	1	1	1	1	1	1
141	1	1	1	1	1	1	1
142	1	1	1	1	1	1	1
143	1	1	1	1	1	1	1
144	1	1	1	1	1	1	1
145	1	1	1	1	1	1	1
146	1	1	1	1	1	1	1
147	1	1	1	1	1	1	1
148	1	1	1	1	1	1	1
149	1	1	1	1	1	1	1
150	1	1	1	1	1	1	1
151	1	1	1	1	1	1	1
152	1	1	1	1	1	1	1
153	1	1	1	1	1	1	1
154	1	1	1	1	1	1	1
155	1	1	1	1	1	1	1
156	1	1	1	1	1	1	1
157	1	1	1	1	1	1	1
158	1	1	1	1	1	1	1
159	1	1	1	1	1	1	1
160	1	1	1	1	1	1	1
161	1	1	1	1	1	1	1
162	1	1	1	1	1	1	1
163	1	1	1	1	1	1	1
164	1	1	1	1	1	1	1
165	1	1	1	1	1	1	1
166	1	1	1	1	1	1	1
167	1	1	1	1	1	1	1
168	1	1	1	1	1	1	1
169	1	1	1	1	1	1	1
170	1	1	1	1	1	1	1
171	1	1	1	1	1	1	1
172	1	1	1	1	1	1	1
173	1	1	1	1	1	1	1
174	1	1	1	1	1	1	1
175	1	1	1	1	1	1	1
176	1	1	1	1	1	1	1
177	1	1	1	1	1	1	1
178	1	1	1	1	1	1	1
179	1	1	1	1	1	1	1
180	1	1	1	1	1	1	1
181	1	1	1	1	1	1	1
182	1	1	1	1	1	1	1
183	1	1	1	1	1	1	1
184	1	1	1	1	1	1	1
185	1	1	1	1	1	1	1
186	1	1	1	1	1	1	1
187	1	1	1	1	1	1	1
188	1	1	1	1	1	1	1
189							

4. 結果

4-1 各字種の比率

まず、使われる字種のうちのどれが、作品（抽出した個々の調査対象）による差が大きいのかを調べる。表2は、各字種の平均値、分散、最大値、最小値の値を示している。なお、句点・読点・スペース（段落の改行部における一字下げの空白）を記号から独立させて計算している。

表2 各字種比率の平均値・分散・最大値・最小値

	漢字	ひらがな	カタカナ	英文字	洋数字	記号	句点	読点	スペース
平均値	26.52	57.29	6.95	0.22	0.29	1.49	2.27	3.96	1.00
分散	45.78	65.79	50.95	0.29	0.93	1.45	0.41	1.10	0.55
最大値	46.58	73.76	30.74	3.15	6.15	6.16	4.20	7.15	3.85
最小値	10.47	33.42	0.00	0.00	0.00	0.00	0.84	1.38	0.00

ある字種の比率が作品によってさまざまに異なる場合、その字種比率の分散値は大きくなる。表2から、ひらがな、カタカナ、漢字の3種が他に比べて分散の値がはるかに大きいことがわかる。他の字種はどれも平均値、分散ともに小さい。すなわち、それらの字種比率はどの作品においても、あまり大きな値をとることがないし、作品による差も大きくないということである。このことは、最大値・最小値の欄を見ても明らかである。つまり、これらの字種は無視してもよい、あるいは少なくとも漢字含有率に大きな影響を与えることはないと考えられるのである。したがって、以下においては、ひらがな・カタカナ・漢字の3種の比率を中心にながめていくことにする。

4-2 雑誌記事別の字種比率

字種比率を文章記事の種類という観点から分析してみよう。雑誌の記事を分類して七つの項目別に集計したのが表3である。

表3の漢字、ひらがな、カタカナの3種の値を示したものが図1である。横軸に漢字含有率をとり、縦軸にひらがな、カタカナそれぞれの含有率をとっている。大きい×印（図の上方に位置しているもの）がひらがな、小さい

表3 雑誌記事別の字種比率

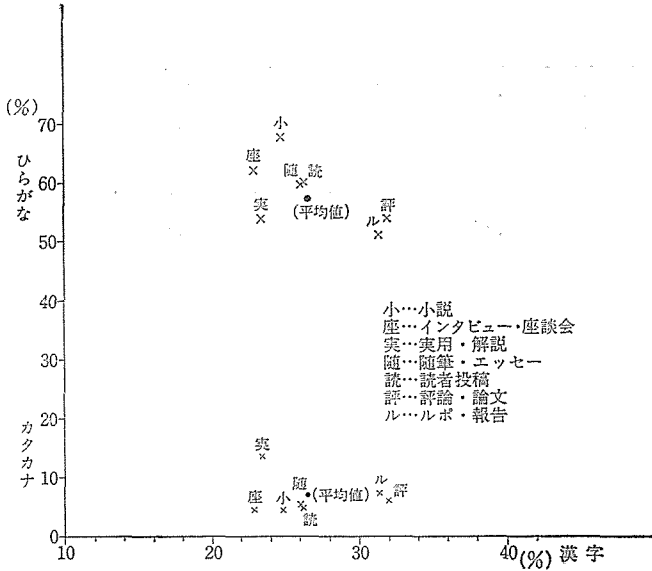
	漢	字	ひらがな	カタカナ	英文字	洋数字	記号	句点	読点	スペース
小 説	24.78	60.72	4.44	0.05	0.01	1.88	2.44	4.56	1.12	
評論・論文	31.99	53.80	6.13	0.30	0.30	1.55	1.63	3.86	0.43	
実用・解説	23.39	53.91	13.75	0.31	0.73	1.27	2.11	3.78	0.73	
ルポ・報告	31.41	51.27	7.34	0.38	0.57	1.99	2.45	3.64	0.94	
インタビュー・座談会	22.89	62.10	4.69	0.15	0.11	1.54	2.64	3.54	2.35	
随筆 エッセー	26.05	59.59	5.63	0.07	0.00	1.10	2.38	4.47	0.70	
読者投稿	26.24	60.27	5.03	0.27	0.18	1.09	2.33	3.78	0.80	
全 体	26.52	57.29	6.95	0.22	0.29	1.49	2.27	3.96	1.00	

×印 (図の下方に位置しているもの) がカタカナを示す。

この表示法 (図2 以下も同様の表示法) の意味は次の通りである。もし、漢字とひらがなの字種比率の和が一定であると仮定する。つまり、漢字・ひらがな以外の字種比率が常に一定の値をとり、漢字比率に影響を与えないと仮定すれば、図の上方の×印は右下がりの一直線上に並ぶことになる。ところが実際には直線上に並ばない。それは、カタカナなどの漢字・ひらがな以外の字種の比率が影響を与えているからである。そのうちのカタカナの影響力は図における下方の小さい×印で確認できるようになっている。つまり、カタカナが多くて、漢字とひらがなの比率が相対的に低くなった場合、その作品は周辺の作品群に比べてひらがな比率の位置が下方になり、その低い分だけカタカナ比率の位置が高くなるはずである。もし、漢字比率の割にひらがな比率の位置が低いにもかかわらず、カタカナ比率も低い場合は、この3種以外の字種比率が高いことを意味する。

図1で、ひらがなの比率は実用・解説を除くとほぼ右下がりに並んでいる。そして、その右下がりの並びから見ると下方に位置する実用・解説は、カタカナ比率の高いことが確認できる。また漢字比率については、評論・論文とルポ・報告が他の項目よりも高い値を示している。随筆・エッセーと読者投稿が非常に近い位置にあることも注目される。

図 1



次に雑誌の記事別に作品ごとの字種比率を図示したのを見てみよう。図2～図7である。図2が小説，図3は評論・論文，図4が実用・解説，図5がルポ・報告，図6がインタビュー・座談会の記事，図7が随筆と読者投稿とを合わせたものである。図の×印やその他の記号の区別は各図に示している通りであり，図中の線分については後述するが，ここでは，記号の区別も線分も無視して見てもらいたい。

ひらがなの比率の位置について見ると，図4・図5以外はどれも似たような右下がりの傾向を読み取ることができる。それに対して図4・図5，特に図4はひらがな比率が散在している感じが強い。カタカナ比率は，逆に図4・図5において，それほど強くはないが右下がりの傾向がうかがわれ，図4・図5以外は横軸に近いところに集まっている。なお，ここで“散在している”というのは，字種の比率がその作品によってさまざまであることを意味し，逆に“集中している”というときは，字種の比率がいずれの作品も似たものであることを意味している。

図 2

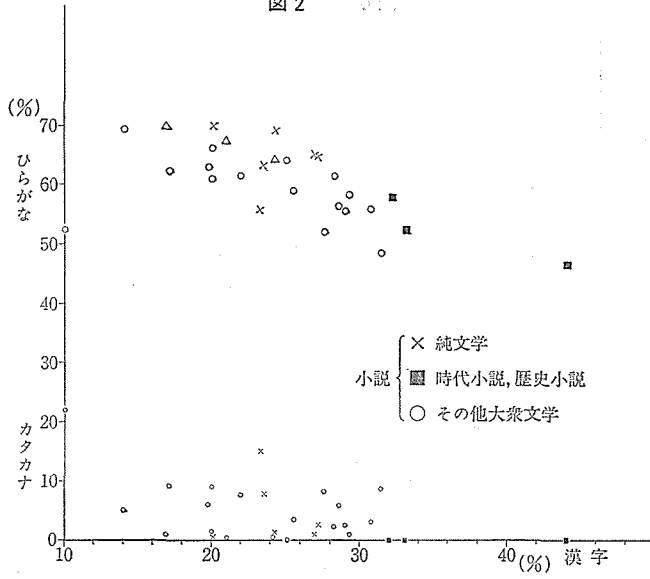


図 3

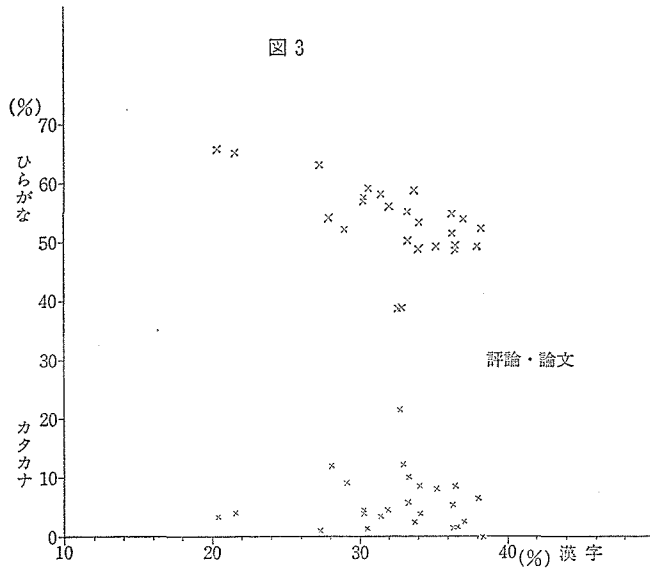


図 4

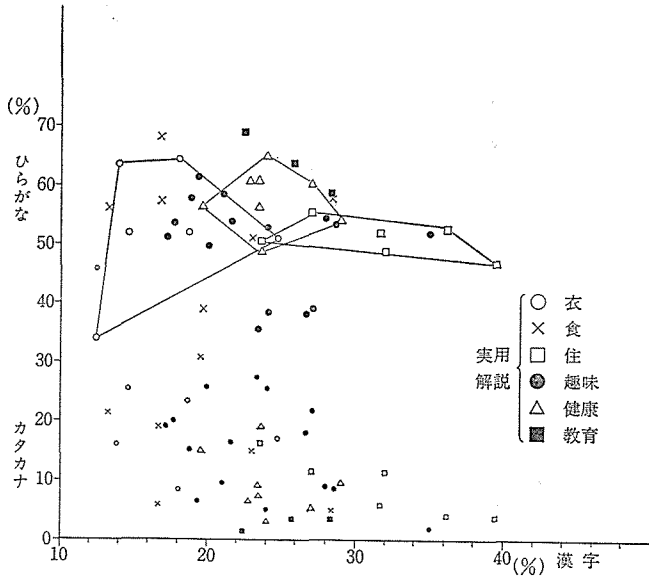


図 5

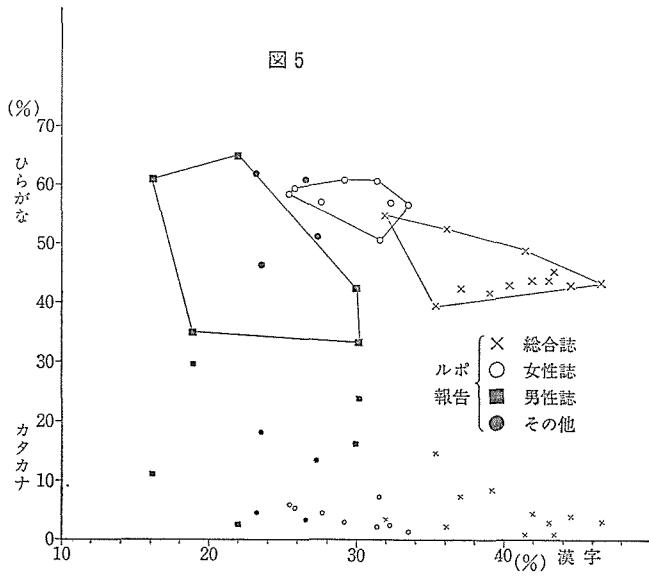


図 6

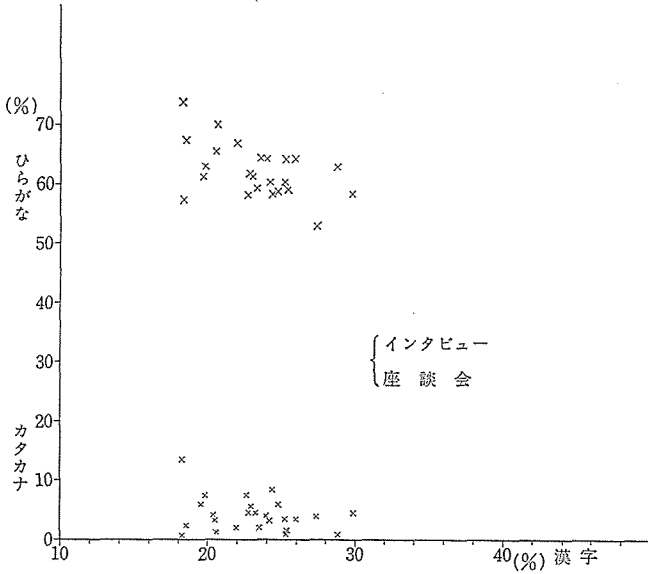


図 7

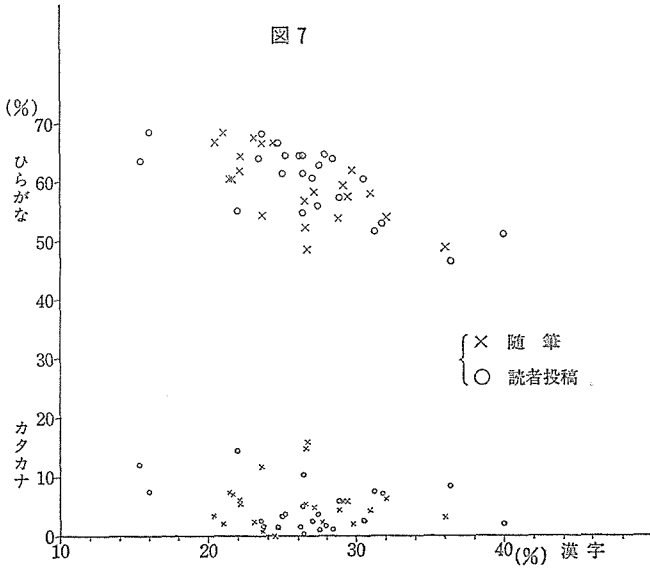


図4・図5以外の各図からは次のことに注目される。

- (1) 小説は漢字比率の最大から最小までその幅が広い。(→図2)
- (2) 評論・論文は漢字比率の高い方にかなり集中している。(→図3)
- (3) インタビュー・座談会記事の位置は、他に比べて集中している度合いが高い。(→図6)
- (4) 随筆と読者投稿とは、ほとんど重なりあっている。ただし随筆の方が集中度が高い。(→図7)

以上のことを総合すると、「字種比率の分布パターンは雑誌記事の種類によって差があり、さらにその種類の中でも、それに含まれる作品ごとに字種の比率がさまざまな値をとるものと比較的似かよった値をとるものがある」と言えよう。

4-3 作品別の字種比率

雑誌記事の種類により分布パターンに差があり、また作品によって差の多少があるとして、その差は作品の内容とどのようにかかわっているのだろうか。

図2の小説を純文学と大衆文学とに分けてみると、純文学の方が集中度が高い。大衆文学で漢字比率の高い作品が出現するのは、歴史小説や時代小説が含まれているからである。図中の△印も小説の舞台となる時代が江戸以前なのであるが、作者が都筑道夫と田辺聖子で、文中の会話部分などかなり現代風に書かれている。それに対して●印の方は、池波正太郎、村上元三といった典型的な時代小説であり、漢字の多いことが十分推測されよう。逆に漢字比率の低いものには現代風俗を扱ったものが多いようである。たとえば、泉大八、川上宗薫、富島健夫のものなどである。

図3の評論・論文は作品ごとの差が小さい。つまり、どれも似かよっているのである。作品の内容としては、政治・経済・文学・哲学・科学など種々の分野にわたっているが、特に分野と関連した差も見出せない。これは、評論・論文には、それらしい文体があるように、表記にもそれらしいスタイルがあり、それが全体に似かよりの結果をもたらしているのではないだろう

か。

図4の実用・解説の場合は、かなりバラエティに富んでいる。図では、作品の内容に応じて、衣・食・住・趣味・健康・教育と六つに分けて、記号を変えて示している。そして、衣・住・趣味・健康については、それぞれに属する作品が含まれるように線分で囲んで示した。これを見ると、内容によって作品の分布にかたよりが見られるものがある。特に住に関するもの、健康に関するものはその傾向が強い。また衣に関するもの場合は、漢字比率が低く、かつカタカナの比率が高いものがある。たとえば、図の最左端に位置する作品は、ひらがなよりもカタカナ比率の方が高い（縦軸50%近くの○印はカタカナ比率である）。これはファッション用語に外来語の多いことを考えあわせれば納得されよう。他方、趣味に関するものが比較的散らばっているのは、その作品に、囲碁・ゴルフ・ドライブ・旅・釣りなどさまざまな内容をもつものを含んでいるためと思われる。

図5はルポや報告の文章を示したものである。ここでは内容の違いを、総合誌・女性誌・男性誌という雑誌の種類の違いで示した。総合誌には、政治・経済に関する社会的な事件のルポやドキュメントの記事が含まれている。女性誌の場合は、教育施設の訪問記や俳優の半生の伝記的なルポおよび家庭問題のレポートが中心である。男性誌の場合は、宇宙船や空母についてのレポートや、スポーツ選手の取材記事などである。このように総合誌・女性誌・男性誌のルポ・報告の記事にはそれぞれ雑誌の種類ごとに内容上の差異が認められる。そこで、図5は雑誌の種類別に記号を変え、その記号を含む線分を示したのである。図の雑誌の種類別の字種比率の分布には、かなり明白な差異が認められる。総合誌では漢字比率が高く、男性誌は漢字比率が低くてカタカナ比率の高いものがあり、女性誌はその両者の中間に位置する格好になっている。

図6はインタビュー・座談会の記事であるが、これは非常に集中度が高い。この中には、政治・経済に関するもの、文芸作品に関するもの、医学・健康に関するもの、芸能人へのインタビューなどさまざまなものが含まれて

いるが、それらの内容との関連は見出せない。これは、談話を文字化し表記する人間がジャーナリストであり、そのジャーナリストたちの間には、ある程度標準的な表記法が存在するためと推測できる。

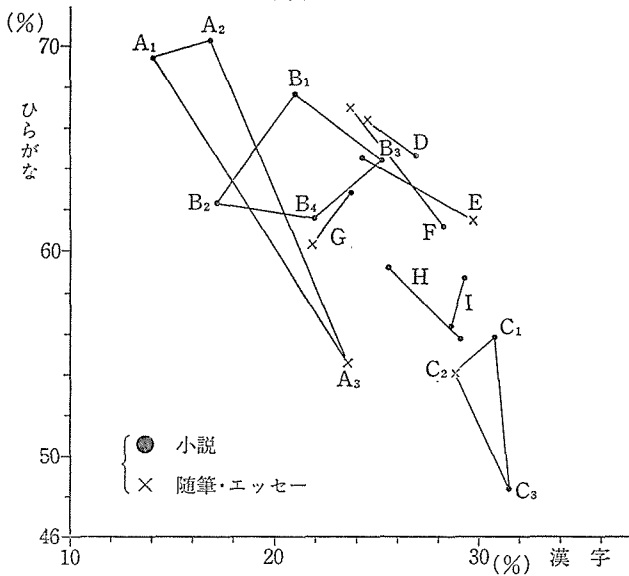
図7の随筆と読者投稿に関しても、内容の違いと字種比率の間には関連性を見出せなかった。もっとも、ここでの随筆と読者投稿とは、ともに筆者の身の周りの日常的なことを述べたものが多く、その意味においては内容の共通性があると言うこともできよう。ただし、随筆と読者投稿とを比べると、後者の方が散らばっている。これは、随筆の筆者の多くが物書きのプロであるのに対して読者投稿の筆者がアマチュアであり、さまざまな表記がなされる可能性が高いことと関連があるかもしれない。

上に述べたことから、実用・解説やルポ・報告の記事のように、同じ種類の記事であっても字種比率の点で種々のものが出現するのは、作品の内容が影響を与えているからだと考えざるを得ない。一方、インタビュー・座談会などの談話記事では、内容以前にジャーナリストの表記法というような制約が働くと考えられるし、評論・論文の場合も同様に評論らしい、あるいは論文らしいという表記上のスタイルにのっとった表記法が行われると考えられる。このように表記を行う際に表記法的な制約が働く場合には、内容による字種比率への影響がそれほど大きくならないのである。

4-4 筆者と字種比率

ここまで、主として作品の面から字種比率を分析してきた。最初に指摘したように、分析の観点としては作品と表記主体の二つが必要である。そこで、ここでは具体的な筆者個人における字種比率の問題を取り上げる。つまり同一の筆者の場合、作品によって字種比率がどれほど異なるのかという問題である。ところで、調査方法でも述べたように、作品の抽出では筆者のことを意識していない。そのため同一の筆者が複数の作品に出てくるのはあまり多くない。したがって、以下に述べるのは、偶然に取り出された人物の、作品における字種比率のあり方であり、結果の一般化には問題があるだろうが、今後の研究への手がかりを示す意味で触れておく。

図 8



上述の調査方法の事情から、ここで対象となるのは、多くの作品を書いている作家、いわゆる人気作家が多い。2作品以上に筆者として登場したものを図8に示す。図8は図7までと表示法は同様であるが、図を見やすくするために縦軸の目盛りを変え、カタカナ比率を省いた点が異なっている。線分をつないでいるのは同一作家で、次の通りである。

- A：都筑道夫 B：田辺聖子 C：藤本義一 D：宇野千代
 E：源氏鶏太 F：水上 勉 G：三田誠広 H：井上ひさし
 I：野坂昭如

Aの都筑道夫の場合、A₁はコミカルな探偵小説、A₂は江戸を舞台とする捕物帖で、A₃は玩具に関するエッセーである。Bの田辺聖子の場合、すべて小説であるが、B₁は忍者が登場する時代のもの、B₂は平安の王朝を舞台とするものであり、B₃とB₄は現代のものである。Cの藤本義一の場合は、C₁は女囚を主人公にした小説、C₂はハードボイルド小説で、C₃は自分の趣味について述べた随筆である。D～Gはそれぞれ小説と随筆・エッセーとの

2 作品、HとIの2作品はともに小説である。

以上のことから特に結論めいたことを言うことはできないが、個人内における差がそれほど大きくない例が見られることから、少なくとも字種比率に対する表記主体の影響力は存在すると思われる。もちろん個人差はあろうが、「文体的特徴」の強い作家がいるように、「表記体的特徴」を持つ作家がいる可能性は高いと思われる。

4-5 雑誌別の字種比率

ところで、調査した雑誌63誌（作品を抽出した雑誌は55誌）は、六つのジャンル（総合誌・小説誌・女性誌・男性誌・実用科学誌・情報誌）に分類することができた。この雑誌のジャンルは、雑誌の性格によって分けたものである。雑誌の性格は、その雑誌が、どのような記事や内容を中心にして、どのような読者を対象としているかということで決定される。記事や内容については、すでに述べてきた記事の種類や作品の内容と深い関連がある。たとえば、小説誌は記事として多くの小説を含んでいるし、図5のルポ・報告で見たように、同種の記事であっても、総合誌・女性誌・男性誌といった雑誌のジャンルによって内容に差が見られた。したがって、雑誌のジャンル別の分析結果は、上述の記事別や作品別の結果と重なる部分があるわけであるが、さらに、雑誌の性格に関与するもう一つの条件である読者に関する面の影響も加わってくると考えられる。そこで、その読者の問題も含めて雑誌の性格が表記（字種比率）とどのように関連しているかを見るために、雑誌別の字種比率を取り上げる。

雑誌の種類としては、六つのジャンルのうち、数の少ない情報誌を除き、小説誌を純文学系と大衆文学系に分けた6種である。表1におけるA、B、B'、C、D、Eにあたる。これらを次の二つに分けて雑誌ごとの字種比率を図9、図10に示した。

図9：総合誌・純文学系・実用科学誌

図10：大衆文学系・女性誌・男性誌

図9の方は、評論・芸術・科学記事を中心とする雑誌であり、図10の方は、

図 9

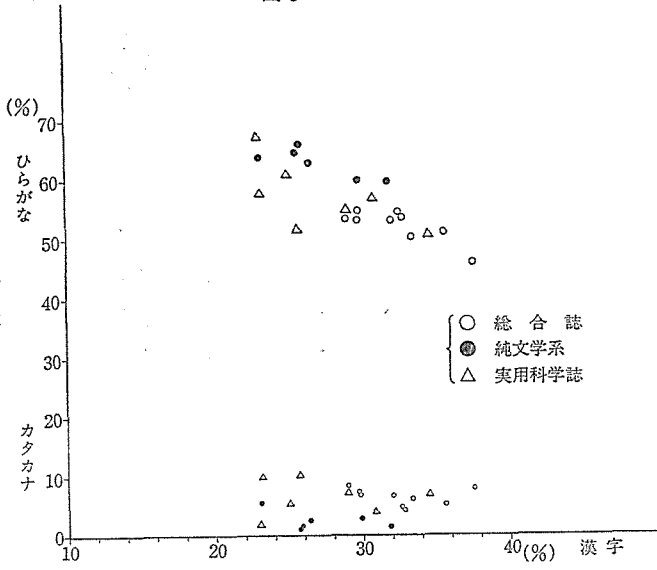
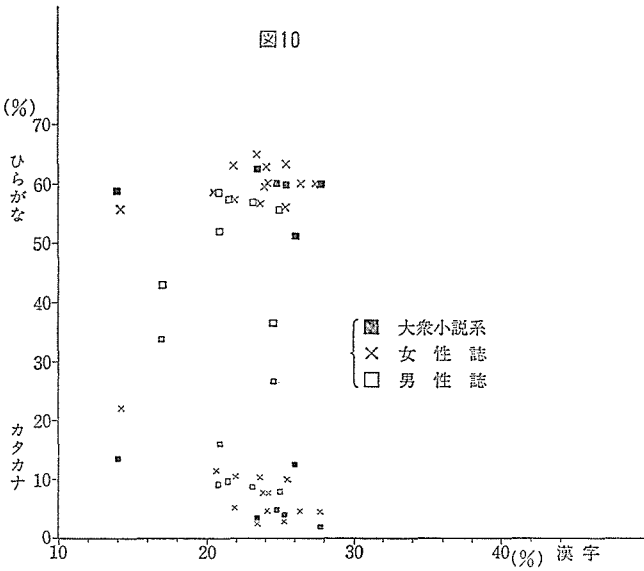


図 10



趣味・娯楽などを中心とする雑誌である。両者を比較した場合、図10の方がより大衆的だと言えよう。なお、図に示した雑誌は、1誌から2作品以上抽出したものに限った。これは、抽出した作品が1作品だけだと、その作品の値がそのまま雑誌の値となるわけで、その作品が特殊な例である場合に生じる結果のゆがみを考慮したためである。

図9では、各雑誌のひらがな比率が右下がりの線を描くかのように並び、カタカナ比率が横軸にほぼ平行して分布しているのに対し、図10の方は、ひらがな比率は水平に近く並び、カタカナ比率は右下がりの傾向が見られる。つまり図9の方は雑誌の間でカタカナ比率がほぼ一定で、漢字が減少するとひらがなが増加するという関係が見られるのに対して、図10の方は、漢字が減少するとひらがなが増えずにカタカナが増えるという現象が認められるのである。

また雑誌の種類に着目すると、図9では総合誌が全般に漢字比率が高く、純文学系や実用科学誌は漢字比率が低くひらがな比率が高い傾向にあることが見てとれる。他方図10の方は、男性誌にカタカナの多いものがあるが、これらは若者世代を読者とする、現代の風俗流行を扱っている雑誌である。

先にも述べたように、雑誌の種類は記事の種類や作品内容と無関係ではあり得ない。しかし、図9と図10における分布パターンの明らかな差異や、それぞれの雑誌の種類による分布のかたよりを見ると、雑誌の性格というものも字種比率に与えている影響は少なくないと考えられる。

5. 考察と仮説

これまで見てきたように、漢字含有率を含む字種比率は、雑誌記事の種類や作品の内容によって種々の値をとってきた。また雑誌の種類によっても字種比率はさまざまな値をとった。しかし、字種比率はただ単にさまざまな値をとるというのではなく、記事の種類によって、ある種の傾向を見せてもいた。

ただし、ここで述べた材料は、必ずしも客観的なデータとは言いがたいも

のも入っており、その点、法則性と言うには無理があるかもしれない。そこで、全体を通した考察のまとめを、文章表記における字種比率に関する仮説として述べることにする。

〔仮説〕

- a. 漢字含有率の比較的高い文章、たとえば評論や論文などでは、漢字とひらがなの間に補完性がある。
- b. 漢字含有率の比較的低い文章、たとえば実用記事や解説文などでは、漢字とカタカナとの間に補完性がある。
- c. 文章において、漢字が減少して仮名が増加する現象が生じる場合、まずひらがなが増加するが、ある程度の限界があり、その限界を越えるとカタカナが増加する。

上のaとbの考えは、図2～7および図9・10をもとにしている。文章の種類によって漢字含有率に差があり、その高い方の文章と低い方の文章では、仮名の使われ方に違いがあると見たのである。ここで漢字含有率が高い（または低い）文章というのは、結果として高い（低い）というより、文章の種類、性格として漢字を比較的多く（少なく）使って書かれやすい文章という意味である。また補完性というのは一方の字種比率が多くなると他方が少なくなるという関係をいう。

cはaとbの考えを拡張したものである。つまり、漢字含有率を基準にして仮名の比率との関係を見たときの仮説である。漢字含有率が減少していく、換言すれば、漢字表記されていた語が仮名書きの語にかわっていく場合、どのような過程が考えられるかについて、巨視的な立場で見たものである。漢字表記語が減少して仮名表記語が増える初期の段階では、ひらがな書き化される。その傾向が進むと次第にひらがなが多くなり読みにくい文章となる。その手当てとしてカタカナ書きが増加して3種の文字によるバランスをとろうとする。これがcの仮説の考え方である。ただし、これは模式的に説明しただけで、実際に一つの文章に対して漢字表記を仮名書き化する場合のことを述べたものではないし、漢字含有率が将来減少していくことを予言

しているものでもない。字種比率における漢字・ひらがな・カタカナの構成比の法則性における仮説なのである。最後にこの仮説を図式化すると次のようになる。

