

国立国語研究所学術情報リポジトリ

Patterns of compound words of modern Japanese

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/10.15084/00001020

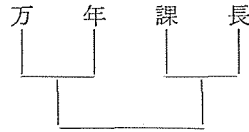
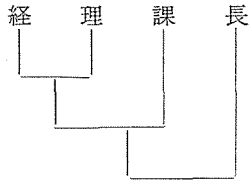
複次結合語の構造

野 村 雅 昭

1. はじめに

現代語を、その語彙的な側面からとらえようとする時に、避けられないのが語構成の問題である。それは、単に、語彙論における各種の分野と同等の重要性を持つことを意味するだけではない。語彙研究の基本となる、語（あるいは単語）の認定と深いかかわりを持つ問題だからである。特に、本研究所において実施してきた語彙調査のように、その量的構造を明らかにすることを目的とする場合に、調査単位をどのように設定するかということは、語構成の問題と不可分の関係にある。また、それは、語彙調査にとどまらず、漢字調査においても、同様な関係を持っている。漢字調査を単なる文字調査として考えるのではなく、漢字をある言語単位に相当する文字とみる立場によって考えるならば、次のような例は、語彙調査、漢字調査のいずれにおいても、かかわりのある問題である。

たとえば、「経理課長」と「万年課長」という二つの語が、さらに小さな単位の結合体であるという意識は、一般人の多くに、共通のものであると思われる。多くの人は、この二語を、「経理／課長」、「万年／課長」という、それぞれ二つの単位に分け、ともに「課長」という単位を含んでいることを指摘するであろう。しかし、これまで行なってきた語彙調査で、われわれは、このような方法をとってはこなかった。「経理課長」からは、「経理」・「課」・「長」の三つの単位、「万年課長」からは、「万年」・「課長」の二つの単位を採集するというのが、現行の語彙調査で言えば、「短単位」とよぶものを抽出する作業手順である。このような手順をとる理由は、われわれが、この二つの語を構成する四つの単位が、それぞれ、次のような関係で結合しているとみるからである。



すなわち、「万年課長」の場合には、「独身課長」・「ワンマン課長」などの系列の上にそれを置いて考え、「経理課長」の場合には、「人事課長」・「庶務課長」・「会計課長」などの系列と、「経理係長」・「経理部長」・「経理局長」などの系列との交点にそれを置いて考えようとするところに、その差異が生まれるのである。このように考える時に、語彙調査あるいは漢字調査と語構成の問題との間に不可分な関係があることは、容易に首肯されるであろう。

ところで、われわれがこれまで進めてきた、「電子計算機による新聞の語彙調査」では、「長単位」および「短単位」という二つの単位を設定している。これらの単位をとりだす手続きを詳しく説明する余裕はないが、概念的に言えば、「長単位」とは、文節から付属語を除いたものであり、「短単位」とは、現代語において意味をなう最小の言語単位を設定し、それがあつた条件の下に結合したもの、および、単独で用いられたものを言う。実際には、現在の調査では、まず、長単位語を抽出し、次に、それを分解して短単位語をとり出すという方法をとっているわけだが、この際に、種種の問題が生ずる。すなわち、先の「経理課長」の例にみたように、作業者が語の結合関係をどのようにとらえるかによって、とりだされる単位が異なる場合もまれではない。もちろん、考えられるかぎりの場合を想定した作業規則をつくるわけであるが、それでもなお、単なるミスではなく、結合の認定に苦しむ例は、決して少なくない。特に、今回の調査対象である新聞には、漢字・漢語にたよつた、複雑な結合体が出現し、作業者を困らせることになる。

このような事態は、単に、語彙調査における作業上の問題にとどまらず、言語情報処理におけるさまざまな分野でも起こりうる。自動構文解析、自動単位分割、情報検索などにおいて、処理対象となる語を、どのように認定するかということは、なお未解決の問題である。そうした事態から少しでも前進するた

めには、現代語の語構成を考えるのに十分なデータとそれに基づく分析が必要なわけであるが、現在までのところ、われわれの研究所が語彙調査の際に行なった分析を除いては、ほとんど行なわれていない⁽¹⁾といってよい。語彙調査の前提には、当然、その裏づけとなる語構成論があるべきであるがそれを組み立てる資料となる膨大なデータは、語彙調査によらなければ得られないという自家撞着的な事態からのがれるために、小稿では、一つの試みを行なってみようと思う。先に行なわれた、婦人雑誌や総合雑誌の調査の際の分析では、比較的、結合回数の少ない単位に焦点を合わせた形で、構成要素の語種や性格を分析することが中心になっていたが、ここでは、結合回数⁽²⁾の大きなものについて、それがどのような結合形態を持っているかということの分析を全体の視点とする。また、ここで対象とするデータは、きわめて小量であって、これをもって、新聞における、あるいは現代語における語構成の実態をとらえるということは、不可能であり、むしろ、その実態を記述する方法に対する、一つの試みと、筆者は、考えている。

なお、これまで、語構成ということばを用いてきたが、以下では、すでに形成されたある言語単位が、どのような部分要素の結合によって構成されているかという事実を「語構造」とよぶ、阪倉篤義氏の立場に従い、「語構造」ということばを使うことにする。また、ここでいう「構造」とは、この調査に出現した長単位とよばれる言語単位について、その部分要素があるパターンによって結合しているスタティックな形態をいう。

2. 調査の対象

ここで対象とするデータは、新聞語彙調査の一部分である、昭和41年の朝日新聞朝刊（1月～6月）の社会面に出現した、約8,000語（異なり数）である。その中から、次の(i)～(iii)に該当するものを対象から除いた。

-
- (1) 「語構造に関する分析」（国研報告4『婦人雑誌の用語』所収）
「語構成に関する分析」（国研報告13『総合雑誌の用語・後編』所収）
「複合語」の分析（国研報告25『現代雑誌九十種の用語用字』所収）
(2) 阪倉篤義『語構成の研究』

- (i) 記号, および, 記号を含むもの
- (ii) 数詞, および, 数詞を含むもの
- (iii) 固有名詞, および, 固有名詞を含むもの (固有名詞には, 人名・地名のほか, 会社名・商店名・学校名・劇場名・政党名・商品名・年号などの類をも含む。)

以上の操作を経て残ったものの中から, 漢語(字音語)を構成要素に含まないものを除き, さらに, 二つ以下の構成要素からなるものを対象外とした。このようにして得られた, 約 1,500 語が, 語構造分析の直接の対象となるデータである。

漢語構成要素を含まないものを除いた理由は, 和語のみからなる長単位語は, 概して, 少ない構成要素からなっており, 結合回数が少ないものが多いので, 分析の際の繁雑さをさげようとしたためである。また, 二つ以下の最小単位からなるものを除いたのは, いわゆる一字漢語や, 二字漢語の構造については, 先の分析で行なわれており, 今回の結合の順序のパターンを調べるという直接の目的からは, 除いてもよいと考えたからである。すなわち, 三つ以上の構成要素からなる長単位語が, ここでいう複次結合語の対象ということになる。

なお, サ変動詞や形容動詞は, その語幹部分の構成要素のみに注目したので「勉強する」「大変だ」などの長単位語は, 二つの構成要素からなるものとして除いた。「健康的だ」・「合理化する」のような場合は, 三つの構成要素からなるものと考え, 対象に含むことになる。

また, 最終次結合の一方に接辞がくるものは, それを除いた部分にのみ着目した。(接辞とは, 「お・ご・み」, 「たち・ら」など, 狭義のそれをさす。ただし, 「お役所仕事」のように, 接辞を含む単位が, さらに大きな結合単位の部分となっているものは, 構成要素に含める。)

したがって, 以上の約 1,500 語は, 漢語構成要素のみからなるもの, および漢語構成要素を含むものということになり, 後者には, 「救助袋」・「工事現場」・「海員組合」・「ママさん社長」のような結合単位も含まれることになる。

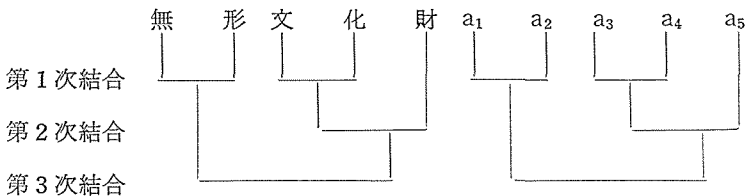
3. 語構造の記述

前章で、構成要素とよんだものを、先に第一章でもふれたように、われわれは、最小単位とよんでいる。最小単位とは、現代語で意味をになっている最小の言語単位で、現代人の意識では、それ以上、分割が不可能なものをいう⁽³⁾。たとえば、和語では、「母・山・草・走る・読む・高い・静か」などがそれにあたり、漢語（字音語）では、原則として、漢字で表記される1字分がそれに相当する。「運動」の「運」と「動」, 「新幹線」の「新」・「幹」・「線」などがそれである。（なお、以下では、長単位語を構成している個々の要素を、最小単位あるいは構成単位（または、単位）とよび、長単位語の内部における、最小単位の結合体を、構成要素とよぶことにする。）

今、この最小単位を「a」という記号で表わすことにし、n個の最小単位からなる語 P(n) を、次のような一般式で表わすことにする。

$$P(n) = a_1 + a_2 + a_3 + \dots + a_{n-1} + a_n$$

しかし、実際は、このn個の最小単位は、この語の中で、対等の関係にあるのではなく、意味的な結びつきによって、階層的な構造を持っている。たとえば、「無形文化財」という5個の最小単位からなる語の構造は、次のような図式で示すことができる。



このように考えると、n個の最小単位からなる語（以下では、「n単位語」とよぶ）は、一般に、(n-1)回の結合をし、最大、(n-1)次の結合をすることができる。（ $a_1 \cdot a_2 \cdot a_3 \dots$ が、語構造上、等質のものかどうかということは、今は問題にしない。）

上の「無形文化財」のような構造を持った語を、図式によらずに、一般式で

(3) 最小単位の詳しい定義は、『電子計算機による新聞の語彙調査』（国研報告）37に述べてある。

書き表わすならば、次のようになる。

$$P_{(5)} = (a_1 + a_2) + \{(a_3 + a_4) + a_5\} : \text{無形文化財}$$

同様にして、比較的、簡単な構造を持った語であれば、いずれも、一般式で表わすことができる。以下に、その例をあげる。

$$P_{(4)} = \{(a_1 + a_2) + a_3\} + a_4 : \text{庶務課長}$$

$$P_{(5)} = \{(a_1 + a_2) + (a_3 + a_4)\} + a_5 : \text{座席指定車}$$

$$P_{(5)} = \{a_1 + (a_2 + a_3)\} + (a_4 + a_5) : \text{核兵器実験}$$

$$P_{(6)} = (a_1 + a_2) + \{(a_3 + a_4) + (a_5 + a_6)\} : \text{後期中等教育}$$

このようにして、カッコの種類をふやしていけば、どのような構造を持った語でも、一般式で表わすことができるわけであるが、繁雑でもあり、実際には不可能である。そこで、以下では、次のような方法をとることにした。

まず、 $a = u$ と置き、最小次の結合 $u + u$ を U とする。そして、それに、 u が一つずつ加わることによって、結合次元がふえていく場合は、以下のように表わすことにする。なお、最終次結合の u が後部分にくる場合を v_p 、前部分にくる場合を v_q などのように示すものとする。

$$a = u \quad u + u = U \quad (\text{ただし, } u_1 + u_2 \neq u_2 + u_1)$$

$$(u + u) + u = U + u = v_p$$

$$u + (u + u) = u + U = v_q$$

$$\{(u + u) + u\} + u = v + u = w_p$$

$$u + \{(u + u) + u\} = u + v = w_q$$

$$[\{(u + u) + u\} + u] + u = w + u = x_p$$

$$u + [\{(u + u) + u\} + u] = u + w = x_q$$

⋮

以上のほかに、次のような結合の場合は、 V で置き換えることにする。

$$(u + u) + (u + u) = U + U = V$$

また、 $u \cdot v \cdot w \cdot x \cdot y \cdots$ の系列のものは、最終次結合の一方である場合

を除いて、w以降の記号を用いないことにする。たとえば、次のような場合である。

$$(u+w_p)+U=\{u+(v_p+u)\}+U$$

以上で定義した記号を用いて、長単位語の構造を、最終次結合を示す二項式で表わすことにすれば、先にあげた例は、次のように書き換えることになる。

$$P_{(5)}=(a_1+a_2)+\{(a_3+a_4)+a_5\}\rightarrow P_{(5)}=U+v_p:\text{無形文化財}$$

$$P_{(4)}=\{(a_1+a_2)+a_3\}+a_4\rightarrow P_{(4)}=v_p+u:\text{庶務課長}$$

$$P_{(5)}=\{(a_1+a_2)+(a_3+a_4)\}+a_5\rightarrow P_{(5)}=V+u:\text{座席指定車}$$

$$P_{(5)}=\{a_1+(a_2+a_3)\}+(a_4+a_5)\rightarrow P_{(5)}=v_q+u:\text{核兵器実験}$$

$$P_{(6)}=(a_1+a_2)+\{(a_3+a_4)+(a_5+a_6)\}\rightarrow P_{(6)}=U+V:\text{後期中等教育}$$

4. 出現した長単位語の構造

この調査に出現した長単位語を構成単位数別に示したのが表1である。3単位語が半数近くを占め、4単位語と5単位語を含めると、全体の90パーセントを越える。すなわち、複次結合語といっても、6単位以上の語が出現する割合は、きわめて低いと言えよう。しかも、

これは異なり語数の比率であって、延べ数で比較する場合には、さらに、その比率は下がるものと思われる。構成単位数が最大でも10というのは、少ないようであり、実際には、それ以上の単位数からなるものも出現しうるはずであるが、調査データが少なかったため、たまたまこのような結果になったものと考えられる。また、先に述べたように、固有名詞

表1 構成単位数別の出現語数

単位数	出現語数
3	720 (47.8)
4	525 (34.9)
5	153 (10.1)
6	66 (4.4)
7	28 (1.9)
8	10 (0.7)
9	2 (0.1)
10	1 (0.1)
計	1,505 (100.0)

()内の数字は百分比。

や数詞を含んだ語は、対象から除いてあるので、その部分を省いたものをこの表に加えれば、全体の比率は、多少、変わるものと思われる。ただし、6単位以上の語の比率が大幅に増加することはないであろう。

表2は、出現したすべての語の構造を二項式で表わしたものを、最終次結合

の一方の項の結合次数の小さいものに合わせて、分類したものである。(表1と総語数が一致しないのは、二項式では、書き表わせないものを除いたためである。それらについては、次の章で述べる。)

この表を概観して気がつくことは、u および U を最終次結合の一方に持つものが圧倒的に多いことである。かりに、この u および U を、漢語構成単位とすれば、いわゆる一字漢語と二字漢語が、語構造の中核となっているわけで、この点に関しては、これまでの語構成の調査とあまり異ならない常識的な結果を示している。結合パターン別の語数では、1 単位語を最終次結合の一方に持つ語のほうが多いことが注目される。(ただし、3 単位語の u+U、U+u、4 単位語の U+U を除けば、ほとんど差はなくなる。) また、結合のパターンの種類も、両者同数であり、出現語数の多いパターンには、 $u \leftrightarrow U$ の関係にあるものが多いことも、u あるいは U を最終次結合の一方に持つパターンの多いことを裏書きしている。そのほかは、v を一方に持つものが少数存在するのがめだつだけで、構成単位数の多い語でも、比較的、小さな結合パターンが累加的に組み合わさった構造を持っていることがわかる。たとえば、7 単位語に属する28語のうち、 $v+V$ あるいは $V+v$ のパターンに属するのは、9 語である。また、8 単位語で、 $V+V$ のパターンに属するものは、1 語もない。

次に、前部分と後部分の関係でみると、最終次結合の後部分に、結合次数の低いもののくる型のほうがはるかに多い。もっとも、4 単位語の U+U、6 単位語の v+v のように、結合次数が前後とも等しいものも、後部分に合わせて整理してあるので、それを考慮する必要があるが、それでも、パターン数、出現語数ともに、 $\square+u$ 、または、 $\square+U$ という型に属するものが多くなっている。ただし、このように、前部分が重く、後部分が軽いという一般的な構造が、構造要素の意味的な強弱関係をそのまま表わしているとみるわけにはいかない。 $\square+U$ の場合には、その長単位語の性格を決定する中核となる構成要素は、むしろ、U の部分にあるものが多いと言える。

U+U: 電話番号, 定期列車, 主任技師, 反対運動

v+U: 区議会議長, 消費者物価, 道德的觀念, 不起訴処分

V+U: 期末手当条例, 生活保護世帯, 遗体収容作業

表2 複次結合語のパターンおよび出現数

型		単位数				
		3	4	5	6	
二項式の結合次数の小さいほうの構成要素の型	u	前	u+U 142	u+v _p 22	u+V u+(v _p +u) 3/2	
		後	U+u 574	v _p +u 44 v _q +u 8 {(u·u)×u}+u 2	V+u 57	(V+u)+u 4 (U+v _p)+u 5 (U+v _q)+u 2 (v _p +U)+u 2 (v _q +U)+u 3 x _p +u 1
		小計	2/716	4/76	3/62	6/17
	U	前			U+v _p 39 U+v _q 3 U+(u+u+u) 1	U+V 8
		後		U+U 443	v _p +U 32 v _q +U 13 (u+u+u)+U 3	V+U 35
		小計		1/443	6/91	2/43
	v	前				
		後				v _p +v _p 3 v _q +v _p 2 v _p +v _q 1
		小計				3/6
	V	前				
		小計				
	V+u	前				
		小計				
	合計	前	1/142	1/22	5/48	2/9
		後	1/574	4/497	4/105	9/57
計		2/716	5/519	9/153	11/66	

*1 一般式の右横の数字は，出現語数を示す。

*2 6/171 のように示した，-6はパターン数を示し，171は出現語数を示す。

7	8	9	10	計	
$u+(U+V)$ $u+(V+U)$	1 1			6/171	
$(V+U)+u$ $(U+V)+u$ y_q+u	2 1 1	$(V+v_p)+u$ $\{(V+U)+u\}+u$	2 1	16/709	
5/6	2/3			22/880	
$U+(V+u)$ $U+(U+v_p)$ $U+(v_q+U)$	2 1 1	$U+(V+U)$	1	8/56	
$(V+u)+U$ $(v_p+U)+U$ $(v_q+U)+U$ $(U+v_p)+U$ $(u+V)+U$	2 3 2 1 1	$(V+U)+U$	2 $(V+v_p)+U$	1 $\{(V+U)+U\}+U$	13/539
8/13	2/3	1/1	1/1	21/595	
v_p+V v_q+V	2 1	$v_p+(V+u)$	1	3/4	
$V+v_p$	6	$(v_p+U)+v_p$ $(V+u)+v_p$	1 1	6/14	
3/9	3/3			9/18	
	$V+(u+v_p)$	1		1/1	
	1/1			1/1	
		$(V+u)+(v_p+u)$	1	1/1	
		1/1		1/1	
7/9	3/3	1/1		20/235	
9/19	5/7	1/1	1/1	34/1260	
16/28	8/10	2/2	1/1	54/1495	

(v+U)+U：核実験探知衛星，新生活運動協会

それに対して， $\square+u$ の型では， u の性格によって，主たる構成要素が u の部分にあるものと，そうでないものがある。後者に属するものは，ほとんどが漢語構成要素で，形式的な機能を持つものが多い。一般に， $\square+u$ の \square の部分の結合回数が多いものほど，その傾向は強く， u は， \square で表わされる内容を統括して，文法的な意味を添加する，いわゆる接辞性を帯びたものが多い。

U+u：指定券，大臣席 \leftrightarrow 安全性，革新系

v+u：焼死者数，送信用アンテナ \leftrightarrow 自治会側，成人式用

V+u：入試改善策，道路清掃車 \leftrightarrow 年中行事化，教育理論的

この表2で，さらに気がつくことは，理論的には出現しうるはずのパターンが1例もないか，あるいは，きわめて少数である場合が見られることである。その理由として考えられるのは，調査データの数が少ないことであり，データの量をふやせば，この表の空白部分のいくつかをうめることが可能であろう。しかし，必ずしも，そのような理由のみで説明しきれないケースも存在する。たとえば，4単位語では，U+Uのパターンに属するものももっとも多いが，u+vあるいはv+uのパターンに属するものも，497例中74例あるのに対し，5単位語で， $u+\square$ と $\square+u$ のパターンに属する62例のうち， \square の部分が，v+uの構造を持つものは2例あるだけである。特に， $\square+u$ の型では57例すべてがV+uとなっているのは，単なる偶然とは言えないものがありそうである。同様の例は，6単位語にもみられる。U+ \square と $\square+u$ の型に属する44例のうち， \square の部分にv+uがくるものは1例だけで，あとは，U+VまたはV+Uの構造を持つものである。

このような現象は，u+vあるいはv+uというような4単位の結合形態は自己完結的で，より大きな単位数からなる語の構成要素となる力に欠けることを物語っている。逆に言えば，この章の初めに述べたことを，次のような言いかたで補強することになる。多くの構成単位からなる複合語は，その構成要素であるU（2単位語）あるいはv（3単位語）の組み合わせからなっているもの，および，それにu（1単位語）が付加したものからなる構造を，その特徴

表 3 構成単位数と結合次数の関係 (数字は、パターン数)

構成単位数 結合次数	3	4	5	6	7	8	9	10	計
2	2	2	2	—	—	—	—	—	6
3	×	3	6	5	3	—	—	—	17
4	×	×	1	6	12	7	2	—	28
5	×	×	×	1	—	1	—	1	3
6	×	×	×	×	1	—	—	—	1

とする。

今、述べたことは、表 3 によっても説明することができる。先にも述べたように、 n 単位語は、最大、 $(n-1)$ 次結合をすることが可能であるから、調査に出現した語では、9 次結合まで可能であるのに、実際には、7 次結合以上の例は、出現していない。大部分が、3 次結合と 4 次結合で占められている。すなわち、構成単位数が多くなっても、それに比例して、結合次数の大きいものが多くなることは、それだけ低い次元で結合した構成要素を内包していることにほかならない。そのような構成要素が、 U や v であることは、くりかえすまでもない。

5. 各単位語の構造

前章では、複次結合語の構造についての全体的な傾向を概観したが、この章では、各単位語ごとに、出現した語例について、その特徴をながめることにしたい。

5.1 3 単位語の構造

3 単位語が複次結合をするのは、組み合わせ上から、 $u+U$ または $U+u$ のどちらかに限られる。すなわち、1 次結合をしている 2 単位語に、さらに、その前あるいは後に、1 最小単位が結合したもので、複次結合語としては、もっとも小さな結合パターンである。 $u+U$ のパターンに属するものと、 $U+u$ の

表 4 3 単位語の語種によるパターンの分類

U+u	(○+○)+○	(○+○)+●	(○+○)+△	(●+●)+○	計
	473 (82.6)	63 (11.0)	22 (3.8)	15 (2.6)	574 (100.0)
u+U	○+(○+○)	●+(○+○)	△+(○+○)	○+(●+●)	計
	105 (73.9)	19 (13.4)	16 (11.3)	2 (1.4)	142 (100.0)

() 内の数字は、パーセント。 ○=漢語構成単位
●=和語 " △=外来語 "

パターンに属するものとは、約 1 : 4 の比率で、後者のほうが多く、前部分に高次の構成要素がきやすい、全体的な傾向と一致している。また、4 単位以上の語の構成要素として、3 単位語が現われる場合の比率をみると、 $v_q(=u+U)$ が 36 語、 $v_p(=U+u)$ が 177 語と、単独で現われる場合の比率と似かよった値を示している。

3 単位語を構成単位の語種別に示したのが表 4 である。漢語構成単位のみからなるものが圧倒的に多いが、U+u と u+U では、多少その比率に違いがある。この違いは、u の部分に漢語以外の構成単位がくる割合と関係がある。それぞれの型に含まれている二字漢語 (○+○) の割合は、U+u では、97.4 パーセント、u+U では、98.6 パーセントと、ほとんど差がない。ところが、一字漢語 (○) の割合は、前者が 85.2 パーセント、後者が 75.3 パーセントで、約 10 パーセントの開きがある。すなわち、u+U のパターンでは、u の部分に、和語や外来語が現われる率が、U+u の場合より高いことを示している。このことは、u の部分に現われる漢語構成単位にも、なんらかの違いがあるのではないかということ予想させるが、以下、実際に出現した語例を比較してみる。(なお U の部分には、表にあげたもののほかに、いわゆる混種語も出現しうるはずであるが、今回のデータには、1 例もなかった。)

U+u のパターンに属するものの例を次にあげる。

- (i) (○+○)+○：現代人，占領軍，伝染病，品評会，昨年末，新築中，家族間，自主性，家庭用，具体化，絶望視，全国的

- (ii) (○+○)+●：救助袋，私書籍，製造元，現金払い，時間つぶし
- (iii) (○+○)+△：全学スト，石油ストーブ，非常ボタン，郵便ポスト
- (iv) (●+●)+○：受取人，組合長，組立式，大型機，小売店，締切分

(i) に属するものの u は，もっとも種類が多く，純体言的なもの，サ変動詞や形容動詞の語幹を形成するもの，連体修飾語的機能を付加するものなど，さまざまである。(ii) の u の部分には，純体言的なものと，動詞の連用形から転じた，いわゆる居体言とが含まれる。

u+U に属するものは，次のとおりである。

- (i) ○+(○+○)：都議会，市庁舎，核軍縮，食中毒，正反対，総工費，
全理事，大規模，不明朗，非能率，各選手，同理事
- (ii) ●+(○+○)：ひな人形，朝火事，初会合，荒れ模様，小道具
- (iii) △+(○+○)：ガス爆発，ダム建設，タクシー会社，スピード解決
- (iv) ○+(●+●)：同建物，不手際

(i) の u には，接頭辞的なもののほか，「都・市・核」のように名詞性の強いもの，「同，各」のように，連体詞に近い性格を持つものなどがある。(ii) には，居体言的なものは少なく，名詞性や自立性の強いものが多くみられる。

以上を比較してみると，一般的な特徴として，次のようなことが指摘できる。(iii) や (ii) の u の部分は，外来語最小単位や和語最小単位であるが，(i) の漢語最小単位と比較した場合，その自立性には，大きな差がある。漢語最小単位の多くが他の最小単位と結合して，はじめて安定した単位になりうるのに対し，外来語最小単位や和語最小単位は，単独でも安定した単位として用いられる。すなわち，(iii) や (ii) に属するものの比率が，u+U のパターンに多いことは，それだけ，u+U という結合形態の不安定性を示し，臨時的な結合が多いことを意味する。これと同様のことが (i) の類についても見られることは，上にあげたわずかな例からも察せられる。U+u の u が前部分と密接な関係を持ち，全体を一つの単位としてまとめあげる性質を持つものに対し，u+U の場合には，u は連体修飾的機能を持つものが多く，「市庁舎→市の庁舎」，「各選手→各の選手」，「全理事→全ての理事」というように，分離する傾向を内在している。このことが，一般に，□+u の型よりも，u+□ の型を

出現しにくくし、さらに、□の部分がより高次の結合になるほど、その前部分に、u が現われにくくなる現象の原因と考えられるのである。

しかし、上に述べたことを立証するためには、以上のような簡単な説明では不十分であることは言うまでもない。そのためには、多くの紙幅を費す必要があるが、小稿の本来の目的に鑑みて、以上の指摘にとどめることにする。それについては、別の機会に、稿を改めて論ずることにしたい。

なお、3単位語には、二項式で書き表わせない語が4例存在する。これを、二項式の例にならって、かりに、次のような一般式で表わす。

$(u \cdot u) \times u$: 重軽傷, 青少年, 新增設

$(u+u+u)$: 経団連

前者は、たとえば、「重軽傷」を例にとると、「重傷」と「軽傷」の二つが合わさったもので、一般式で書くならば、

$$(a+m)+(b+m)=(a+b) \times m$$

のようにも表わせるはずであるが、一般の結合形式である a+b の場合と、この場合とは、やや異なると思われるので、上のようにした。ただし、このパターンによる結合が安定し、語源意識がうすれてくると、「陶磁器」のように、

$$(u \cdot u) \times u \rightarrow (u+u)+u$$

語構造に対する意識が、上に示したように、変化する場合もある。

後者は、略語の例であるが、「松竹梅」、「天地人」などの、いわゆる三字漢語と同様に考えた。この場合は、三つの構成単位が対等の資格で結合しており、厳密に言えば、1次結合であり、複次結合とは言えないものである。

5.2 4単位語の構造

先にもふれたように、4単位語では、2次結合の U+U というパターンに属する語が圧倒的に多く、3次結合の語は、あまり多くない。また、3次結合のパターンでは、v+u のパターンに属する語のほうが、u+v のそれに倍している。

$u+v_p$: 総調達額, 軽飛行機, 各管制塔, 同動物園, ジェット旅客機

v_p+u : 幼稚園児, 郵便局長, 警備員室, 解剖学者, 不動産屋

v_q+u : 中学校長, 大御所格, 不景気風, 法医学者, クレー射撃用
 $U+U$: 火災防止, 公共事業, 住所氏名, 期末手当, ニュータウン区域
 $\{(u \cdot u) \times u\} + u$: 青少年法, 小中学生

このうち、 v_p+u のパターンには、 $U+U$ のパターンとの区別があいまいなものはいくつかみられる。上にあげた、「幼稚園児」、「郵便局長」、「解剖学者」などがそれである。これらは、「 $○○+○○$ 」のように分割しても、それぞれの後部分は、1次結合の漢語として存在しており、一般人の語構造意識としては、むしろ、そのほうが普通であるかもしれない。そうした意識を支えるのは、4単位語の大部分が2単位語プラス2単位語という形態を持っていることからくる、潜在的な語構成意識であろうと思われる。「法案」、「用地」、「業界」というような二字漢語が通用する背景には、その本来の語形成の事実とは別に、「 $○○$ 法案」、「 $○○$ 用地」、「 $○○$ 業界」といった、3次結合の4単位語が存在したことも、考慮に入れる必要があるであろう。

$U+U$ のパターンに属するものを、前部分と後部分の語種によって分けると、表5のようになる。前後とも、漢語からなるものは、90パーセント近くを占め、3単位語のすべて漢語構成要素からなるものの比率よりも、やや高い数値を示している。このことは、多くの構成単位からなる複次結合語ほど、漢語最小単位によって構成される傾向を示している。和語が前後にくる場合が少数あるほかは、ほかの語種が前後にくる割合は、きわめて少ない。外来語が結合する率が3単位語の場合にくらべてわずかしかならないのは、例にあげた、「ニュータウン」のように、外来語構成単位どうしが結合することが少なく、外来語は、1最小単位が、漢語最小単位の1次結合したものとほぼ同じ機能を持っていることによるものとみられる。これらの前部分と後部分がどのような意味的關係にあるかを分析することも、重要な課題であるが、今は、それについては、ふれない。

表5 $u+u$ のパターンに属する4単位語の語種別分類
 () 内の数字はパーセント。

(前部分) + (後部分)	出現語数
(漢語) + (混種語)	7 (1.6)
(漢語) + (外来語)	1 (0.2)
(漢語) + (和語)	15 (3.4)
(漢語) + (漢語)	393 (88.7)
(和語) + (漢語)	21 (4.8)
(外来語) + (漢語)	1 (0.2)
(混種語) + (漢語)	5 (1.1)
計	443 (100.0)

なお、4単位語で、2項式で表わせないものは、以下の3種6語である。

$(u \cdot u) \times U$: 小中学校

$(u+u+u) \times u$: 市町村長

$(u+u+u+u)$: 市区町村, 都道府県, 環衛組連, パン協組連

5.3 5単位語の構造

V+uのパターンに属するものが57例(37.6パーセント)ともっとも多く、4単位語でもっとも多いU+Uに、最小単位がうしろから結合したものが、5単位語でも多数を占めている。ついで、U+vが42例(27.5パーセント)、v+Uが45例(29.4パーセント)と、2単位語と2次結合した3単位語との組み合わせからなるものが多く、この二つを合わせると、半数以上を占める。これにもっとも多いV+uを加えると、5単位語全体の約94パーセントに達する。一方、4単位語にみられた、u+v、v+uのパターンに、uがさらに結合したものは、u+(v_p+u)の2例しか現われていない。

また、3、4、6単位語に比較すると、前部分が重いパターンと後部分が重いパターンとがほぼ同数であることも特色といえる。出現語数も、105:48と小さな比になっている。このことは、上に述べた、2単位語と3単位語の結合が、前後の別なく起こりやすいことによるものとみられる。

$u+(v_p+u)$: 同刑事部長, 同事務所長

$u+V$: 各郵便局番, 都清掃条例, 準保護世帯

$V+u$: 試験放送中, 有線放送網, 行政区画名, 売春防止法

$U+v$: 地下核実験, 大型外航船, 組合委員長, 女子作業員

$v+U$: 再編成論議, 道徳的観念, 大使館公邸, 旅客機事故

$U+(u+u+u)$: 船員中労委

$(u+u+u)+U$: 経団連会長

5単位語で問題になるのは、V+uとU+v_pのどちらに属するかあいまいなものかなりみられることである。たとえば、次の例は、いずれも、この調査では、V+uに所属すると判定したものであるが、U+v_pのように考えることも、そう不自然ではない。

債券所有者，宇宙飛行士，沿岸警備隊，地震観測所，入試改善策

このような問題は，4 単位語にみられた，「郵便+局+長→郵便+局長」の現象と似た性格を持っているが，構成要素の面からみると，やや異なる点がある。4 単位語の場合は，vp+u のパターンとみれば，1 次結合単位一つと，1 最小単位二つ，計三つの要素から構成されていることになり，V+V とみれば二つの 1 次結合単位から構成されていることになる。つまり，どちらのパターンに属するかによって，構成要素の数に変化があるわけである。しかし，5 単位語の場合には，V+u，U+vp のどちらのパターンに属するとみても，その語の構成要素として，二つの 1 次結合単位と一つの最小単位をとりだすことには変わりがない。したがって，従来の語彙調査で，短単位に分割する場合に，このパターンの違いは，あまり問題とはされなかった。しかし，この最後につく最小単位の語構成力を調べるためには，それが直前の 1 次結合単位だけを用いるのか，それとも，全体を統括しているのかということは，結果に大きな違いをもたらすことになる。

この現象を説明するためには，多少，意味的な関係に立ち入らなければならない。すなわち，V+u に属する 5 単位語が，すべて，U+vp と似かよった構造を持つわけではなく，同じ V+u に属するものでも，「婦人児童館」，「行政区画名」，「世界一周線」などの語例とくらべると，上にあげた「債券所有者」以下の例は，その構成要素間の意味的な関係に，ある種の類似を持っていることがわかる。

今，この V+u を，二つの 1 次結合単位と一つの最小単位からなると考え，A+B+c という式で表わすならば，これらは，「Aニツイテ，ソレヲBスルトコロノc」という言いかたで，その意味的な関係で表わせるものが多い。ところで，この c に相当する最小単位には，人間活動の主体を表わす，職業・地位・身分・組織・機関などに関する意味内容を持ったものが現われやすい。このような最小単位は，直前に，それを説明・限定する 2 単位語を伴って，比較的安定した。2 次結合の 3 単位語をつくるものが多い。たとえば，「所」を例にとれば，「裁判所・事務所・営業所・保健所・製鉄所・発電所」のような場合は，それ自体で安定した単位であり，さらに 2 単位語と結合して，「△△○

〇所」のような結合をしても、これらは、A+(B+c)という構造からなることが明瞭である。

ところが、上にあげた「観測所」や「研究所・養成所・貯蔵所・安定所・収容所」のような場合は、単独で用いられる場合もあるが、直前の2単位語が、「観測・研究・養成・……」のように、動作の対象をとりうる性格を持っており、本来は、(A+B)+cのような構造を持つものが、B+cの結合度いかんによっては、A+(B+c)のような解釈を可能にするわけである。たとえば、「研究所」などは、結合度が強く、「〇〇研究所」のような結合形態を持っていても、(A+B)+cという意識は強くない。それに対して、「捕虜収容所」のような場合は、A+(B+c)のように考えることは無理で、単独で、「収容所」と使われる場合も、略語的な印象をまったく伴わないわけにはいかない。5単位語における、V+u ⇔ U+v_pの問題は、以上のように考えることができる。

5.4 6単位語の構造

6単位語では、V+Uのパターンに属するものが、約半数を占め、□+uに属するものがそれに続いている。u+□に属するものは、1例もない。また、v+vのように、2次結合の3単位語二つからなるものがあることも注目される。以下に、主な語例を示す。

(V+u)+u：都市交通課長，宇宙開発史上

(U+v)+u：身体不自由児，航空代理店員

(v+U)+u：小学校入学児，自動車修理工

x_p+u：中学校長会長

U+V：脳性小児マヒ，総合警備本部，事故合同会議

V+U：都市交通問題，消費減退傾向，高校全入運動

v+v：社会科副読本，野鳥用給水池，都教組委員長

U+VとV+Uに属するものを合わせると、43例になり、約三分の二を占める。このパターンは、構成要素として、1次結合の2単位語三つを持つ点で、3単位語と似た構造を持つ。3単位語を、a₁+(a₂+a₃)，(a₁+a₂)+a₃のよう

に書き表わすならば、6単位語の場合は、 $A_1+(A_2+A_3)$ 、 $(A_1+A_2)+A_3$ のように表わすことができる。3単位語で、 $u+U : U+u$ が142:573であるのと同様に、6単位語の場合も、 $U+V : V : U$ が8:35と、比の値がほぼ等しく、前部分が重い型のほうが多く出現している。

$(V+u)+u$ に属するものには、上の「都市交通課長」のように、見かけ上、最後にくるuが $(V+u)$ のuと結合して、 $V+U$ のように見えるものがある。これは、4単位語の $v+u \rightleftharpoons U+U$ の現象と同様の傾向である。

また、 $(v+U)+u$ に属するものには、 $v+v$ と似ているものがある。上にあげた「自動車修理工」のような例である。この類は、5単位語の $V+u \rightleftharpoons U+v$ の現象とまったく同様に説明することができる。すなわち、 $(v+U)+u$ のvが、「自動車」、「小学校」のように、 $U+v$ のUに相当する安定した3単位語で、「修理」、「入学」のような動作性を伴った2単位語の対象になっている場合である。また、「たばこ耕作組合員」のように、uの直前のUが動作性を伴ったものでなくても、 $U+u$ の結合度が高い場合にも、これと同様の傾向がみられる。

4, 5, 6単位語にみえてきた、 $\square+u$ の型のuが、見かけ上、前にくる構成要素と結合しているように見える傾向は、語構造におけるuの性格と関係を持っているように思われる。特に、uが漢語最小単位の場合に、そうした傾向が見られるのは、語構成要素としての漢語最小単位の自立性とかかわりを持つ問題である。すなわち、漢語最小単位は、そのみで単独に用いられることが少なく、他の最小単位と結合することによって、安定した形態を持つことができるものが多いため、 $\square+u$ の型は、もともと不安定な性質を持っている。これは、 $u+\square$ の型においても同様で、むしろ、その不安定性は、 $\square+u$ の場合よりも強いために、 $u+\square$ という型が出現しにくいのであろうということは、前に述べた。 $\square+u$ の場合でも、 \square が高次な結合を持つほど、uは不安定になり、直前のuまたはUと結合して、見かけ上でも、Uあるいはvという形態をとりたがる傾向が、今までに問題としてきた現象としてあらわれたものと考えられる。

5.5 7単位以上の語の構造

7単位以上の語については、出現例も少ないので、一括して、特徴的な例のみをあげることにする。

$u + \square$, $\square + u$ のパターンに属するものは、構成単位数が多くなるほど、現われにくくなる。その理由は、今までに、たびたび述べてきたので、ここでは、くりかえさない。

$u + (U + V)$: 新長期経済計画

$u + (V + U)$: 都児童福祉協会

$(V + U) + u$: 議員手当条例中

$(V + v) + u$: 強風波浪注意報下

$y_q + u$: 都中学校長会長

U を二項式の一方に持つパターンは、もっとも多く、7単位語では、他の一方が、 $v + U$ あるいは $U + v$ である場合が普通である。また、 $\square + U$ に属する語が $U + \square$ に属するものよりも多いことは、他の単位語と同様である。

$U + (v + U)$: 国鉄新幹線総局

$U + (V + u)$: 都立工業奨励館

$U + (V + U)$: 国際航空輸送協会

$(v + U) + U$: 業務上過失傷害

$(U + v) + U$: 経済審議会会長

$(V + U) + U$: 設置反対期成同盟

$(V + v) + U$: 公正取引協議会会長

$\{(V + U) + U\} + U$: 物価安定対策推進本部

そのほかでは、 v を一方に持つものがあるが、それほど多くはない。 V および $V + u$ を結合次数の低いほうに持つものは、それぞれ、1例だけであった。

$v + V$: 代表的航空企業, 核兵器完全禁止

$V + v$: 公営企業委員会, 皮膚紋理研究会

$v + (V + u)$: 警視庁交通規制課

$(V + u) + v$: 姿勢制御用小型ロケット

$V + (u + v)$: 現場調査小委員会

(V+u)+(v+u)：姿勢制御用窒素ガスジェット

これら、多くの構成単位からなる複次結合語には、組織・団体などに関するものが多く見られる。また、一般的な内容を表わす場合にも、臨時的な結合関係にあるものがほとんどである。そして、すでに存在する結合次元の低いパターンの組み合わせからなるもので、ほとんどが占められ、新しいパターンをつくり出す力は、しだいに乏しくなるように思われる。

6. おわりに

はじめにも述べたように、小稿の目的は、語構造を一般式で表わすことによって、その構造上の特徴を明らかにすることにあった。しかしながら、調査データそのものが、量的に不十分だったこともあって、必ずしも、その目的を果たしたとは言いがたい。さらに多くのデータを対象とすることによって、今回の分析の空白部分をうめる必要があると思われる。

また、ここでは、構造のパターン化に焦点をしばったため、構造上の特徴をつくり出す要因となる。語構成要素の性格や、その意味的な結合関係については、問題点の指摘にとどまり、それを、十分に検討することはあえてしなかった。それについては、別の機会に譲ることにしたい。小稿が、現代語の語構造を記述する方法論の一つの試みたりうれば、幸いである。

(47. 12. 1)

〔付記〕

本稿の主旨については、48年2月17日に行なわれた、本研究所の研究発表会でも、「現代漢語の語構造」と題して発表した。その際、本稿の論旨にやや修正を加える必要があることに気づいたが、期限の関係もあり、最小限の手直しにとどめた。いずれ、本稿に示したような方法によって、現代語の語構成について、あらためて述べる機会があると思うので、その折に、修正を含めて、新しい考え方を提示することにした。