

国立国語研究所学術情報リポジトリ

Orthographic patterns in Japanese Kanji-Kana strings

メタデータ	言語: jpn 出版者: 公開日: 2017-03-31 キーワード (Ja): キーワード (En): 作成者: 野村, 雅昭, NOMURA, Masaaki メールアドレス: 所属:
URL	https://doi.org/10.15084/00001011

漢字かなまじり文の文字連続

野 村 雅 昭

1. はじめに

漢字の機能には、いろいろなものが考えられるが、その一つとして、漢字かなまじり文において、漢字が分かち書きに相当する機能を持っていることがあげられる。この機能は、漢字とかな*が、それぞれ、どのような語を表現するのに用いられるかということ、および、日本語のシンタックス上の特徴とに關係を持つ。

いま、ここに、

/amenofuruhiwatenkinawarui/

という音列からなる文があるとする。この音列を、かなのみで表記すると、

㊦ あめのふるひはてんきがわるい。

となる。この文字列を読みにくいと感じるか否かは、人によって異なるであろうが、

㊧ あめの ふる ひは てんきが わるい。

というように分かち書きされていれば、意味の切れ目は、わかりやすくなる。しかし、われわれが、ふだん読み書きしているのは、このような、かなのみからなる文章ではない。次のように、漢字をまじえて書かれるのが普通である。

㊨ 雨の降る日は天気が悪い。

(㊦と㊨には、実際は、幾種類ものバリエーションが考えられよう。また、㊦と㊨のどちらが読みやすいかということは、ここでは、問題としない。)

㊨で、漢字で書かれている語は、「雨・降る・日・天気・悪い」で、いずれも実質的な意味内容を持ったものであり、それぞれの漢字は、㊦で分かち書き

* 広義には、「かな」は、「ひらがな」と「かたかな」をさすが、以下で特に断わらない場合は、「ひらがな」をさすこととする。

された、それぞれの句の冒頭に位置する。一方、ひらがなで書かれたのは、「の・は・が」の助詞と「降る・悪い」の活用語尾であって、おのおのの句の終わりにあらわれる。漢字かなまじり文が読みやすいか意味がとらえやすいとかいわれるのは、ふつう、このように、漢字が実質的な内容を持つ部分に用いられ、かなが形式的部分に使われていることによる。すなわち、書き手が述べようとしている事柄の中核的部分を漢字が表わし、それらの関係を明確にする部分に、かなが用いられていることをさしていると考えられる。また、日本語の文の構造を、いわゆる文節というような概念で考える場合、上に述べた実質的な内容を持った部分は、自立語に相当し、形式的な部分とは、付属語に相当する*。したがって、もし、自立語がつねに漢字で表記されるならば、普通の漢字かなまじり文においては、文節は、漢字のみ、あるいは、漢字プラスひらがなで表記されることになる。また、体言や用言は、付属語や活用語尾を伴うことが多いところから、それらの文節は、漢字で始まり、ひらがなで終わるという形式をとることが予想される。

ところで、戦後に告示された、「当用漢字表」や「当用漢字音訓表」は、上述のような文章表記の性格に、大きな変革をもたらした。もちろん、戦前の表記において、上に述べたような、漢字かなまじり文の原則ともいうべきものが完全に守られていたわけではない。自立語に漢字が当てられる傾向が、現在よりも強かったことは予想されるが、必ずしも、すべての自立語が漢字で書かれていたとは限らない。また、付属語でも、漢字で書かれる場合は、かなりあり漢字かなまじり文の原則が一貫していなかった面も多いと思われる。戦前の文章表記の特徴を一言でいうならば、漢字で書ける部分はなるべく漢字で、というのが、それに当たるかもしれない。むしろ、実質的な部分は漢字で、形式的な部分はかなで、という原則は、戦後になってから明確にされた理念とさえみ

* 実質的な部分と自立語、形式的な部分と付属語が、完全に一致するものでないことは、いうまでもない。自立語の中でも、接続詞、連体詞、副詞などは、名詞や動詞にくらべれば、具体的な概念との結びつきは、あまり強いとは言えない。石綿敏雄氏は、漢字で表記されやすい語を、content word、かなで表記されやすい語を、structure word というように呼んでいる。(末尾参考文献①)

ることができる。

「当用漢字表」や「当用漢字音訓表」に、この理念が生きているかどうかはともかくとして、実質的部分と形式的部分という分類そのものが、明確な区分とはなりえない以上、この原則にはずれるものがあらわれるのも、また、当然と言える。「当用漢字表」の「まえがき」には、代名詞・副詞・接続詞・感動詞などは、なるべく、かな書きにするようにとの一項があり、自立語一付属語という概念では、律しえないことを物語っている。また、動詞・形容詞の中でも、「ある・いる・ない」など、補助用言としての用法が頻繁なものは、独立の動詞・形容詞として用いられる場合にも、かな表記される傾向が戦前からあったが、現在では、「当用漢字音訓表」に、「有（あ）る・居（い）る・無（な）い」の訓が認められているにもかかわらず、ほとんど、かなで書かれるということも、その一つの例である。また、実質的部分を、名詞・動詞・形容詞などにかぎってみても、「当用漢字表」や「当用漢字音訓表」に忠実に表記しようとすれば、漢字で書こうとしても書けない語があらわれるのは、当然の結果である。このような傾向の及ぶところ、漢字で書ける語でも、かなで書くという現象がしだいに定着しつつある*。すなわち、戦前と戦後の表記上の特徴の変化は、「漢字で書けるものは漢字で」から「かなで書けるものはかなで」という言い方によって、とらえることができる。さきに、「変革」ということばを用いたのは、この意味である。

ここで述べようとする小調査の報告は、新聞の文章を対象として、上に述べたような漢字かなまじり文の機能が、どのようにいかされているかを明らかにしようとするものである。新聞を選んだ理由は、現在進行している語彙調査・表記調査のデータであり、利用しやすいためであることは、いうまでもないが、戦後における表記の傾向を調べるのに、もっとも適した対象と考えられるからである。周知のように、新聞は、「当用漢字表」や「同音訓表」にもっとも忠実であり、その範囲で文章を表記するために、種々の工夫をこらしてきた。今、語彙調査の対象としているデータは、昭和41年のものであるが、「当

* 現代語で、かな表記される語が増加しつつあり、特に、和語にその傾向が著しいということは、多くの研究者によって指摘されている。(たとえば、末尾文献②)

用漢字表」が告示されてから、二十年近くがたち、新聞の表記にも、一応の安定したスタイルが生まれていると考えることができる。したがって、戦後の表記の変遷によって、漢字かなまじり文の原則がどのような形で存在するかをみるためには、新聞は、恰好の対象であると思われる。かさねて、この調査で目的とするところをまとめてみると、次の二つになる。一つは、漢字かなまじり文で、漢字は、意味の切れ目やまとまりを示す役割を果たしているか、ということである。そして、もう一つは、もし、漢字からその機能が失われつつあるとするならば、それにかわるものは何かということである。この二点を中心に、調査したところを以下に報告する。

2. 調査対象と調査単位

調査の対象としたのは、現在行なわれている「新聞語彙調査」の中間集計までのデータである。この語彙調査は、昭和41年の朝日・毎日・読売三紙の一年分を対象とし、 $\frac{1}{60}$ の面積比で抽出された、14220ブロック（1ブロックは、新聞紙面1ページの約 $\frac{1}{30}$ ）が標本ということになる。今回の調査では、その $\frac{1}{5}$ にあたる中間集計までの4740ブロックから、120ブロックを再抽出して、第二次標本とした。抽出の方法は、文章の種類によって分けた、6層から（各20ブロックずつ）のランダム・サンプリングによった。文章の種類による層別とは、次の6種類である。

- I ニュース
- II ニュース解説・ニュース関連記事・ニュース展望
- III 特別読物・探訪記事・実用記事
- IV 社説・コラム・評論・投書
- V 小説
- VI 広告（「案内広告（三行広告）」を除く）

層別による方法をとったのは、抽出標本数が少ないことから、サンプリングによるかたよりが生じるのを少しでも防ぐためである。したがって、以下の分析では、層による比較は行なっていない。また、VとVIは、今回の分析からは除いたので、直接には、今回の分析では、I～IVの層の80ブロックを対象とし

ている。

漢字かなまじり文中の漢字の機能を見るためには、なんらかの尺度が必要である。ここでは、はじめに述べたような理由によって、文節を単位にして、集計・分析をすることにした。ここでいう文節とは、いわゆる学校文法のそれをさす。ただし、「…ている・…である・…である・…でない・…てみる」などの「いる」「ある」「ない」「みる」の類は、1単位とせず、前の文節に含まれるものとする。また、句読点をはじめとする記号類は、すべて1単位とした。したがって、文中の意味の続きぐあいに関係なく、記号のところで分割されることになる。

以上のような手つづきによって取り出された単位のうち、文字を含むものを、以下、文節とよび、記号のみからなるものは、文節とはよばない。（この場合の「文字」には、漢字・ひらがな・かたかなのほか、アルファベット・洋数字を含む。）

また、文節がどのような文字（表記記号）列から構成されているかをみるために、文字の種類によって符号化して処理をした。以下に、単位分割と符号化の例、ならびに、符号の種類をあげる。

例：首相は、六日、|「|「こん回の事件はイカンである」|と述べた。|
●●●○ ●●● |「○○●●●●●○△△△○○○」|○●○○○。|
●=漢字 ○=ひらがな △=かたかな 、=読点
。=句点 ⊙=段落のはじめの一行下げなどの空間

3. 調査データ中の表記記号の数

今回の調査データ中に、どのような表記記号（文字および狭義の記号）があらわれたかを表1に示す。漢字の含有率という点からみると、従来の調査に比べて、比較的、低い数値を示している。同じ新聞語彙調査のデータについて計算した結果がある（末尾文献⑨）が、それによると、漢字とひらがなを除いた表記記号の比率は、ほとんどかわりがないが、漢字は42.1パーセント、ひらがなは43.8パーセントと、漢字の比率が高くなっている。この差は、一つは、サンプリングの誤差によるものと考えられるが、先に示した層別が、実際の出現

表1 調査対象中に出現した表記記号の数

表記記号の種類	出現数	(百分比)
漢 字	8,357	(36.0)
ひらがな	11,824	(51.0)
かたかな	1,274	(5.5)
読 点	812	(3.5)
句 点	443	(1.9)
アルファベット	44	(0.2)
洋 数 字	21	(0.1)
記 号	421	(1.8)
計	23,196	(100.0)

率と一致していないことによる。すなわち、実際の紙面では、ニュースの全体に占める比率が約31パーセントである（語彙数の比率による）のに対し、ここでは、80ブロック中の20ブロック（25パーセント）しかない。ニュースにおける漢字の比率は高く、この20ブロック分では、41.5パーセントになっている。

次に、文中において、文節より大きな表記単位として、文節群という

ものを考えてみることにする。文節群とは、文中で、句読点とスペースによって、他の文字列と区別され、1個以上の文節からなる文字列のこととする。この場合の文とは、句点またはスペースから句点までの文字列をさす。スペースとは、段落のはじめの一字下げの空間をさすが、行の途中で、句読点を伴わずに改行された場合にも、最後の文字の次に、スペースが存在すると考える。したがって、一つの文において、読点があられない場合は、その文は、一文節群によって構成されることになる。

表2は、文節群の種類と平均字数を示したものである。出現数をもっとも多

表2 文節群の類型と平均字数（字数には、「・などの記号も含む）

類 型	字数	出現数	平均 字数	類 型	字数	出現数	平均 字数	類 型	字数	出現数	平均 字数
㊦～、	2,116	148	14.3	、～、	7,387	466	15.9	。～、	2,485	198	12.6
㊦～。	1,184	40	29.6	、～。	7,200	342	21.1	。～。	1,417	61	23.2
㊦～㊦	21	6	3.5	、～㊦	60	4	15.0	。～㊦	71	3	23.7

いのが、文中にあらわれる「、～、」で、次に多いのが、文末にくる「、～。」である。文頭にくる文節群の「㊦～、」や「。～、」は、文中や文末にくる文節群にくらべて、比較的、短い文字数からなっている。また、一文節群で一文が構成される「㊦～。」「。～。」を平均すると、文字数は、25.0となる。

一文の長さの平均を、かりに、総文字数を句点数で割ったものとする、表1により、52.4字となるから、文の長さ、文節群の数との間には、なんらかの相関がありそうである。また、文節群と漢字の数との間にもなんらかの関係があると思われるが、ここでは、これ以上の分析はしなかった。

4. 文節を構成する文字列の類型

次に、文節を構成する文字列が、どのような組み合わせによってできあがっているかを調べてみよう。調査対象中に出現した文節の総数は、5,310である。そのうち、漢字とひらがな以外の文字を含む文節は、384例あり、その大部分は、かたかなを含むものである。問題のありかをはっきりさせるために、以下では、漢字とひらがな以外の文字を含む文節を除いて、漢字のみ、ひらがなのみ、および、漢字とひらがなからなる文節を対象として分析することにする。

文字列を類型化するために、ここでは、次のような方法をとった。対象とする文節は、漢字とひらがなのみからなっている。したがって、これらの文節の文字列は、二種類の符号によって表わすことができる。これを文字数に関係なく、漢字が連続する部分、ひらがなが連続する部分というようにまとめてみると、次の八種類に分けることができる。

- ■ ■ ■ ■ 昨夜 (●●●) ・ 結局 (●●●)
- ■ ■ □ □ 作った (●○○) ・ 安保理事会で (●●●●●○)
- □ ■ □ □ 打ち上げた (●○○○○) ・ 取り扱い方を (●○●○○○)
- □ □ ■ ■ 忘れ物 (●○○) ・ 阿佐が谷 (●●○○)
- □ ■ ■ ■ その上 (○○●) ・ この頃 (○○●)
- ■ ■ □ □ お弁当を (○●●○) ・ くり拾いに (○○●○○)
- ■ □ ■ ■ 該当例なし
- □ □ □ □ しかし (○○○) ・ きびしかった (○○○○○○)

表3は、漢字とひらがなのみからなる文節を、上の8種(実際には、7種)の類型に分類したものである。文節の始まりが漢字であるものの合計が69.5パーセントと約7:3の割合で、ひらがなで始まる文節をしのいでいる。また、文節の終わりが、ひらがなであるものの割合は、92.1パーセントとなってい

表3 漢字とひらがなのみからなる
文節の類型と出現数

類	型	出現数	(百分比)
漢字で始まる文節	■ ■ ■ ■	379	(7.7)
	■ ■ □ □	2,978	(60.4)
	■ □ ■ □	62	(1.3)
	■ □ □ ■	3	(0.1)
	小計	3,422	(69.5)
ひらがなで始まる文節	□ □ ■ ■	7	(0.1)
	□ ■ ■ □	82	(1.7)
	□ ■ □ ■	—	—
	□ □ □ □	1,415	(28.7)
	小計	1,504	(30.5)
計	4,926	(100.0)	

る。したがって、一般に文節は漢字で始まりかなで終わるという原則は、かなり高い割合で守られていることになる。こうした全体の傾向を裏書きするように、個々の類型では、漢字で始まり、途中でかなにかわって、そのまま終わる ■ ■ ■ □ のタイプがもっとも多く、約60パーセントを占めている。ついで、ひらがなのみからなる □ □ □ □ が28.7パーセントを占めているのが注目される。そのほかは、すべて漢字からなる ■ ■ ■ ■ がやや多いほかは、複合動詞などにあらわれる ■ □ ■ □、まぜ書きの前部分がかなの場合にあらわれやすい □ ■ ■ □ が1パー

セント台で、■ □ □ ■ や □ □ ■ ■ などは、きわめて低い割合しか占めていない。このようにみると、文節を構成する文字列は、■ ■ ■ □、■ ■ ■ ■、□ □ □ □ という単純なパターンで、95パーセント以上が占められており、そのほかの複雑なパターンは、あらわれにくい、あるいは、それを避けようとする傾向があるということがわかる。以下に、出現数の少ないものの例を示す。

- ①二人は戦地の思い出にひたったが、数日後彦治は望月に裏切られてしまう。 (■ ■ ■ ■ の例)
- ②一般のおとなも「血液型登録証」を持ち歩くことを励行すればよいと思う。(■ □ ■ □ の例)
- ③日本刀を持ってあばれている男を、けん銃で威圧して逮捕したことがありました。(□ ■ ■ □ の例)
- ④身辺の整理、忘れ物、落し物のないように、小さくとも、礼儀をわきまえることなどに重点をおいて、(■ □ □ ■ の例)
- ⑤医者が登録者のところへ行く費用も自弁で、その上、手ぶらではいけないからお香典を持っていく。(□ □ ■ ■ の例)

このように、文節は、漢字で始まりかなで終わるといった一般的なタイプを持つ。しかしながら、ひらがなで始まる文節も、約30パーセントを占めており、文節の終わりにひらがながきやすい以上、ひらがなが連続する場合も、また、同様に、漢字が連続する場合もありうる。そのような場合に、読みにくさを避けるために、表記上、なんらかの工夫が施されていることも、予想できよう。そこで、文節の類型ごとに、その直前、直後に、どのような表記記号があらわれるかをみる必要がある。表4-1は、前部分（直前の表記記号）、表4-2は、後部分（直後の記号）によって分類した結果である。

まず、前部分との関係からみると、先に述べたように、かなで終わる文節は、全体の約92パーセントであるが、実際に文節の前部分にひらがながあらわれる場合は、69.9パーセントであり、句読点・スペース・記号などが、文節の

表4-1 漢字とひらがなのみからなる文節の文字環境（前部分）

類 型		●	○	△	、	。	㊦	その他	計
漢字で始まる文節	■ ■ ■ ■ ■	23	187	—	67	15	14	73	379
	■ ■ ■ □ □	47	2,218	3	434	105	88	83	2,978
	■ □ ■ □ □	1	47	—	7	3	4	—	62
	■ □ □ ■	—	—	—	2	—	—	1	3
	小 計	71 (2.1)	2,452 (71.6)	3 (0.1)	510 (14.9)	123 (3.6)	106 (3.1)	157 (4.6)	3,422 (100.0)
ひらがなで始まる文節	□ □ ■ ■ ■	—	—	—	1	2	1	3	7
	□ ■ ■ ■ □	—	57	—	13	3	7	2	82
	□ □ □ □ □	19	935	—	170	103	40	148	1,415
	小 計	19 (1.3)	992 (65.9)	—	184 (12.2)	108 (7.2)	48 (3.2)	153 (10.2)	1,504 (100.0)
計	90 (1.8)	3,444 (69.9)	3 (0.1)	694 (14.1)	231 (4.7)	154 (3.1)	310 (6.3)	4,926 (100.0)	

（前部分の「その他」は、おもに、「」・（ ）などの記号からなる。）

表4-2 漢字とひらがなのみからなる文節の文字環境（後部分）

類 型		●	○	△	、	。	㊦	その他	計
漢字で終わる文節	■■■■■	83	23	5	137	21	2	108	379
	■□□■	—	—	—	2	—	—	1	3
	□□■■■	—	—	—	6	—	—	1	7
	小 計	83 (21.3)	23 (5.9)	5 (1.3)	145 (37.3)	21 (5.4)	2 (0.5)	110 (28.3)	389 (100.0)
ひらがなで終わる文節	■■■□□	1,617	71	48	344	223	2	73	2,978
	■□■■□	26	17	1	12	5	—	1	62
	□■■■■□	40	17	4	13	7	—	1	82
	□□□□□	666	254	31	221	156	7	82	1,415
	小 計	2,349 (51.8)	957 (21.1)	84 (1.9)	590 (13.0)	391 (8.6)	9 (0.2)	157 (3.4)	4,537 (100.0)
計	2,432 (49.4)	980 (19.9)	89 (1.8)	735 (14.9)	412 (8.4)	11 (0.2)	267 (5.4)	4,926 (100.0)	

直前にくる場合もあることを示している。これを、漢字で始まる文節とかなで始まる文節とに分けてみると、前者では、ひらがなに続く割合が71.6パーセントであるのに対し、後者では65.9パーセントと、わずかではあるが、ひらがなの連続するのを避ける傾向がみられる。

漢字で始まる文節のうち、すべて漢字からなる■■■■■は、漢字が前にくる場合が379例中23回（6.1パーセント）と、やや高い比率を示している。たとえば、次のような例である。

⑥北爆停止後九日たっても共産側からまだ“よい返事”がないため、

⑦これを知ったPTAが昨年六月一台買入れ、

ただし、これを後部分との関係からみると、すべて漢字からなる文節は、あとに読点や記号をともなうことが多く、次の例のような形であられやすい。

⑧六七年、ジョンソン米大統領はベルリンを訪問、ポーランドにオーデル

＝ナイセの国境を保障するむね演説。

⑨つまり“持ち家主義”の住宅政策を“借家主義”に切りかえるのである。

ひらがなのみからなる文節で、その他の記号が前にくるものが多いのは、次の章でも述べるが、次のような場合である。

⑩「絶対に彼を藤山支持にさせない」と放言したり、

次に、後部分についてみると、漢字が続く場合がもっとも多く、全体では、49.4パーセントである。ただし、かなで終わる文節では、51.8パーセントであるのに対し、漢字で終わる文節では、21.3パーセントである。すなわち、漢字で終わる文節のうち、約70パーセントは、句読点と記号があとにくるものであり、かなで終わる文節のあとに句読点や記号がくる場合が約28パーセントであるのと、はっきりした対照をみせている。句読点や記号が後にくる場合とは、上にあげた⑧や⑨のような例である。

かなで終わる文節のうち、■■■■□□では、ひらがながあとにくる場合が、2,978例中、671例(22.5パーセント)であるのに対し、□□□□では、1,415例中、254例(18.0パーセント)である。また、句読点とその他の記号が続く割合は、■■■■□□では、644例(21.6パーセント)であるが、□□□□では、459例(32.4パーセント)となっており、ひらがなのみからなる文節は、うしろに、漢字とかな以外の記号類がきやすいことを示している。たとえば、次のような例である。

⑪…との現地電も伝えられている。また、辞表を提出した関係のうちには、

⑫現在は流動的にしか捕え得ないものなのだ。それこそ、最近の未来論の流行の真の理由なのである。

■■■■□□や□□■■■■は、先にあげた④・⑤や、⑨の「持ち家主義」(■■■■□□)のように、ほとんどが特殊な場合で、前部分や後部分は、すべて、記号類の場合である。普通の漢字かなまじり文では、出現しにくいパターンと言えよう。

5. ひらがなで始まる文節の分析

以上にみてきたように、漢字かなまじり文においては、漢字で始まりかなで終わるといふ文節がもっとも一般的なタイプである。そして、前部分や後部分

には、漢字やひらがなのみでなく、句読点などの記号類がくることによって、意味のまとまりや切れ目を示していることが予想される。しかしながら、なおここで注目されるのは、ひらがなで始まる文節が1,504例と全体の約30パーセントを占め、そのうち、前部分にひらがながくる場合が、約66パーセントに達していることである。なかでも、ひらがなのみからなる文節は、1,415例あり、そのうち、前部分にひらがながくる場合が935例(70.6パーセント)、後部分にひらがなが続く場合が254例(18.0パーセント)となっている。いわば、この類は、漢字かなまじり文の原則にはずれたものである。漢字かなまじり文の機能を考えるということは、裏がえせば、その原則にはずれたものの実体を明らかにすることでもある。以下、本論では、ひらがなで始まる文節の分析を中心に論をすすめることにしたい。

表5は、ひらがなで始まる文節を、それぞれの文節を構成する自立語の品詞によって、前部分の表記記号別に分類したものである。いわゆる形容動詞は、すべて、名詞の中に含めた。品詞の「その他」というのは、例⑬のように、文

表5 ひらがなで始まる文節の品詞別分類

前部分 品 詞	●	○	、	。	Ⓢ	その他	計
名 詞	3 (0.7)	296 (72.4)	71 (17.4)	20 (4.9)	12 (2.9)	7 (1.7)	409 (100.0)
動 詞	6 (1.2)	490 (97.4)	3 (0.6)	3 (0.6)	1 (1.2)	—	503 (100.0)
形 容 詞	—	70 (94.5)	1 (1.4)	2 (2.7)	—	1 (1.4)	74 (100.0)
副 詞	2 (1.1)	90 (48.9)	63 (34.2)	20 (10.9)	3 (1.6)	6 (3.3)	184 (100.0)
連 体 詞	3 (2.2)	46 (33.8)	42 (30.9)	23 (16.9)	15 (11.0)	7 (5.2)	136 (100.0)
接 続 詞	5 (7.4)	—	4 (5.9)	40 (58.8)	17 (25.0)	2 (2.9)	68 (100.0)
そ の 他	—	—	—	—	—	130 (100.0)	130 (100.0)
計	19 (1.3)	992 (65.9)	184 (12.2)	108 (7.2)	48 (3.2)	153 (10.2)	1,504 (100.0)

節の途中で記号がはいったため、独立して1単位としたもので、すべて、助詞である。引用の「と」がもっとも多いが、次のような場合もある。

⑬椎名外相とラスク長官の間で“極秘の討議”が行なわれたようだが、
接続詞は、ひらがなに続く場合は一回もなく、前の語との切れ目がもっともはっきりしている。ほとんどが⑭・⑮のように、文頭にあらわれることが多いが、⑯・⑰のように、文中にあらわれる場合でも、切れ目は明瞭である。

⑭辞典の売場がにぎわっている。／だが、辞書はちょっと見ただけではその内容はわからないものである。

⑮…防ぐようにしたいものです。なぜなら、これからは日本脳炎の季節だからで、

⑯それを相殺し、あるいは上回る勢いで物価が上昇、

⑰施政権の無条件かつ全面返還の一步としての
連体詞は、⑳のように、かなに続く場合が33.8パーセントあるが、文頭にあらわれる場合が27.9パーセント、読点のあとが30.9パーセントと、接続詞について、切れ目がはっきりしている。

⑱「この壇上で理事会を開け」などと詰めよった。

⑲国民の半分以上が、生活が苦しくなったと感じている。その苦しくなった理由の大部分が…

⑳反政府派が、さる二十一日の全国大会で巻返しをはかり、

㉑米軍基地の維持をその前提としているので、

副詞になると、ひらがなに続く率がほぼ半数に近づくが、名詞や動詞にくらべれば、句読点のあとにあらわれる割合が高く、40パーセントを越えている。

㉒それはいうまい。とにかく一日もはやく十万円たまることを祈ろう。

㉓黒28とキラれては、とても勝ちみがないし、

㉔兄弟の答えもほとんど声にならず、

㉕ワシントンにはそろそろ悲観論が出始めている。

㉔と㉕は、たまたま直前の文節には、漢字やかたかなが含まれ、後の文節にも漢字が含まれている。このように、かな書きの副詞の直前にひらがながあらわれる場合でも、前後の文節に漢字やかたかなが含まれていたり、あるいは、例㉒

表6 ひらがな表記の副詞の前後にくる文節の文字列構成

後 \ 前	ひらがなのみからなるもの	その他
ひらがなのみからなるもの	8	22
その他	10	50

「その他」には、ひらがな以外の文字を含む文節と句読点などの記号がくる場合が含まれる。

前後またはそのどちらかに、漢字やかたかななどが含まれるか、句読点がかかる。前後とも、ひらがなのみからなる文節があらわれるのは、8例にすぎない。㉗と㉘は、その例である。

㉘ こうして、公選制はまた、派閥政治の根源ともなった。

㉗ 車のかけるしぶきも、ふだんよりずっとはげしいので、

㉘ インドの外交的威信が今月ほど低下したことはかつてない。

以上の接続詞・連体詞・副詞にくらべると、名詞・動詞・形容詞は、直前にかながくる割合が圧倒的に高い。名詞の場合は、読点に続く場合が、17.4パーセントとやや多く、ひらがなに続く場合が72.4パーセントであるが、動詞・形容詞では、ひらがなに続く場合がいずれも90パーセントを越えている。句読点に続く名詞の中では、「これ」・「それ」などの、いわゆる「こそあど」の類が多いのがめだつ。読点に続く71例のうち、30例がそれにあたる。たとえば、㉙・㉚のような例である。いっぽう、ひらがなに続く名詞296例のうち、「こそあど」は、15例にしかすぎない。以下に、名詞・動詞・形容詞が、ひらがな以外の表記記号に続いてあらわれる例をあげる。

㉙ いろいろな投書の中から焦点をしぼり、それを一般的な話題にして聞かせる。

㉚ 左下スミを白30とめしあげられるのも、べらぼうな損害で、

㉛ 団地には保育園が完備している。ここの保育料は半年で四元。

㉜ 北爆停止後九日たつても、

㉝ (通知簿ヲ) なん度もながめては、ほほえんでいた。

のように、読点や記号があれば意味の切れ目は、比較的わかりやすいと思われる。そこで、前後の文節がどのような文字列からなっているかを、ひらがなのみからなるものと、それ以外のものに分けて示したのが表6である。90例中、82例までが、

㊸…などが主な産地。すっぱいが、ビタミンCが豊富だし、

以上みてきたように、接続詞・連体詞・副詞などは、ひらがなで文節が始まっても、句読点などが前後にくることが多く、それによって、読みやすさが失われることは少ない。また、副詞を除けば、これらの語は、ほとんど、ひらがなで表記されることが多いが、それが、漢字かなまじり文の中で、読みにくさを増すことにはなっていないことがわかる。副詞には、漢字表記されるものもあるが、すべて漢字からなる文節 379 例のうち、前部分に漢字がくるものそのうち 23 例であるところから、漢字表記の副詞の場合にも、前の文節との切れ目がわかりにくくなる場合は、そう多くないと考えられる。

一方、名詞・動詞・形容詞は、漢字表記されることが多いと思われるが、ひらがな表記されるものの大部分が、名詞の一部を除けば、ひらがなに続いてあらわれる。それが、文字列中において、どのような環境にあらわれるかについては、次章以下で分析する。

6. 出現確率の高い文字連続

表 6 で示した、ひらがなに続く名詞 296 例の内訳を表 7 に示す。出現頻度の高いものは、ほとんどが、形式名

表 7 「ひらがな」で始まる名詞の内訳

詞か、それに準ずるものであり、普通名詞として使われたものは、あまり多くはない。また、調査対象が新聞であることを考えると、表外字や表外音訓に該当するため、ひらがな表記となったもの

こと	61	とも(に)	5
もの	46	とき	4
ため	17	ほか	4
ところ	12	わけ	4
これ	9	(その他)	129
うち	5	計	296

(数字は出現回数)

(例えば、㊸・㊹)も、いくつかみられるが、それほど多くはない。また、その他の 129 例のうち、13 例は、接頭辞「お」を伴うもので、「お」の次の文字が漢字であるものは、そのうち 10 例であった。

㊺何らの合意もみなかったことは残念である。

㊻軍事基地の維持と引き替えになされるものであれば、反対する。

㊼大会は混乱のうちに中止になったといわれる。

㉔通学する子供たちはほんとうに危険である。

㉕インド提案のあいまいさと、情勢分析の甘さを指摘されて、

㉖汚水のため海水のよごれがひどくなっており、

㉗病気で寝ている人のお見舞に買ってゆけば、

表7で、さらに気がつくことは、総数296のうち、比較的少数の語によって、全体のかなりの割合が占められていることである。出現頻度の高い「こと・もの・ため・ところ・これ」の5語で、全体の49.0パーセントと約半数が占められていることになる。このことは、名詞がかな書きされる場合、その多くは、出現頻度の高いものであり、また、それが形式名詞やそれに準ずるものであるところから、それを含む文字列には、しばしば出現するものが多いのではないかということ予想させる。

もっとも出現頻度の高い「こと」を例にとると、㉘のように、普通の名詞として使われる場合もあるが、その多くは、㉙・㉚のように、形式名詞的なものである。そこで、このように、「…ことになる」「…ことがある」というよう

㉘午後から夕方になると疲れやすく、一つのことに集中できません。

㉙与野党が真向から対立することになろう。

㉚ベトナム政府軍の司令官も勝手に使うことがある。

な文字列が、もしも、ひんぱんに出現するならば、読み手にとって、ゲシタルトが構成されやすく、かなが連続しても、読みにくさを感じるものが少ないのではないだろうかと思われる。それを立証するためには、ここで対象としたデータでは、量的に不足なので、言語計量調査室で作成された、「朝日新聞朝刊(1月～6月)文字列表」* (以下略称「文字列表」)を援用することにする。

表8と表9は、「文字列表」に出現した「こと」という文字列の前後の表記記号、および、「こと」を含む4文字列のうち、出現頻度の高いものをあげたものである。この場合の「こと」は、「事」で表わされる名詞とはかぎらず、「こと(琴)」「ことば」「いかとたことを食べる」などのように、ひらがな表記さ

* この「文字列表」は、今回の語彙調査データのうち、朝日新聞朝刊(1月～6月)に出現した、本文に相当する部分、約20万字について、文字列の出現確率を集計したものである。(末尾参考文献④)

表8 文字列「こと」の前後に出現する
確立の高い表記記号

前 部 分		後 部 分	
る	294	に	135
た	87	を	111
う	61	が	93
の	56	は	83
い	43	で	57
な	29	も	48
(その他26種)	109	(その他21種)	152
計 (32種)	679	計 (27種)	679

(数字は出現回数)

表9 文字列「こと」を含む
出現頻度の高い4文字列

ることに	68
ることを	56
ることが	46
ることは	44
ることも	26
うことに	17
のことで	15
たことが	13
たことは	13
うことを	12
たことを	12
ることで	12
(その他147種)	345
計 (159種)	679

(数字は出現回数)

れた「こと」という文字列をすべて含んでいる。ひらがなは、すべて、異なる文字ごとに1種類と数え、句読点も、それぞれ、1種類とした。また、漢字・かたかな・アルファベット・洋数字・句読点以外の記号は、異なる種類のものがあっても、それぞれのグループを1種類と数えた。

表8によれば、「こと」の前にくる文字のうち、「る」がもっとも多く、全体の約43パーセントを占める。また、上位の「る・た・う・の・い・な」の6字で、全体の約85パーセントにあたる。また、後部分では、上位の「に・を・が・は・で・も」の6字で、全体の約78パーセントを占めている。そして、「こと」を含む4文字列では、これらの出現頻度の高い文字が前後にくるものが上位を占め、159種類のうち、12種類で、全体の約50パーセントに達する。これらが必ずしも名詞の「こと」とはかぎらないことは、先に述べたとおりであるが、まず、そのほとんどが名詞であることは、次の例からも察せられる。

- ④早急に手続きをすすめることにしており、
- ④捕虜になることを耐えがたく思いながらも
- ④別の病院に申込むようにとのことでした。
- ④ベトコンの襲撃を受けて大損害を出したことがあった。

表10 「ひらがなで」始まる動詞の内訳

なる	76	とる	9
する	63	できる	9
ある	50	わかる	7
いう	42	はかる	7
よる	21	もつ	6
つく	17	つくる	5
いる	9	(その他)	159
みる	9	計	490

(数字は出現回数)

が読みやすさを妨げないだけでなく、むしろ、意味の把握を容易にしている面さえあると考えられる。

名詞について述べたことは、動詞についても当てはまる。表10は、ひらがなで始まる動詞のうち、直前にひらがながくる490例を整理したものである。上位の「なる・する・ある・いう」の4語で、全体の約47パーセントを占める。

「いう」の42例のうち、34例は、㉔のように、「と」が前にくる。「よる」の21例は、すべて「～による・～によって・～へによれば」という慣用的な使い方である。「つく」の17例のうち、4例は、「調整がつく・汚れがつく・気がつく・座につく」のように使われ、他の13例は、「～について」という助詞に相当する言い方の場合である。しかも、はじめに述べた規則によって、「…ている・…てみる・…である・〇〇する」などの補助動詞的なものは、ここに含まれていない。それでも、上位を占めるものの中には、準補助動詞的なものや慣用的なものが多いことが予想される。以下に、頻度の高い動詞を含む例をあげる。

- ㉙ ヤジ合戦も始って、混乱状態になった。
- ㉚ 結婚を前提としないこのような交際はやはり誠実でない
- ㉛ その土地が水びたしになっていたこともあったが、
- ㉜ 北ベトナムを国連に呼ぼうという非現実的な構想を
- ㉝ 日米貝類衛生協定によってきびしい条件をつけている。
- ㉞ 決意がある旨を示唆したものとみられる。

さきほどの「こと」と同様に、頻度数の最も多い「なる」例をにとって

以上をまとめれば、ひらがなに続く名詞の大部分は、形式名詞またはそれに準ずるものである。そして、その前後にくる文字は、比較的、少ない種類のものがひんばんにあらわれることが多く、一定の文字列を作りやすい。したがって、読み手にとって、かなの連続

表11 文字列〔なる〕の前後に出現する
確率の高い表記記号

前 部 分		後 部 分	
に	325	て	164
と	131	た	159
く	94	、	76
ば	55	●	76
●	42	な	75
は	17	。	35
る	13	と	30
(その他25種)	128	(その他33種)	190
計 (32種)	805	計 (40種)	805

(数字は出現回数)

表12 文字列〔なる〕を含む
出現頻度の高い4文字列

になって	107
になった	73
ばならな	44
となった	39
くなつた	29
となつて	24
となり、	23
になる。	21
になり、	18
くなつて	16
●なら●	14
になりま	13
くなり、	13
(その他216種)	371
計 (229種)	805

(数字は出現回数)

「文字列表」の集計結果を、表11と表12に示す。ただし、「なる」は活用語なので、活用形として考えられる「なら・なり・なる・なれ・なる・なつ」をすべて合わせて集計した。動詞「なる」以外のものも含まれていることは、「こと」の場合と同様である。

前部分では、「に・と・く」の3字で、約68パーセントを占める。後部分では、「て・た」の2字で、約40パーセントに達する。句読点が多くあらわれることも、「こと」とは異なる特徴である。4文字列の「ばならな」は、「…なければならぬ」であり、「●なら●」は、「雨なら中止だ」のように、助動詞「だ」の仮定形「なら」が多いものと思われる。〔なる〕を含む4文字列では、「になって」「になった」の2種類で、約22パーセントを、上位8種類では、約45パーセントを占めている。〔なる〕を含む4文字列の頻度の高いものの例を次にあげる。

- ㊦政界全体が腐敗する原因なつているという事実を
- ㊦身軽なつた二人は通勤バスに乗って、
- ㊦国民の半分以上が、生活が苦しくなつたと感じている。

⑤森構想を支持するのは与党民主党だけなった。

この結果を、さきの「こと」を含む4文字列と合わせてみると、7文字列では、「ることになって」「ることになった」「うことになって」というような文字連続があらわれやすいということが言えよう。次の⑨・⑩のように、かなが連続する場合でも、そう読みにくさを感じないとするならば、それは、出現回数が多いことによって、かな連続のゲシタルトができやすいことによると思われる。

⑨計量法によって一年一回検査するこ。こ。な。な。っており、都内を順々に回って調べている。

⑩通説では佐藤・池田・藤山の三派の間で使われたカネは十数億というこ。こ。な。な。っている。

なお、形容詞は、70例のうち、「ない」(26例)と「よい」(9例)で、全体の半数を占める。他の語は、特に、かな書きになりやすいというものはなく、分散しているようである。全般に、形容詞は、動詞・名詞にくらべて、「ない」を除けば、出現頻度の高い文字列を作りそうなものはない。かな表記される形容詞は、⑪・⑫のように、漢字で書かれるのが普通なのに、かな表記になる場合は少なく、⑬・⑭のように、現代語では、もともと、かな書きが普通であるものが多いように思われる。

⑪とにかく一日もはやく十万円たまることを祈ろう。

⑫出版社があたらしくなった機会に

⑬日本の奥さん方にはうらやましい話である。

⑭団交について即答できないのはおかしい。

7. ひらがなで始まる動詞の前の文節の文字列

ひらがなで表記される名詞や動詞の大部分が出現頻度の高いもので占められ一定の文字連続を作りやすいとしても、それだけでは、かな表記される理由が説明できない例は、たくさんある。表10にあがっている、「みる・とる・もつ・つくる」などは、上位の「なる・する・ある」などどちがって、「見る・取る・持つ・作る」のように漢字表記されることも多いはずである。また、新聞

では、漢字表記したくても、表外字であるために、かな書きすることもあると思われる。

そこで、かな書き文節の分析として、次のようなことを調べてみた。すなわち、かな書き動詞の直前の文節が、どのような文字列からなっているかということである。「気をつける」という文字列で、「つける」が「付ける」と書かれることが少ないのは、一つには、「気をつける」が一語のように意識されているので、「付ける」と書くことによって、「●○／●○○」のように、二つの単位として意識されることを避けようとする気持ちがはたらくためであろう。そして、「気」が、かな書きされることがほとんどないということも、「つける」を漢字表記しないでもすむ、一つの要因であろうと思われる。したがって、かな書きの文節が連続することを避けようとするために、直前の文節は、漢字書きにするなどの方法をとることが考えられるし、逆に、漢字書きの文節のあとにはかな表記の動詞があらわれやすいのではないかという推測をしてみた。

表13は、ひらがなが前部分にくる、ひらがなで始まる動詞 490 例を、直前の文節の文字列を構成する表記記号の種類によって、分類したものである。すべて、ひらがなを含むことは当然であるが、そのほかの表記要素に何があるかという基準でわけてみた。

表13 「ひらがな」で始まる動詞の直前の文節の文字列構成

漢字とひらがなからなるもの	351
かたかなを含むもの	23
記号を含むもの	28
ひらがなのみからなるもの	86
その他	2
計	490

まず、漢字とひらがなからなるものが圧倒的に多い。全体の約70パーセントを占めている。㉔のように、漢語の名詞は、ほとんど漢字表記される。㉕のような和語の名詞や、㉖のような動詞でも、漢字を含み、ひらがな以外の表記要素を含まないものは、このグループに入れた。㉗の場合には、後に述べるように、「を」という文字の役割を無視するわけにはいかない。また、㉘は、慣用句なもので、「かかわる」が漢字表記されているか否かは、それほど大きな要素ではないと思われる。

㉔生産者側との調整をはかる方法は

㉕泥しぶきをあびせけて通りすぎる自動車、

⑥佐藤政治は、好むと好まざるとにかかわらず、当面続くのである。

次に、かたかなを含むものは、⑥のように外来語の場合はもちろんであるが⑨・⑩の場合は、表外音訓に該当するもので、かたかなで書くことによって、漢字と同じ機能を持たせようとしたものとみられる。

⑨話し合いをモットーにしていながら、

⑩そのままかかずと、あとでシミになることがあります。

⑪今年の米価問題の重要なカギとなろう。

記号を含むものは、⑪・⑫のように、直前の文節が、作業規則では、記号のところで分割されているが、記号によって強調されたものに助詞がついたケースである。また、その他の2例は、数字やアルファベットを含むものである。

⑪この“誤解”をとくため、

⑫「秋の空となんとか」といわれるように

⑬主人公がC I Aという実在機関のスパイだとはっきり言っているのに興味があるんだそうです。

以上にあげた4種類の文節は、いずれも、直前の文節に、ひらがな表記の文節との連続をさけるなんらかの目じるしが含まれているもので、全体の約82パーセントを占める。したがって、かな表記の動詞の直前の文節には、意味のまとまりを示す表記記号が含まれている確率が高いといえよう。

残りの約18パーセントは、直前の文節がひらがなのみからなるものであり、490例中86例を数える。これらの文節は、名詞を含む文節、動詞をふくむ文節、形容詞からなる文節、副詞からなる文節の4種類に分けることができる。

名詞を含む文節は、45例で、そのうち、「こと」を含むもの21例、「もの」を含むもの7例であった。そのほかは、例⑭・⑮のように、かな表記が普通であるものが多いが、⑯のように、漢字表記できるものもある。

⑭常識的な考え方を示したものといえよう。

⑮交渉へのきっかけがつかめるとの見方は少ないようだ。

⑯自分がいやになることもあります。

⑰社会不安のたねになっていることは間違いない。

動詞を含む文節の場合は、後にくる動詞が出現頻度の高いものが多く、13例

のうち、「する」4例、「いう」3例、「なる」3例、その他3例であった。

㊸内容を多角的にしようとしている。

㊹西洋の白人の古い使命がいまやアメリカの手にゆだねられたという考え方である。

㊺いちばん最初に出たお茶の葉をつんでつくったもの。

形容詞からなる文節8例のうち、5例が後部分に「なる」がくるもので、その他が3例であった。

㊻海水のよごれがひどくなっており、

㊼党をよくするのが私の念願である。

副詞からなる文節は、前に ページで述べたように、前に読点がきたり、漢字やかたかなを含んだ文節であったりすることが多いが、この場合の20例のうち、前部分に、漢字を含んだ文節がくるもの14例、句読点・記号がくるもの4例、ひらがなのみからなる文節に続くもの2例であった。したがって、かな書きの動詞の直前にくる、かな書きの副詞は、その前に、なんらかの切れ目をもなうことが多いと言えよう。

㊽清浄海域がだんだんせばまっている。

㊾社会党の主張をほとんどいれてきた。

㊿刻印のはっきりした物は二つだけ。よくわからないが刻印のあるもの一つ。

㊿本格的に活発化するかどうかは、まだはっきりとしない。

これらの、かなで始まる動詞の前に、ひらがなのみからなる文節がくる場合をまとめてみると、名詞・動詞・形容詞などがくる場合は、出現頻度の高い語とともに、出現確率の高い文字列が作られる可能性が強く、副詞の場合には、その直前の文節に、意味のきれ目を示す表記記号が出現することが多い、ということになる。つまり、かな書きの文節にかな書きの動詞が続く場合、意味の切れ目がわかりにくくなるために、抵抗が感じられるというケースは、あまり多くない、という結果になる。(名詞の場合には、多少問題はあるが。)

形容詞についても、動詞と同様の傾向がみられるが、総数が少ないので、動詞ほど細かい分析はできなかった。ひらがなに続く、かなで始まる形容詞70例の

うち、前の文節が漢字・かたかな・記号などを含むものは51例、ひらがなのみからなるもの19例である。かなのみからなる文節が前にくる例を以下に示す。

㉞ 帰ってくればこなうらしいことはないが

㉟ 国の手をかりなればむつかしいでしょうね。

㊱ 諸君の良識に待つほかはない。

なお、ここで、いわゆるまぜ書きについてふれておく。まぜ書きとは、一般に、複合語の語構成単位のいずれかが（複数の場合もある）漢字表記され、ほかの単位がかな書きされるものをいうのであるが、ここでは、いちばんはじめの単位が漢字書きで、あとがかな表記されたものは、対象としない。文節の文字列の類型で言えば、

ひとまとまりを示す機能を持っているのではないかということが考えられる。たとえば、㊸の場合、「帰」という漢字があることによって、その文節の終わりに「て」があらわれることに何らかの予期を与えているのではないかということである。（この場合には、「っ」や「ひ」の文字連続中における機能をも考えなければならないが。）なお、いわゆる漢語サ変動詞の前部分がかな書きになったのは、㊸の1例だけであった。

- ㊸学校から帰^てて^てひと^と休^まみ^ましたのち、
- ㊹「兄^をさんをさが^して^くだ^さい^い」と^とふる^え声^{こゑ}で^で訴^うえ^たた。
- ㊺近^く二^ひ百^ご十^万ド^ルを^を自^ら発^的に^にき^ょ出^する。
- ㊻共^に関^与会^議派^の学^生が^が会^場に^にな^だれ^込み^み、
- ㊼この不^信が^がす^ぐに^にぬ^ぐい^去れ^ない^いのは^は当^然で^であ^る。

8. 意味の切れ目を示す表記記号

これまで、文節の始まりが漢字であることを原則として、それがかな表記される場合、意味の切れ目を示す機能に、どのようなかわりを持つかという立場から述べてきた。いわば、漢字の機能のほうに重点があったわけである。しかし、文節が漢字で始まりかなで終わるというタイプをとることを原則とするならば、文節の終わりにあらわれるかなにも、句読点のように、意味のまとまりや切れ目を示す機能を持つものがあるのではないかという推測は不可能ではなからう。漢字がほぼ無限といってよいくらいあるのに対し、かなは有限であり、しかも、たかだか数十種類にすぎない。それらを、文節のはじめにあらわれやすいもの、途中にしかあらわれないもの、文節末にあらわれやすいものといったような特徴によって分けたり、漢字のあとによくあらわれるもの、句点の前にくることの多いものという分類をすることもできよう。

表14は、先に引用した「文字列表」をもとに、ひらがな・句点・読点について、その前後にどのような種類の表記記号が出現したかを集計したものである。ただし、これは、出現頻度を示すものではなく、4文字列の第2字目をキーとして排列したものをもとにしたので、ある文字が4文字列の第2字目に来たときに、どれだけの種類の文字列を構成しているかということを示すもの

である。たとえば、「るこ[。]が」「たこ[。]が」という文字列は、前者が56回、後者が13回出現しているが、それぞれを1回と数えることになる。しかも、4文字列の表を使ったため、「るこ[。]が」と「るこ[。]を」は、「こ」を中心にしてみれば、「ること」という文字列になるが、これをそれぞれ1回と数えてあるその意味では、この表は不完全なものであるが、3文字列の後部分に重点を置いて分類した類型とみることもできよう。

表14 多くの出現類型を持つひらがなの文字環境（句読点を含む）

文字	出現 類型数	前 部 分					後 部 分					確 率 の 高 い 環 境	
		●	○	、	。	その他	●	○	、	。	その他	前	後
の	7,428	4,620 (62.2)	2,176 (29.3)	8	2	622	5,576 (75.1)	1,294 (17.4)	17	20	521	●	●
、	6,416	1,853 (28.9)	4,317 (67.3)	—	—	246	4,651 (72.5)	1,037 (16.2)	—	—	728	○	●
に	4,319	2,804 (64.9)	1,251 (29.7)	1	1	262	2,167 (50.2)	1,749 (40.5)	206 (4.8)	4	193	●	●
は	3,791	1,993 (52.6)	1,560 (41.2)	12	3	223	1,913 (50.5)	774 (20.4)	754 (19.9)	—	350	●	●
を	3,472	2,606 (75.1)	517 (14.9)	—	—	349	2,384 (68.7)	943 (27.2)	55	5	85	●	●
が	3,125	1,783 (57.1)	1,107 (35.4)	4	—	231	1,554 (49.7)	875 (28.0)	536 (17.2)	2	158	●	●
と	2,842	939 (33.0)	1,396 (49.1)	71	13	423	879 (30.9)	1,588 (55.9)	248 (8.7)	16	111	○	○
で	2,672	1,465 (54.8)	978 (36.6)	4	5	220	897 (33.6)	1,266 (47.4)	390 (14.6)	8	111	●	○
。	2,647	344 (13.0)	2,169 (81.9)	—	—	134	1,589 (60.0)	510 (19.3)	—	—	548	○	●
い	2,587	616 (23.8)	1,876 (72.2)	60	26	18	409 (15.8)	1,783 (68.9)	71 (9.4)	242	82	○	○
し	2,474	1,380 (55.8)	1,024 (41.4)	14	6	50	126 (5.1)	2,023 (81.8)	283 (11.4)	—	42	●	○
た	2,470	164 (6.6)	2,270 (91.9)	3	14	19	623 (25.2)	1,021 (41.3)	23 (26.4)	651	152	○	○
る	2,270	160 (7.2)	2,059 (92.7)	—	—	51	804 (36.2)	704 (31.7)	43 (22.8)	507	212	○	●
て	1,949	113 (5.8)	1,833 (94.0)	—	1	2	814 (41.8)	728 (37.4)	308 (15.7)	—	99	○	●
な	1,962	710 (36.2)	1,110 (56.6)	40	15	87	479 (24.4)	1,424 (72.6)	8	2	49	○	○
か	1,913	583 (30.5)	1,233 (64.5)	18	11	68	89 (4.6)	1,619 (84.6)	77	83	45	○	○
も	1,685	534 (31.7)	1,045 (62.0)	26	9	71	705 (41.8)	721 (42.8)	194 (11.5)	—	65	○	○
ら	1,445	162 (11.1)	1,287 (88.5)	—	—	6	537 (36.9)	727 (50.0)	129	7	55	○	○
れ	1,261	223 (17.7)	1,038 (82.3)	—	—	—	58 (4.6)	1,080 (85.6)	112	4	7	○	○
り	1,231	388 (31.5)	841 (68.3)	1	—	1	288 (23.4)	582 (47.3)	294 (23.9)	22	45	○	○

表14は、出現類型数の多いもの、上位20字までについて、その前後の表記記号の種類を示したものである。右端に、●と○で示したのは、前後で、それぞれ、もっとも出現確率の高い表記記号である。この表によって、全体を概観すると、「ゝ」を除けば、上位に、前後に漢字のきやすいものが集中し、15位ぐらいまでに、前後のどちらかに漢字がきやすいものがあらわれている。それ以下は、すべて、前後とも、ひらがながあらわれやすいものということになる。前後に漢字があらわれやすい「の・に・は・を・が」は、いずれも、1字が1助詞に相当する場合が多いものであり、出現頻度も多いと考えられる。これら、多くの出現類型を持つ漢字ほど、文節の切れ目を示す役割を果たしていることが予想される。

前部分に漢字がきやすいものの中で、比率が高いのは、「を・に・の」などで、逆に、「て・る・た」などは、ひらがながくる確率が90パーセントを越え、「ら・れ・。」がそれに続いている。後部分に漢字がきやすいのは、「の・ゝ・を」などで、ひらがながあらわれやすいのは、「し・か・れ」などである。句読点の後ろにくるひらがなは、平均していて、特定の文字への集中度は少ないが、句点の前にくるひらがなは、「た・る・い・だ・う・す」の6種類で、総類型数の約90パーセントに達する。また、読点の前にくるひらがなは、「は・が・で・て・り」の5種類で、総類型数の50パーセントを占める。このような結果に、今回は行なわなかったが、文節のどの位置にきやすいかという集計を合わせると、意味の切れ目を表わすひらがなの性格を明らかにすることができるであろう。

「を」という字は、前部分と後部分のどちらにも漢字のくる確率の高い文字であるが、現代かなづかいでは、この文字は、助詞の「を」を表わす以外に用いられず、きわめて表意（語）性の強いひらがなである。したがって、たとえば、前後に漢字があらわれなくても、すなわち、ひらがなが連続しても、この「を」によって、文節の切れ目は、明瞭なはずである。とすれば、「を」の前後では、ほかの文字の場合よりも、ひらがな表記される確率が高いのではないかという推定がなりたつ。

この推定を確かめるために、次のような集計を行なった。「を」のつぎにひ

らながきやすいかどうかを、「を」と同じように格助詞を表わすことの多い「が」を選んで、比較をしてみた。「が」を選んだ理由は三つある。「が」は、格助詞のほかにも、接続助詞にも用いられるし、語中にあらわれることも多く、「を」に比べれば、表語機能は劣ると考えられること。また、同じ条件で比較するためには、「を」も「が」も、格助詞として用いられる場合に、前部分に名詞を含む文節、後部分に動詞を含む文節という形をとりやすいこと。そして、それぞれの後にくる動詞を、いわゆる自動詞と他動詞という概念でとらえやすいことである。

そこで、「を」「が」の直後に、和語の動詞がくる例を、今回のデータからすべて抜き出して比較をしてみた。その結果、「を」の後に、和語動詞がくる例は、227例あり、そのうち、漢字表記されたもの-137例、かな表記されたもの-90例であった。また、「が」に続く和語動詞は、85例で、そのうち、漢字表記されたもの-35例、かな表記されたもの-50例であった。なお、ませ書きの複合動詞は、すべて、前部分の表記によって分類した。この数字だけを見ると、先の推定とは逆に、「を」のあとには、漢字があらわれやすく、「が」のあとには、ひらがながあらわれやすいかのような印象を受ける。しかし、実は、これだけでは、厳密な比較とは言えない。なぜならば、和語動詞といっても、語の性質によって、その表記には、さまざまな違いがあり、一概に、漢字かかなかといった色分けをすることは危険である。そして、対象が新聞の文章であることを考えると、和語動詞の表記には、四つのタイプがあると思われる。

その基準となるのは、「当用漢字表」である。まず、①のグループとして、「当用漢字表」、「同音訓表」の範囲で、漢字表記することが可能であり、実際に漢字表記されたものがあげられる。「当用漢字表」以外の漢字を用いて表記されたものは、1例もなかった。したがって、漢字表記されたものは、すべて、このグループにはいる。実例は、多くあげるには及ばないであろう。

⑤これに反対の立場から政府追求を続けてきた。

⑥私が住んでいる古河市は、

⑦のグループは、上のグループとまったく同じ事情にありながら、実際には、ひらがな表記されたものである。すなわち、一般に漢字表記が普通であ

り、漢字表記できるのに、ひらがな表記されたものということになる。この中でも、㊸のように、ひらがな表記されることがあまりないものから、・のように、かな表記がそう珍しくないものまで、各種のレベルが存在する。

㊸折箱料理をたのしむゆとりができたんじゃないからしら。

㊹ことしからアマ・スポーツ番組をつくるはずだったが、

㊺マクナマラ米国防長官の暗殺をはかって捕えられ、

㊻タム・チャウ師に不満をもつ反政府派が、

㊼おじいちゃんにズケズケとものをいい、

㊽ちょうど台風がきて、その土地が水びたしになっていたこともあって、

㊾沈滞していた中小企業の設備投資意欲に、回復のきざしがみえてきた。

㊿日本政府の熱意がたりないからだ。

㊸のグループは、漢字で表記することが、まったくないというわけではないが、戦前から、かな表記が行なわれ、現在では、かな表記が普通と思われるもの、および、漢字表記の形式が考えられないものからなる。たとえば、「当用漢字音訓表」で訓が認められている「ある(有る)・いる(居る)」のようなもの、ここに含めてある。したがって、㊸のグループとの境界は、必ずしも明確ではない。

㊽気分転換をなさってみませんか。

㊾ずいぶんコウをやったようだけど、

㊿全党的に再批判する必要がある。

㊽起きる前がつらくて頭痛がします。

㊸のグループは、表外字・表外音訓に該当するもので、かな表記が普通であるという点では、㊸のグループと共通する。新聞では、ほとんど漢字表記されることはないが、新聞以外では、漢字表記されることもありうるという意味で、㊸のグループとは区別した。以上㊸・㊹・㊺が、実際になな表記されたものである。

㊽アズキ色の丸いスタンプをはっていた。

㊾朝の出勤時の混雑をねらったものとみられている。

㊿生ガキの中毒がふえている。

表15 「を」「が」に続く和語動詞の表記

		を		が	
		ことなり	の べ	ことなり	の べ
漢字表記が 一般のもの	①漢字で表記されたもの	87	137	28	35
	②かなで表記されたもの	33	50	8	10
かな表記が 一般のもの	①現代語では、かな表記が普通のもの	6	15	5	29
	②表外字・表外音訓に該当するもの	21	25	9	11

(112)…という人々で辞典の売場がにぎわっている。

「を」「が」に続く和語動詞を、以上の四つのグループに分類して、異なり語数と延べ語数別に示したが表15である。「が」の①のグループは、他のグループにくらべて、異なり語数と延べ語数の比に著しく差があるが、これは、「ある」が24例を占めているためである。すなわち、「が」について、ひらがな表記された50例のうち、約半数が「ある」ということになる。このように考えると、「を」「が」のあとにくる動詞が、漢字表記されやすいか、かな表記されやすいかという比較は、結局、①のグループと②のグループについて行なうのが、妥当であるということになる。異なり語数についてみれば、「を」の場合は、かな表記される割合が、120例中33例(27.5パーセント)、「が」の場合は、36例中8例(22.2パーセント)である。延べ語数では、「を」の場合、かな表記される割合が、187例中50例(26.7パーセント)、「が」の場合は、45例中10例(22.2パーセント)と、どちらの場合にも、わずかながら、「を」のあ

表16 「を」「が」に続く漢字表記できる動詞がかな表記される割合(数字は、パーセント)

	を	が
異 な り	27.5	22.2
延 べ	26.7	22.2

とにくる和語動詞が、ひらがな表記される割合が、「が」の場合よりも上回っていることになる。(ただし、カイ自乗検定では、有意な差はみられなかった。)

次に、ある特定の動詞について、それが前にくる文字によって、どのよう

に表記形式が変わるかを調べてみた。そのためには、ある語について、多くの用例を集める必要がある。そこで、今回の語彙調査の約39に当たるデータのK W I C索引によって、「つくる」という語を選び、120例を得た。この120例を表記別にみると、次のようになる。

作る 62例
造る 2例
つくる 56例

これを、直前の表記記号によって分類すると、「を」につづくものが65例で、そのほかは、「で」の12例をはじめとして、ひらがな13種類、漢字3種類で、合計が55例となる。そこで、「を」とそれ以外のものに分け、そのあとの「つくる」がどのような表記になっているかに分類してみたのが、表17である。

「を」に続く場合に、ひらがな表記が、あらわれる割合は、47.7パーセント、「を」以外の文字に続く場合は、45.5パーセントで、わずかながら、「を」に続く場合のほうが、ひらがな表記になる率が高い。(この差は、もちろん、統計的に有意な差ではない。)

表17 動詞「つくる」と直前の文字の関係

表記 前部分	作る・造る	つくる
を	34	31
を以外	30	25

以上、「を」に続く和語動詞がひらがな表記になりやすいか否かを見たが、わずかながら、ひらがな表記になりやすいという結果が出ている。しかしながら、この違いは、僅少であって、はっきりした傾向としては、とらえることができなかった。この程度の少量のデータでは、これ以上の比較を試みることは、無意味と思われるが、少なくとも、この結果から、「を」につづく和語動詞がひらがな表記になりやすいという推定を否定することはできないであろう。

9. おわりに

この調査でいままでに述べてきたことをまとめると、次のようになる。

- (i) 文節は、漢字で始まり、ひらがなで終わる傾向が強く、とくに、■■■■□□

という類型を持つものは、漢字とひらがなのみからなる文節の約60パーセントをしめる。

- (ii) ひらがなで始まる文節のうち、接続詞・連体詞・副詞などは、大部分がかなのみからなる類型□□□□に属し、文のはじめや、読点のあとに用いられることが多い。
- (iii) ひらがなではじまる文節のうち、名詞・動詞・形容詞は、直前に、ひらがなのくる確率が高い。
- (iv) ひらがな表記される、名詞・動詞の大部分は、出現頻度の高い語によって占められ、その語を含む、出現確率の高い文字列を作りやすい。
- (v) ひらがな表記の動詞・形容詞の直前には、漢字やかたかなを含む文節がきやすい。
- (vi) ひらがなの中には、文節の終わりに用いられることの多いものがあり、ひらがなが連続する場合に、意味の切れ目を示す機能を持つことが推測される。

はじめにあげた、この調査の目的は、漢字かなまじり文の中で、漢字が、意味の切れ目やまとまりを示す役割を果たしているか、もし、漢字からその機能が失われているならば、それにかわるものは何かということであった。まず、前者についていうならば、現代の新聞においては、少くとも量的には、その機能は失われていない。しかし、それは、名詞・動詞・形容詞といった範囲に限られつつあるというのが正確であろう。二番目の目的については、一般にかな表記される傾向の強い、接続詞・連体詞・副詞などは、文字連続中にあらわれる位置によって、語のはじまりを示す機能がみられることを指摘することができた。また、名詞・動詞などでも、かな書きされることの多いものは、文字連続の出現確率や前後の文節の文字列構成、表意性の高いひらがななどによって、漢字の機能がカバーされている傾向をとらえることができた。

すなわち、本論でとった方法は、文字連続の中で、かな表記語が、どのような環境にあらわれるかということを手がかりにしたものである。なぜ、これらの語が、かな書きにされるかという理由は、ほかにもいろいろあるはずである。意味の面から言えば、同じ語でも、本来の意味から転化したり、実質的な意味

のうすいものは、かな書きになりやすいということもある*。また一つの訓に対して、種々の漢字表記がある動詞、(たとえば「かえるー変える・代える・替える・換える」)は、かな書きになりやすい***。そのほか、漢字の字形の複雑さ、漢字の表わす音節数****なども、かな表記と関係があるとされている。ある語が、かな表記されるか否かは、実際には、これらの要素のいくつかを、複雑にかみ合わさって決まるものであろう。この調査では、その要素のいくつかを、文字連続という観点から、とらえることができたと思う。

最後に、反省を含めて、この調査の問題点をふりかえることにする。まず、漢字表記語の分析を行わずに、かな表記語の分析を中心として、論を進めたことがあげられる。漢字表記語については、いくつかの部分を除いては、推測という結果にとどまらざるをえなかった。つぎは、いま述べたこととも関係するが、この種の調査でよく行なわれる。漢語・和語といった語種の面からの比較をしなかったことである。さらに、このような調査を行なうためには、規模がやや小さかったという事実を認めざるを得ない。全体の傾向をとらえることはできたと思うが、細かい問題になると、他の結果を援用しなければならず、統計的には無意味になってしまうことも多かった。そして、この調査の対象が新聞の文章であるということから、この調査の結果には、必ずしも一般の文章と一致しない点が多いはずである。それについては、深く検討する余裕がなかった。

以上のような問題点はあるが、このような調査を、さらに規模を大きくして行なうことによって、より精度の高い結果を得ることは不可能ではない。この

* 宮島達夫氏によれば、「みる」という動詞が、「絵をみる」のように視覚的な意味の強いものは、漢字表記される傾向が強いが、「ようすをみる」のような場合は、かな表記される傾向が強いとされる。(巻末文献⑤)

** この調査で出現した「はかる」という動詞は、7回あるが、音訓表では、「図る・計る・測る・量る」の訓がみとめられているにもかかわらず、「タイムをはかる」「調整をはかる」のように、すべて、かな書きである。

*** 「行なう」という動詞は、「現代雑誌九十種の用語用字調査」では、175例中、かな表記されたのは、7例である。「もつ」は、436例中、197例がひらがな表記されている。

ような調査が、正書法や基本漢字を考えるために役立つ資料となることは、もちろんであるが、とくに言語情報処理における、自動構文解析や、自動かな漢字変換において、未解決の問題を解明するためにはぜひ必要であると思われる。

参 考 文 献

- ①石綿敏雄「機械処理の対象としての漢字」(『情報処理』10巻9号)
- ②土屋信一「雑誌『太陽』の用字の変遷」(『言語生活』193号)
「明治・大正・昭和の漢字漢語の変遷」(森岡健二編著『近代語の成立』所収)
- ③斎藤秀紀「漢字かな混り文のエントロピー」(『計量国語学』43/44号)
- ④斎藤秀紀「漢字かな混り文の文字列」(『国研LDP』8号)
- ⑤宮島達夫「和語の漢字表記」(『教育国語』1970—12)