

国立国語研究所学術情報リポジトリ

使用頻度“1”の語と文章：
高校『物理』教科書を例に

メタデータ	言語: Japanese 出版者: 公開日: 2016-06-15 キーワード (Ja): キーワード (En): quantitative theory of vocabulary, low frequency words, vocabulary survey, demonstrative, technical terms, nonce formation 作成者: 石井, 正彦, ISHII, Masahiko メールアドレス: 所属:
URL	https://doi.org/10.15084/00001362

使用頻度 “1” の語と文章

— 高校『物理』教科書を例に —

石 井 正 彦

要旨：ある程度の規模の語彙調査を行うと、得られた語彙（語の集合）は少数の高頻度語と大多数の低頻度語とに分離することが知られている。しかし、文章においてなぜ多くの低頻度語が用いられるのかという問いに、いまだ明確な解答は与えられていない。それは、計量語彙論の主要な関心が高頻度語に注がれたためであり、また、より本質的には、低頻度語を語彙論的に特徴づけることが困難であるためである。小稿は、低頻度語の出現（使用）を規定する「機構」は、語彙にではなく、文章の側に存在すると仮定した上で、そのような文章上の「機構」を明らかにするためには、当面、大きな文章の全数語彙調査によって得た低頻度語を対象に、その使用を具体的な文章表現の中で見ていくことが必要であると主張する。この主張の妥当性を確かめるために、国語研究所が実施した「高校教科書の語彙調査」中の『物理』教科書における頻度1の語に注目し、その出現に有意に関連する文章上の語「特徴」を見出す。さらに、それら諸「特徴」をもとに低頻度語の出現を規定する文章上の「機構」を解明する見通しと、そのための課題について述べる。

キーワード：計量語彙論、低頻度語、語彙調査、指示語、専門用語、臨時一語

Abstract : Vocabulary surveys in the past have clarified the general fact that a vocabulary may be divided into two parts, that is, a small number of high frequency words and a large number of low frequency words. However, the answer has not yet been given to the question of why so many low frequency words would be used in a text, because of the difficulties in characterizing the use of low frequency words lexically.

This paper advances the hypothesis that the mechanism of a text governing the appearance or use of low frequency words resides in the text itself, and claims that to explain this mechanism it is necessary: 1) to identify low frequency words through a large scale population vocabulary survey; and 2) to identify the characteristics of a text by studying how low frequency words are used in actual expressions in the text.

In order to substantiate the above hypothesis and methodology, this paper points out some characteristics of a text, based on a high school physics textbook, which result in use of words appearing only once in a text.

Key words : quantitative theory of vocabulary, low frequency words, vocabulary survey, demonstrative, technical terms, nonce formation

I. 問題提起

ある程度の規模の語彙調査を行うと、得られた語彙（語の集合）は少数の高頻度語と大多数の低頻度語とに分離することが知られている^{#1)}。語彙のそのような量的構成は、語の累積使用率分布が「対数型分布」になるとか、語の使用頻度分布が「L字型分布」になるといったことを通して確認されている。なぜ延べ語数の大部分が少数の高頻度語によってカバーされるのか、なぜ異なり語の大部分が頻度1の語をはじめとする低頻度語によって占められるのか、ということは、語彙を量的に解明しようとする計量語彙論の中心的課題の一つである。

そのような問いの前段に、高頻度語とはどのような語であり、また、低頻度語とはどのような語であるのか、という問いが設定できよう。高頻度語については、日本語の文章（談話）を記述する以上、その繰り返しての使用が避けられない一群の語、すなわち（日本語の）「基本語彙」や、その文章の主題にかかわる重要な語として繰り返しての使用が不可避である一群の語、いわゆる「キーワード」などが、その候補として想定され得るだろう^{#2)}。しかし、低頻度語については、そのような候補を想定することは困難であるように思われる。

玉村文郎(1984)は、「専門用語、特殊な職務上の語、特定地域の方言語彙、古語、俗語、隠語など」を「使用の範囲が極端に狭く、したがってまた、使用の回数が一般の人々の間では無視されるほど少ない」語としてあげている（pp. 58-59）が、これは、低頻度語を想定したものではなく、一般の文章に用いられにくい語が存在することを述べたものである。低頻度語とは、一般に用いられにくい語がたまたま用いられたために低頻度になったものではなく、用いられるべくして用いられ、しかも、低頻度であるという語である。そのような語に対して、基本語彙やキーワードと同様の、語彙論的説明を与えることはできていない^{#3)}。

たしかに、低頻度語は文章によって大きく異なり、そのことが低頻度語に

偶発的な印象（たまたま低頻度であったにすぎない）を与えている。実際のところ、いろいろな種類の文章に用いられ、しかも、用いられた場合には常に低頻度となるような“安定した低頻度語”というものを想定することは困難である（低頻度語に対する語彙論的規定が難しい理由もここにある。なお、臨時的な合成語については後述）。しかし、それは、低頻度語が、その語彙論的な性格にではなく、それが用いられた文章の性格に規定されているからだと考えられる。

低頻度語が文章の性格に規定されるとすれば、われわれは、低頻度語とはどのような語か、ではなく、文章においてある語を低頻度語とするのはどのような「機構」であるのか、という問いを設定・提起するべきであろう。

II. 方法論

そのような問いに答えるには、第一に、語を文章から切り離して見るのではなく、具体的な文章表現の中での語の使用を見ていかねばならない。言い換えれば、低頻度という使用実態を「見出し語」の語彙論的な特徴に還元するのではなく、それを支える文章の側の事情に注目するということである。

とはいえ、この方法は、いわゆる語彙論的文章論の方法、すなわち、語彙を指標として文章（ないし文体）の特徴づけを行うという方法とは異なる。目的は、あくまで、語の低頻度の使用という計量語彙論的な主題の解明であり、低頻度語を指標とする文章解析ではない。したがって、低頻度語の文章論的な特徴を平面的に連ねるのではなく、ある語がなぜ低頻度となったのかを説明できるような文章上の「機構」を指摘することが重要である。

ただし、語の使用頻度を文章の側から説明しようとする以上、文章のありようによってさまざまな説明が、とりあえずは、可能であると予想される。したがって、当面は、多種多様な文章を対象とし、そこにおいて見られるであろう低頻度語実現の「機構」を、性急に一般化することなく、一つ一つ検討していくことが必要である。

第二に、注目すべき低頻度語とは一つの語彙において高頻度語と対立するものでなければならず、そのためには、高頻度語と低頻度語とが明確に分離する語彙、すなわち、ある程度規模の大きい文章の語彙を対象とする必要がある。そのような語彙をもつ文章にこそ、低頻度語を実現する「機構」はより明確に認められるものと考ええる。

高頻度語と低頻度語とが分離していないような、規模の小さい文章では、高頻度語に対立するものとしての低頻度語を明確に押さえられない場合がある。そのような場合には、「単に頻度の小さい語」と「低頻度語」との区別があいまいになることもある。たとえば、国語辞書における見出しは、原則としてそれ自身の意味記述文では用いられないから、見出しとその意味記述とをあわせて一つの「文章」と考えれば、その頻度はほぼ確実に1である。しかし、この頻度1の見出しは、その「文章」において、かならずしも、高頻度語に対立するものとしての低頻度語ではない。

規模の小さい文章にも低頻度語は存在するはずであり、それを実現する文章上の「機構」も、また、存在するはずである。しかし、研究の初発の段階では、それらを対象とするには至らない。

第三に、大きな文章における語の使用頻度を調べるためには、語彙調査が不可欠である。しかも、文章における低頻度語の具体的な使用を見るためには、文章をまるごと扱う「全数調査」方式の語彙調査でなければならない。文章を抽出単位としての断片に切り取る「標本抽出調査」は、文章における語の具体的な使用を見る上では適当でない。それは、また、低頻度語を確実に特定する上でも、適当ではない。なぜなら、標本抽出調査は、母集団における語の使用率を、一定の幅（信頼区間）をもたせて、ある確率のもとに推定するのであるが、その際、「使用率の小さい見出し語については標本使用度数の変動が大きく、標本に現れたか否かが相当に偶然に左右される」（国語研（1962），p. 21）からである。付言すれば、標本抽出調査というものは、一般に、くりかえし現れるものをとらえるのに有効であって、低頻度語のように、わずかしか現れないようなものには向かない調査法なのである。

Ⅲ. 問題提起・方法論の妥当性の検証

1. 目的

上述の問題提起と方法論——低頻度語の実現を規定する文章上の「機構」を、まずは大きな文章の全数語彙調査によって得た低頻度語を対象に、その使用を具体的な文章表現の中で見ていくことによって、解明するということ——の妥当性は、なお検討する必要がある。とくに、低頻度語実現の「機構」が文章の側に存在するという見通しの妥当性については、具体的な事例に依拠しつつ、確かめておかなければならない。ただし、現段階ではそのような「機構」を具体的に提示することは困難であるので、その存在を示唆する文章上の「特徴」を見出すことによって、見通しの妥当性をはかることにする。すなわち、上述の方法論にかなう調査を実際に行うことによって、低頻度語の使用に有意に関与する文章上の「特徴」を見出すことができれば、上の見通しに妥当性があるものと判断する。

2. 調査

国語研究所が行った「高校教科書の語彙調査」から『物理』（大塚明郎他『標準高等物理Ⅰ』、講談社、1974）を選び、その本文に現れた頻度1の語——当該教科書中でただ一度しか用いられなかった語——について、その出現に関与する文章上の諸「特徴」を見出す。以下、いくつかの注釈を加える。

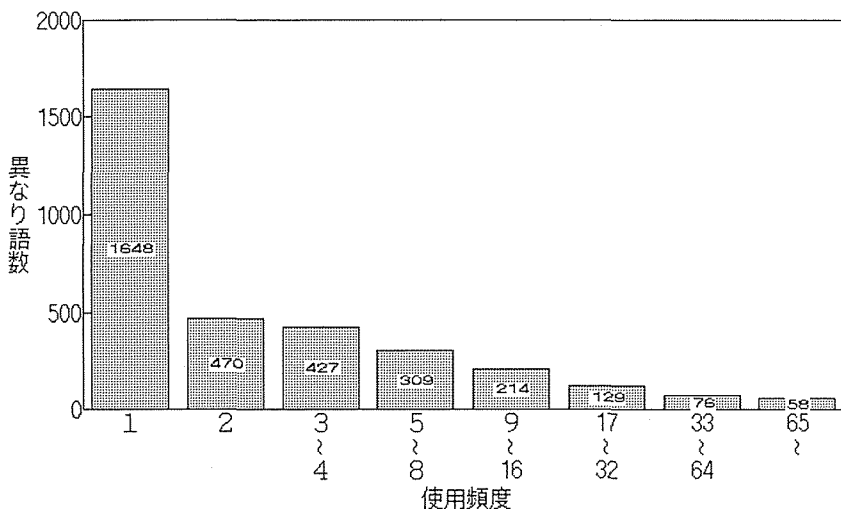
「高校教科書の語彙調査」は、1975年当時の高校の理科・社会科9科目の教科書から各1冊を選び、その本文全文（前書き、[問]などは除く）を対象とした、全数調査としては比較的規模の大きい語彙調査であり、IIで述べた方法論にかなう数少ないデータである^{#4)}。ただし、『物理』を選んだこととくに理由はない。いずれ、すべての科目を対象とし、それらの比較も行いたい。

頻度1の語に注目するのは、後述するように、それが語彙のほぼ半数を占め、また、かならずしも明らかでない低頻度語の認定において、確実に低頻

度であるとみなし得るからである。

何を「語」と認めるかは、「高校教科書の語彙調査」のW単位の規定（国語研（1984），pp. 4-9）に従った。W単位は、いわゆる「長い単位」の一種で、ほぼ文節に相当する（調査の便宜上、数字・記号なども含んでいる）。ただし、ここでは、規定上はW単位である助辞を除いた。また、W単位では、用言の変異形（の一部）がそのまま語彙表の見出し語となっているので、それらを一つの見出し語にまとめた。たとえば、動詞「あげる」は、「あげた」（頻度2）・「あげて」（頻度1）・「あげる」（頻度3）という別々の見出し語となっているが、ここでは、これらを統合して「あげる」（頻度6）とした。したがって、「あげて」を頻度1の語とする、というようなことはしていない。

『物理』全体の語数は、延べ 29781 語、異なり 3331 語である。異なり語の使用頻度分布は、[図1]のようになる。頻度1の語は 1648 語あり、異なり語数全体に占める割合は 49.5%と、ほぼ半数である。



[図1] 『物理』語彙の使用頻度分布

3. 結果

調査により、頻度1の語の出現に関与すると考えられる『物理』の文章上の「特徴」を、いくつか見出すことができた。以下、それらを列挙する。これらの諸「特徴」は、いまだ相互の関連性およびそれらに通底する本質を明らかにし得てはいないものの、頻度1の語の、少なくともあるものについて、それを頻度1たらしめている文章上の「機構」が存在することを示唆するものと解釈できる。その意味において、小稿での問題提起と方法論上の見通しについては、ある程度の妥当性が得られたものとする。

3.1 例示

以下に掲げるのは、『物理』の冒頭部分（I部1章§1 直線運動と速度）である。カッコ内は文の通し番号であり、太字が頻度1の語である（以下同様）。これを見ると、頻度1の語は、その多くが《例示》の文ないし表現に用いられていることがわかる。

(0009) 速さと速度

(0010) 運動のようすを知るうえで、速さは重要な役割を果たしている。

(0011) そのために、自動車や列車の運転台には必ずスピードメーターがついていて、運動のようすを運転者に示すようになっている。

(0012) この計器の多くは、一定時間内の車輪の回転数を、適当な機構で計器面に表わす。

(0013) 車輪の回転数は、その時間内に車が移動した距離を示す。

(0014) つまり、この場合は一定時間内の移動距離を知って、速さをきめる。

(0015) また100 m競走の場合には、一定距離を走るのに要する時間で速さを比較する。

(0016) しかし一般には、単位時間あたりの移動距離で、速さを表わす。

(0017) たとえば、t秒間(s)にl m移動したときの速さを◆○式^(*) [m/s] (I-1) のように表わす。

(0018) [m/s] は、速さを測る単位の一つである。

(0019) 同じように、[cm/s] [k m/時 (km/h)] なども速さの単位になる。

(0020) 2台の車が、同じ地点をどちらも48 k m/hの速さで通過したとしても、

一方が東向き、もう一方は西向きに走ると、ある時間後には両者は全く別の場所に達する。

- (0021) このように運動のようすを表わすときには、速さだけでは不十分で、その運動の向きもいっしょに示す必要がある。
- (0022) そこで、速さと向きをいっしょにした量を考え、これを速度と呼ぶ。
- (0023) 速度の大きさが速さである。
- (0024) 石を真上に投げたときには、はじめ上向きに進んでいた石もやがて下向きに落ちてくる。
- (0025) このように一つの物体の運動でも、速さとともに向きも考えることが必要なのである。
- (0026) 直線運動をしている物体の速度をきめるには、物体が運動している直線上に適当な基準点を取り、これを原点として各瞬間、瞬間の物体の位置を示す座標をしらべればよい。
- (0027) たとえば、時刻 t_1 での物体の位置を x_1 、時刻 t_2 での位置を x_2 とすれば、速度 v は◆○式 (I-2) で表わされる。
- (0028) 直線上の座標は、原点に対し一方の向きを正にきめれば、反対の向きの値は、負の値で表わされる。
- (0029) 同じように直線運動をしている物体の速度は、その向きが座標の正の向きに一致するときは正の値で、その反対向きのものは負の値で示される。
- (0030) 図 I-1 の例で、道路上の標識の位置を座標の原点 O とし、図の右向きを正にとる。
- (0031) いま車 A 、 B が、時刻 t_1 には同じ位置◆○式にいたものが、 20 s 後の時刻 t_2 には、 A は◆○式に、 B は◆○式 (つまり O の左側 280 m) の位置に達したとする。
- (0032) このときの A の速度 v_1 と、 B の速度 v_2 を求めると◆○式 $[\text{m/s}]$ ◆○式 $[\text{m/s}]$ となる。
- (0033) つまり車 A 、 B は同じ速さではあるが、 A は図で右向きに、 B は左向きに運動しているということが示される。

ここで、[表 1] [表 2] は、頻度 2 以上の語について、それが頻度 1 の語と同じ文に現れる (=共起する) 回数と現れない (=共起しない) 回数とを求め、その分布が、頻度 2 以上の語全体についての同様の分布とどの程度異なっているかを、「特化係数」により表したものである^{※6)}。特化係数の値が 1

[表1] 頻度1の語と共起する方に偏って分布する語

語(W単位)	頻度1の語と共起			語(W単位)	頻度1の語と共起		
	する	しない	特化係数		する	しない	特化係数
1 1 s 間	12	0	1.89	48 明らか	18	7	1.36
1 理論*	12	0	1.89	52 はず	10	4	1.35
1 各点	11	0	1.89	52 聞こえる	10	4	1.35
4 ふれる	11	1	1.73	54 間隔	12	5	1.33
4 研究	11	1	1.73	54 電気量*	12	5	1.33
6 ふくむ	27	3	1.70	56 中心	23	10	1.32
7 特に	13	2	1.64	56 達する	14	6	1.32
8 電流	12	2	1.62	56 波長**	14	6	1.32
9 観測者*	11	2	1.60	56 部分	14	6	1.32
9 原子核*	11	2	1.60	56 S	7	3	1.32
9 表面	11	2	1.60	56 ごく	7	3	1.32
12 物質*	35	7	1.57	56 わずか	7	3	1.32
13 1個	9	2	1.54	56 開く	7	3	1.32
13 再び	9	2	1.54	56 耳	7	3	1.32
13 衝突前	9	2	1.54	56 衝突後	7	3	1.32
13 短い	9	2	1.54	56 進む	7	3	1.32
17 原子**	44	11	1.51	56 進行方向	7	3	1.32
17 くりかえす	16	4	1.51	56 測定*	7	3	1.32
17 それ	12	3	1.51	69 O	9	4	1.31
17 差	12	3	1.51	69 関係式*	9	4	1.31
17 粒子*	12	3	1.51	69 普通	9	4	1.31
17 E	8	2	1.51	72 波面**	22	10	1.30
17 事実	8	2	1.51	72 来る	22	10	1.30
17 大きな	8	2	1.51	72 C	11	5	1.30
25 たえば	37	10	1.49	75 さらに	15	7	1.29
26 [29	8	1.48	75 はく	15	7	1.29
26]	29	8	1.48	75 内部	15	7	1.29
26 述べる	11	3	1.48	78 P	17	8	1.28
26 説明する	11	3	1.48	78 描く	17	8	1.28
30 分子*	14	4	1.47	80 N	18	9	1.26
31 測定する*	10	3	1.45	80 面積*	16	8	1.26
32 変位*	13	4	1.44	80 検電器*	10	5	1.26
33 水	19	6	1.43	80 考え方	10	5	1.26
34 ひろがる	9	3	1.42	80 実際	10	5	1.26
34 受けとる	9	3	1.42	80 V	8	4	1.26
34 油滴	9	3	1.42	80 音さ*	8	4	1.26
37 もの	35	12	1.41	80 直接	8	4	1.26
38 電気*	23	8	1.40	80 入れる	8	4	1.26
39 v	14	5	1.39	80 非常になる	8	4	1.26
40 ほぼ	11	4	1.38	90 音源*	15	8	1.23
40 放射線*	11	4	1.38	90 音源*	13	7	1.23
40 理解する	11	4	1.38	92 ため	57	31	1.22
43 中	32	12	1.37	92 電子**	29	16	1.22
43 みなす	8	3	1.37	92 振り子*	11	6	1.22
43 銀電荷*	8	3	1.37	92 新しい	11	6	1.22
43 点電容量**	8	3	1.37	92 帯電する**	11	6	1.22
43 熱容量**	8	3	1.37	97 発生する	18	10	1.21
48 (236	91	1.36	97 用いる	16	9	1.21
48)	236	91	1.36	97 もう	9	5	1.21
48 形	26	10	1.36	97 ストロボ写真	9	5	1.21

[表 2] 頻度 1 の語と共起しない方に偏って分布する語

語(W単位)	頻度 1 の語と共起			語(W単位)	頻度 1 の語と共起		
	する	しない	特化係数		する	しない	特化係数
1 作用**	0	10	0.00	48 単振動**	7	14	0.63
2 静電誘導**	1	10	0.17	48 単振り子**	6	12	0.63
3 反対	2	12	0.27	48 θ	5	10	0.63
4 2	2	11	0.29	48 固有振動数**	5	10	0.63
4 減少する	2	11	0.29	48 車	5	10	0.63
6 a	3	15	0.31	48 動かす	5	10	0.63
6 向く	2	10	0.31	48 内部エネルギー**	5	10	0.63
6 平面内	2	10	0.31	48 つるす	4	8	0.63
9 放物運動*	3	14	0.33	59 重力*	21	40	0.65
10 \blacklozenge \circ \circ	10	45	0.34	59 法則*	13	25	0.65
10 下向き	2	9	0.34	61 運動*	60	112	0.66
10 反作用**	2	9	0.34	62 静止する	9	17	0.65
13 3	2	8	0.38	63 ここ	7	13	0.66
13 干渉**	2	8	0.38	64 学ぶ	6	11	0.67
15 直線運動*	5	19	0.39	64 増す	6	11	0.67
16 ばね	16	55	0.43	64 仕事**	34	62	0.67
17 等速直線運動**	4	14	0.42	64 位相**	5	9	0.67
18 保つ	3	10	0.44	64 鉛直方向	5	9	0.67
18 きまる	7	23	0.44	64 原因	5	9	0.67
20 周期**	5	16	0.45	70 基準	9	16	0.68
21 負	7	21	0.47	71 地面	8	14	0.69
21 つね	4	12	0.47	71 自由落下**	4	7	0.69
21 等加速度直線運動**	4	12	0.47	71 摩擦力**	4	7	0.69
21 力積**	4	12	0.47	74 外	10	17	0.70
21 .	3	9	0.47	74 棒	7	12	0.70
21 原点	3	9	0.47	76 持つ	30	50	0.71
27 運動量**	12	35	0.48	76 移動**	6	10	0.71
28 他	7	20	0.49	76 帯電体*	6	10	0.71
28 電位**	6	17	0.49	76 受ける	17	28	0.71
30 手	6	16	0.51	76 運動エネルギー**	14	23	0.71
30 しまう	3	8	0.51	76 時間*	14	23	0.71
30 軽い	3	8	0.51	82 エネルギー**	42	67	0.73
33 定常波**	5	13	0.52	82 つまり	39	62	0.73
34 波動**	18	46	0.53	82 1	10	16	0.73
35 振動**	16	40	0.54	82 場所	10	16	0.73
35 衝突*	6	15	0.54	82 やはり	5	8	0.73
37 摩擦*	7	17	0.55	87 おもり*	27	42	0.74
37 向き	40	96	0.55	88 抵抗力*	7	11	0.73
39 等速円運動**	3	7	0.57	89 一致する	14	21	0.75
40 もどる	4	9	0.58	89 一つ	10	15	0.75
40 極板間	4	9	0.58	89 あるいは	8	12	0.75
40 座標*	4	9	0.58	89 外部	6	9	0.75
43 和	6	13	0.60	89 床	6	9	0.75
44 たいに	5	11	0.59	89 ともに	4	6	0.75
44 逆向き	5	11	0.59	89 自由	4	6	0.75
46 加速度**	31	66	0.60	89 弱い	4	6	0.75
47 式	20	41	0.62	89 当然	4	6	0.75
48 2つ	13	26	0.63	98 これ	55	81	0.76
48 f	8	16	0.63	99 水平	9	13	0.77
48 Q	7	14	0.63	100 関係	31	44	0.78

よりも大きければ大きいほど、その（頻度 2 以上の）語は、全体の分布傾向と比べて、頻度 1 の語と共起する方に偏り、また、特化係数の値が 1 よりも小さければ小さいほど（0 に近いほど）、頻度 1 の語と共起しない方に偏る傾向をもつということになる。なお、[表 1] [表 2] とともに、頻度 10 以上の語に限り、[表 1] は頻度 1 の語と共起する方に、[表 2] は頻度 1 の語と共起しない方に偏って分布する傾向のあるものを、それぞれ、特化係数の降順に上位 100 語まで掲げた。

さて、[表 1] の特化係数第 25 位「たとえば」は、《例示》の文ないし表現の指標となる副詞であるが、その分布（47 回中 37 回、頻度 1 の語と共起している）は、全体の分布傾向と比べて、頻度 1 の語と共起する方に偏っているといえる。いいかえれば、頻度 1 の語は「たとえば」を含む文に現れやすいということであり、このことから、頻度 1 の語と《例示》とのかかわりが見てとれる（下線筆者、以下同様）。

- (0759) たとえば、バットやラケットで、飛んできたボールを打ち返すときには、短時間ではあるが、その間に複雑に変化する力がボールに対して加えられている。
- (0997) 水以外の物質でも、熱容量は、たとえば銅 mg に比べ 2 mg の銅の熱容量は 2 倍というように、質量に比例した大きさを持つ。
- (1313) たとえば女声のうちでも一番高いソプラノが 300~1000 サイクル/s、男声のうちで一番低いバスが 90~300 サイクル/s 程度である。

一方、[表 3] [表 4] は、W 単位の規定により助辞とされたものについて、頻度 1 の語に下接する回数と、下接しない（頻度 2 以上の語に下接する）回数との分布を求め、[表 1] [表 2] と同様に、その特化係数を計算したものである^{#7)}。[表 3] には頻度 1 の語に下接する方に偏る助辞を、また、[表 4] には下接しない方に偏って分布する助辞を、ともに、頻度 10 以上に限って、上位 10 位まで掲げた。

[表3] 頻度1の語に下接する方に偏って分布する助辞

助辞(W単位)	頻度1の語に下接する			特化係数	助辞(W単位)	頻度1の語に下接しない		
	する	しない	特化係数			する	しない	特化係数
1 など	15	32	3.99	6 で	103	992	1.18	
2 より	3	19	1.70	7 まで	6	59	1.15	
3 や	11	76	1.58	8 に	211	2153	1.12	
4 ならない	1	9	1.25	9 だけ	8	84	1.09	
5 の	343	3239	1.20	10 と	112	1247	1.03	

[表4] 頻度1の語に下接しない方に偏って分布する助辞

助辞(W単位)	頻度1の語に下接する			特化係数	助辞(W単位)	頻度1の語に下接しない		
	する	しない	特化係数			する	しない	特化係数
1 だ	0	31	0.00	6 か	3	62	0.58	
1 ほど	0	25	0.00	7 が	63	1281	0.59	
3 も	14	456	0.37	8 ので	4	68	0.69	
4 は	65	1565	0.50	9 を	115	1608	0.83	
5 な	12	276	0.52	10 へ	1	14	0.83	

[表3]の特化係数第1位は、「たとえば」と同様《例示》の指標となる「など」であり、また、第3位には、同じく《例示》に関係する助辞である「や」がある。これらのことも、また、頻度1の語と《例示》とのかかわりを示唆するものである。

(0684) 腕時計などでは、振り子は利用しないが、一定の大きさのテンプとよばれる部品が、その軸を中心に振動する周期が一定であることを利用している。

(1331) 音も反射することは、建物の中での反響や、こだまなどでよく経験する。

(1489) 金属のほか酸や塩類の水溶液、炭素なども導体である。

もっとも、「など」について見れば、その総数47例に対して頻度1の語に下接しているのは15例であり、絶対数では頻度2以上の語に下接する場合(32例)の方が多い。「など」は、あくまで、助辞全体の分布から見れば頻度1の語に下接する傾向が強い、ということである。しかし、「下接」ではなく

「共起」ということになると、47 例中 31 例が頻度 1 の語を含む文に現れており、「など」と頻度 1 の語とが共起しやすいことをうかがわせる（「や」についても同様）。

- (0398) 空気や水などの流体中で運動するときには、大きな抵抗を受けることはだれでも経験する。
- (1361) 踏切に警笛を鳴らしながら電車が近づき、また遠ざかっていく場合、あるいはこちらですれちがう電車の中で相手の警笛を聞いている場合などには、音の高さが変わって聞こえる。
- (1851) この結晶の密度は、食塩を溶かさない エチルアルコールの中につるして、浮力をはかるなどの方法で測定することができる。

頻度 1 の語が、同一文中において「たとえば」と共起しやすく、また、「など」や「や」の直前に出現する（あるいは共起する）傾向があるということは、頻度 1 の語と《例示》という文章上の「特徴」とが、『物理』においては、強くかかわっていることを示すものであろう。

3.2 結論・まとめ

一方、[表 2] には「つまり」（特化係数第 82 位）があり、また、[表 2] には現れないが、同第 110 位には「したがって」（特化係数 0.80）がある。これらは、「たとえば」が《例示》の文や表現の指標となるように、文章の展開の上で、《結論》あるいは《まとめ》を表す文や表現の指標となる副詞ないし接続詞である²⁸⁾。これらの特化係数が、これらと頻度 1 の語とが共起しにくいということを示すものであるとすれば、『物理』においては、頻度 1 の語は《結論・まとめ》を表す文や表現に現れにくいと考えることができよう。

- (0502) つまり平面運動における加速度は、その物体に作用している力の向きと一致し、その力の大きさに比例した大きさをもつ。
- (1340) つまり音波も干渉して、節が生じているのである。

- (1616) つまり，帯電体を近づけた導体では，帯電体に近い部分には帯電体と異種の，また遠くの部分には同種の電荷が集まる。
- (0202) したがって，少なくとも直線運動の場合には，加速度の向きは力の向きに一致するといつてよい。
- (1077) したがって，物体が外部に対して仕事をし続けるためには，外部からそれに見合う仕事，または熱が物体に与えられていなければならない。
- (1371) したがって，音源の進行方向にいる観測者は，波面の間隔が一様に短くなった波動を，受けとることになる。

頻度 1 の語の出現に関して，《例示》という「特徴」はそれを促進する方向のはたらきに関与し，《結論・まとめ》という「特徴」は，対照的に，それを抑制する方向のはたらきに関与するものと考えられる。

3.3 指示

[表 1] には，カッコ — () ， [] — も見られる。『物理』では，下例のように，教科書中の図やページを《指示》する際にカッコが用いられることが多い。その際，図・ページを指示する語の多くが頻度 1 となる。

- (0182) そのときは，一つの力のベクトルの矢印の先端まで別の力のベクトルを平行移動し，その先端にまた次の力のベクトルを，というように作図をくりかえす (図 I - 15(b))。
- (0520) 2 章，§ 2 (p・18) で，慣性の法則にあたる関係を述べたときに，その法則は，地面または地面に対して等速直線運動をしている物体を基準にして運動を表わした場合に成り立つといった。

『物理』の文章では，このほか，下例のように，説明に要する具体的な概念ないし対象（これらの多くは図中に示されている）を《指示》する場合にも，頻度 1 の語がよく現れるようである。その際，頻度 1 の語は“概念を表す名詞＋記号（式）”の形をとる。

(0094) この値は、上底 v 0 、下底 \diamond \circ 式、高さ T の台形の面積に相当し、図 I-8 (b) のグラフが $0 \sim T$ の間に包む面積に一致する。

(0617) いま 半径 A の円盤の端に 目印 P をつけ、これを周期 T で等速で回転させる。

(0676) つまり、このおもりに作用している力 f は、変位 x の大きさに比例しているが、常に x を減少させる向き、いいかえれば、中心 P O に向かってはたらく。

《指示》に関して注目されるのは、指示語の分布である。[表 1] には第 17 位「それ」があり、[表 2] には第 63 位「ここ」、第 98 位「これ」がある。[表 1] [表 2] に入らなかつたものも含めて、指示語の分布をまとめると、[表 5] のようになる。

[表 5] 頻度 1 の語との共起についての指示語の分布

語(W単位)	頻度1の語と共起		特化係数
	する	しない	
それ	12	3	1.51
それら	6	3	1.26
そこ	23	16	1.11
その	169	133	1.06
ここ	7	13	0.66
これ	55	81	0.76
これら	4	5	0.84
この	270	325	0.86

頻度および特化係数に差はあるものの、いわゆるソ系の指示語は、すべて、頻度 1 の語と共起する場合の方が多く、逆に、コ系の指示語は、すべて、頻度 1 の語とは共起しない場合の方が多くなっている。同じ指示語でありながら、頻度 1 の語との共起という点において、ソ系とコ系とは対照的である。

ただし、このことは、指示語が頻度 1 の語の出現の可否を直接に規定する、ということの意味するわけではない。実際、両者のそのような直接的な関係を見出すことはできていない。ただ、コ系の「これ」「この」については、下

例のように、それらが《結論・まとめ》の文や表現に用いられることが少なくないという現象を介して、頻度 1 の語の出現（の抑制）に關与している可能性を指摘することができる。

(1405) たとえば、同じ振動数の音を 2 個用意し、共鳴箱の口を向かい合わせにしてならべ、一方をたたいて鳴らす (図Ⅲ—30)。

(1406) しばらくして、鳴らしたほうの音を止めてみると、もう一方の音さが鳴っていることがわかる。

(1407) これは相手の音さの出した音に共鳴したのである。

(1649) 導体は、正電荷を多く持てば持つほど、その電位は高くなることが知られている。

(1650) 導体の表面近くの電界も強くなり、したがって、そこまで新しく他の正電荷を運び込むのに必要な仕事量もふえる。

(1651) これは電位が高くなったことを意味する。

(0045) 図 I—2 から直接、速度 v と時間 t のグラフもかける。

(0046) たとえば、 $0 \cdot 1$ s めの位置と $0 \cdot 2$ s めの位置を利用して、その間の速度をきめ、この値を $0 \cdot 1$ s と $0 \cdot 2$ s の中間の t に対する速度として目盛る。

(0047) このようにして作ったのが、図 I—4 のグラフである。

(0153) たとえば、台はかりの皿を手で押せば、はかりがその力を目盛に示す。

(0154) このとき、手は必ず皿が押し返す力を感じる。

(0155) この力は、手が皿に加えた力の反作用に相当する。

3.4 定義・名付け

「これ」「この」は、また、概念を《定義》し、あるいは、《名付け》るための文や表現にも、よく用いられるようである。

(0294) これがこの場合の初速度である。

(0713) このように、時間の経過に関係なく一定に保たれる (◆○式) という量に

注目し、これを運動量とよぶ。

(1781) これは光のエネルギーをもらって飛び出すことができた電子で、光電子とよばれる。

(0044) このような運動を等速直線運動（または等速度運動）という。

(0059) この a を加速度とよぶ。

(0127) このとき、 F_1 、 F_2 を、力 F の成分または分力という。

このような《定義・名付け》の文や表現にも、頻度 1 の語は現れにくいようである。実際、頻度 1 の語は『物理』において最も一般的な《定義・名付け》の形式である「～を～という。」形式の文に現れにくい、という傾向がある。

「～を～という。」形式の文のうち、

頻度 1 の語を含む文 39

頻度 1 の語を含まない文 78

また、専門用語、とくに《定義・名付け》が与えられるような専門用語は頻度 1 の語と同じ文に共起しにくい、という傾向も見られる。[表 1] [表 2] において、語の右肩に*印を付したものは、文部省『学術用語集 物理学編(増訂版)』(1990)にもある語、すなわち、物理学の専門用語である。そのうち、**印のものは、『物理』において、下例のような形で、《定義・名付け》がなされている専門用語である。

(0080) このような運動を、等加速度直線運動という。

(0993) 物体の温度を 1 度高めるのに必要な熱量を、その物体の熱容量といい、カロリー／度 (cal/度) という単位で測る。

(0642) このような振動が単振動とよばれる。

(1160) この値を波動の振動数または周波数とよぶ。

[表1] [表2] におけるそれぞれの内訳は [表6] のようになる。

[表6] 頻度1の語との共起についての専門用語の分布 (100語中)

	[表1] 頻度1の語と共起する 方に偏って分布する語	[表2] 頻度1の語と共起しない 方に偏って分布する語
* 専門用語	18	12
** 《定義・名付け》が行われる 専門用語	7	26

これによれば、専門用語、なかでも、《定義・名付け》がなされるような専門用語は、頻度1の語と共起しない傾向が強い。このことは、頻度1の語が《定義・名付け》の文や表現に現れにくいということにとどまらず、それらが、『物理』の記述においてより重要な概念（これらの多くは《定義・名付け》がなされる専門用語によって表される）と、《定義・名付け》以外の文においても共起しにくい、すなわち、頻度1の語は『物理』の重要な概念を含む文には現れにくい、ということを示唆するものと考えられる。

3.5 注釈

[表1]に見られたカッコは、《指示》だけではなく、《注釈》的な情報を提示する際にも用いられる。しかも、その《注釈》的な情報を表す語は頻度1であることが多い。

(0829) 1 Nの力がはたらいて、その方向に1 m移動したときに、その力のした仕事を1ジュール（記号J）ときめる。

(0956) またばねの強さを示す定数kは、この物体にはたらいている力の内容から考えて◆○式に相当することも確かめられている（P・72 式I-45参照）。

(1481) 習慣で、前者を負電気（陰電気）、後者を正電気（陽電気）と名づける。

たとえば、上例(1481)では、「負電気」「正電気」という用語が、それぞれ、「陰電気」「陽電気」ともいわれることを《注釈》によって示しているが、『物理』では以降の記述において前二者が採用されており、《注釈》として示された後二者は、ここでのみ現れる頻度1の語となっている。

3.6 省略

はじめに提示された語が、語形が長いなどの理由で、その省略形に置き換えられる場合、それ以降の文章では省略形が用いられるために、はじめに提示された語が頻度1となる例が見られる。下例では、最初の文に現れる「力学台車」という語が、以降の文においては単に「台車」とされ、「力学台車」は頻度1の語となっている。

- (0720) 2台の力学台車を、水平な机の上に先端をふれ合わせてならべて置き(図 I-65)、一方の内部に押し込んであるばねが急にのびるときの力で、両者を突き放す。
- (0721) そのときの運動のようすをストロボ写真などでしらべると、台車が離れていくときの速さは、それぞれの質量に反比例していることがわかる。
- (0722) たとえば次ページの図 I-66 は、両者の質量が等しい場合と、質量が2:1の場合を示している。
- (0723) 2台の台車は、おたがいが相手を押しのける力だけで動き出しているから、運動量は保存されるはずである。
- (0724) しかも最初は両者とも静止していたのだから、運動量はどちらも0である。
- (0725) そこで、台車の質量を m_1 、 m_2 、それぞれの反発後の速度を v_1 、 v_2 とすれば、運動量保存の関係は◆○式のように示される。
- (0726) つまり、両者の持つ運動量は、同じ大きさで向きは反対である。

前述の、《指示》に用いられる“概念を表す名詞+記号(式)”という形式は、また、後続する文において省略形によって表されることを、はじめから前提としてつくられた語でもある。下例の「目印P」や「音源S」は、後続の記述において、それぞれ、「P」や「S」とのみ称されることを前提として、

その最初の使用にのみ用意された語である。最初に一度使用されるだけ、ということが頻度 1 につながる。

(1617) いま半径 A の円盤の端に目印 P をつけ、これを周期 T で等速で回転させる。

(1618) この円盤を真横から見たときの P の位置の変化は、図 I-57 右のようになり、図 I-56 (a) のおもりの振動のようすとよく似ている。

(1619) 真横から見たときの P の位置とは、図 I-57 左の円盤上で P の位置を直径 $XX \blacklozenge \bigcirc \bigcirc$ 上に投影した位置に相当する。

(1365) いま音源 S を中心に密の状態を示す波面を描いてみる。

(1366) $1/N_s$ ごとに 1 つの波面が作られるから、最初の波面 W_1 が S を出てから $4/N_s$ 後の状態を考えると、この波面は、 S から 4λ m にまで達している。

(1367) そして音源が静止していれば、その後、 $1/N_s$ ごとに生じた他の波面 W_2 、 W_3 、 W_4 は、 S を中心とする同心円を描いている。

3.7 臨時的な合成語

I では、「いろいろな種類の文章に用いられ、しかも、用いられた場合には常に低頻度となるような“安定した低頻度語”というものを想定することは困難である」とし、「低頻度語は、その語彙論的な性格にではなく、それが用いられた文章の性格に規定される」と述べた。しかし、“安定した低頻度語”というものがまったく想定できないかといえば、そうでもない。

まず、「臨時一語」をあげることができる。「臨時一語」とは、文を構成するそのときにその場限りのものとしてつくられる語である^{*)}から、その本質において、1 回限りのもの、つまり、頻度 1 の語になる「必然性」ないし「可能性」を備えているといえる。たとえば、下例では、「内部エネルギー最低」「イオン発生」が臨時一語であるが、これらは、「内部のエネルギーが最低である」「イオンの発生」といった語の連続による表現を、文構成にあたって、臨時に一語化したものである。

- (1063) これに対して、内部エネルギー最低の状態を0度ときめた目盛を、絶対温度目盛という。
- (1901) 電極の間には、放電が始まる直前の電圧がかけてあるので、イオン発生が原因となって、短時間、放電が起こる。

以下、(a)は、『物理』において、頻度1であった臨時一語であり^{註10}、(b)は頻度2以上であった臨時一語（カッコ内は頻度）である。臨時一語の多くが実際に頻度1であることが確認できる。

- (a) イオン発生，一定距離，一定数，一定量，移動方向，運動物体，鉛直下方，鉛直上方，大型水槽，大きさ一定，温度上昇，回転部分，火山爆発，加速度・力とも，管口付近，近代科学成立，高温物体，構成粒子，撮影時刻，作用時間，参考図，時間的变化，時間範囲，斜面方向，ジュール自身，衝突直後，衝突直前，振動方向，垂直上方，垂直方向，水平・鉛直両方向，静止状態，絶縁物内部，接触面積，測定結果，測定方法，単振動一般，直線グラフ，直角上方，低温物体，同一地点，内部エネルギー最低，半径方向，反対方向，ピストンつき，物質構成，物質探究，落下距離，理論的計算
- (b) 運動方向(5)，鉛直方向(14)，水平・鉛直方向(4)，直角方向(2)，導体内部(2)，入射方向(2)，反射方向(2)，物質内部(3)

臨時的といえば、派生語の中にも臨時的なものが多い。以下に、『物理』において頻度1であった派生語のうち、生産的であったものについて、接辞(補助用言的要素も含む)ごとに示す。これらの中には、すでに固定的な派生語となっているものもあるが、多くは、臨時一語と同様に、臨時的な性格をもつものである。

各～：－時刻，－長方形，－力

全～：－宇宙，－質量

同～：－時刻，－振動数

～方：いい－，うすめ－，変わり－，きめ－，組み合わせ－，ずれ－，でき－，とり－，のばし－，混じり－，やり－

～間：原子－，帯電体－，等電位面－，分子原子－，両者－

～後：1周期－，運動開始－，時間－，時間 t 1－，20 s－，反発－

～ごと：1周期－，各基本単位－，時間－， t s－，電圧－， $1/20$ s－， $1/8$ －，物質－，方向－， $1/4$ 波長－

～上：一線－，一直線－，A－，延長－，円盤－，球面－，曲線－，金属－，グラフ－，コンデンサー－，C－，CD－，実用－，図 I-63－，ストロボ写真－，線－，台－，帯電体－，直径 XX ◆○○－，導体－，等電位面－，道路－，波面－，みかけ－，理論－，レール－

～状：板－，液－，鎖－，正弦波－，同心円－，放射－

～中：運動－，液体－，円筒－，気体－，金属－，原子－，固体－，作業－，図－，接触－，大気－，導体－，流体－

～的：一般－，エネルギー－，化学－，基本－，客観－，近似－，三次元－，時間－，質－，実験－，実証－，集中－，相対－，対称－，統一－，凸レンズ－，爆発－，物理－，本質－，模式－，理想－

～どうし：かたまり－，球－，極－，極板－，氷－，等電位面－，波－，波面－，腹－，物質－

～内：暗室－，鉛直面－，金属円筒－，空間－，系－，コンデンサー－，細胞－，シリンダー－，組織－，単位時間－，電球－，電池－，導線－，等電位面－，ピストン－，範囲－，物質－，物体－，ブラウン管－，丸底フラスコ－，容器－

～面：陰極－，基準－，金属－，計器－，フィルム－

～にくい：上がり－，起こり－，確かめ－，伝え－

～やすい：気化し－，比べ－，しらべ－

～合う：干渉し－，しりぞけ－，接触し－，強め－，反発し－

～続ける：動かし－，動き－，運動し－，加わり－，作用し－，進行し－，析出し－

～はじめる：圧縮し－，動き－，運動し－，描き－，自由落下し－，すべり－，とられ－，測り－，振れ－，減り－，利用され－

臨時一語や臨時的な派生語は、たしかに一面では、“安定した低頻度語”と見ることができる。これらの語は、その成り立ちがもつ臨時性において、すでに低頻度語としての必然性ないし可能性を備えていると考えられるからである。しかしながら、これらの臨時的な合成語がつけられるのは、つねに、実際の文章においてである。具体的な文章の、具体的な文構成の場が、これらの語を形成する場である。臨時的な合成語は、文章と離れて、いわゆるレキシコンとしての語彙の中にあらかじめ（できあいのものとして）存在しているわけではない。語彙が用意しているのは、臨時的な合成語の構成様式（造語法）であろう。その意味で、臨時的な合成語は、語彙のあり方に規定されつつも、より直接的には、文章のあり方に規定されているものと考えることができる。

IV. 課題

1. 「特徴」から「機構」へ

上に列挙した、頻度1の語の出現に關与する文章上の「特徴」は、低頻度語の出現を規定する文章上の「機構」の存在を示唆するものではあるが、いまだ「機構」そのものとはいえない。「機構」は、これらの諸「特徴」がどのような理由によって頻度1の語の出現にかかわっているのかを説明し得てはじめて、見えてくるものであろう。そのような作業が、より多くの調査と吟味を通して、行われなければならない。ただし、現段階でも、「特徴」から「機構」に至るいくつかの見通しを述べることはできる。たとえば、頻度1の語の出現に關する《例示》と《結論・まとめ》との対照的なかわり方については、以下のように考えることもできよう。

『物理』は、物理学によって明らかにされた物理的諸法則について説明する文章である。《結論・まとめ》は、そのような説明の骨組となる文や表現であるが、説明は諸法則相互の関連性や階層性（一つの法則を前提として他の法則を説明する）にもとづいて展開されるので、《結論・まとめ》に用いられ

る語（ここに重要な概念を表す専門用語も加わってくる）は、結果として、繰り返し用いられる、すなわち、頻度2以上の語となる可能性が高くなるものと考えられる。一方、物理的諸法則の説明を《結論・まとめ》の文・表現だけで行おうとすると、記述が抽象的・概括的なものになってしまう。そこで、そのような説明をわかりやすく、また、説得力のあるものにするために、それらの諸法則が反映する具体的な事実を《例示》として示すことになる。しかも、法則の普遍性を示すためには、同じ事実を繰り返して《例示》とするよりも、異なる多様な事実を《例示》とした方が効果的である。その結果として、《例示》に用いられる語は、一回限りのものとなる可能性が高くなるのではないだろうか。《例示》のもつ具体性・多様性にもとづく「一回性」が頻度1の語を生み出すのではないか、ということである（冒頭部分の例（pp.30-31）でも、「列車」「100 m競走」「標識」などの語に、このような傾向を見ることができる）。

また、頻度1の語との共起について見られた、ソ系の指示語とコ系の指示語との対照的な分布についても、正保勇（1981）の「情報の焦点となるものは、『コ』で指示されることが多い」「焦点でない部分は『ソ』で指示するのが普通である」（pp. 94-109）、金水敏・田窪行則（1990）の「『解説のコ』が典型的に現われるまとまった内容についての解説というのは、まさしく聞き手に対しての内容の把握、情報量などの点において優位に立った発話なのである」（p. 140）などの指摘が参考になろう。すなわち、コ系の指示語が用いられるような、つまり、書き手が焦点を当てて解説しているような部分には頻度1の語が現れにくく、ソ系の指示語が用いられるような、つまり、焦点の当てられていない部分には頻度1の語が現れやすいということが、『物理』教科書のような「解説」的な文章にはあるのかもしれない。

今後、これらの見通しをもとに、頻度1の語の出現を規定する「機構」の解明が必要となる。

2. 文章の独自性

上に列挙した、頻度1の語の出現に關与する文章上の「特徴」は、また、すべて『物理』という文章のあり方に依存しているということを強調しておかなければならない。たとえば、《例示》という「特徴」は、同じ高校教科書であっても、『日本史』『世界史』などの文章にはほとんど見られないものであるし、臨時一語の使用状況も『物理』という文章に規定されている²¹¹⁾。さらに、教科書以外の文章を対象とすれば、ここで見たのとはまったく違う諸「特徴」を見出すことにもなろう。頻度1の語と文章とのかかわりを考える場合、その文章の特徴や独自性といったものを無視することはできないのである。だからこそ、まずは、さまざまな文章について、同様の調査を行っていくことが必要である。

3. 「一つの文章」とは

語の使用頻度とは、語の現れた文章がひとまとまりのもの、一つの文章であることを前提として数えられるものである。『物理』における頻度1の語も、また、頻度2以上の語も、『物理』が一つの文章であることを前提に認定されている。もし、『物理』がいくつかの異なる文章から成るのだとすれば、頻度1の語は、それぞれの文章ごとに認定されるから、今よりももっと増えることになる。小稿では、具体的な吟味をすることなく『物理』を「一つの文章」とみなして、頻度1の語について検討してきたが、以下、この点について、やや詳しく考えてみたい。

3.1 内容の独立性

まず、『物理』の内容を、その「目次」から引用して掲げる。全体が大きく4部に分かれ(I~IV)、それぞれがさらに、各章、各セクション(§)に分かれている。

I 運動と力

1 章 直線運動

§ 1 直線運動と速度

§ 2 直線運動の加速度

2 章 直線運動と力

§ 1 作用と反作用

§ 4 物理量とディメンション

§ 2 運動と慣性

§ 5 落下運動

§ 3 直線運動と運動の方程式

§ 6 摩擦や抵抗のあるときの運動

3 章 いろいろな運動

§ 1 放物運動

§ 4 円運動

§ 2 平面運動と運動の法則

§ 5 単振動

§ 3 座標系

§ 6 単振り子の運動

4 章 衝突と運動量

§ 1 運動量保存の法則

§ 3 斜衝突と運動量

§ 2 運動量と運動の法則

II エネルギー

1 章 仕事と力学的エネルギー

§ 1 仕事

§ 4 力学的エネルギー保存の法則

§ 2 重力による位置エネルギー

§ 5 いろいろな運動とエネルギー

§ 3 運動エネルギー

2 章 熱とエネルギー

§ 1 熱量と温度

§ 3 エネルギー保存の法則

§ 2 熱と仕事

III 波動

1 章 波動の特性

§ 1 波動とその伝わり方

§ 2 波動の干渉

2 章 音

§ 1 音波

§ 3 発音体

§ 2 音の伝わり方

IV 電界と電子

1 章 電界

§ 1 電荷と電気力

§ 3 電界中の導体

§ 2 電界

§ 4 コンデンサー (蓄電器)

2 章 電子と原子

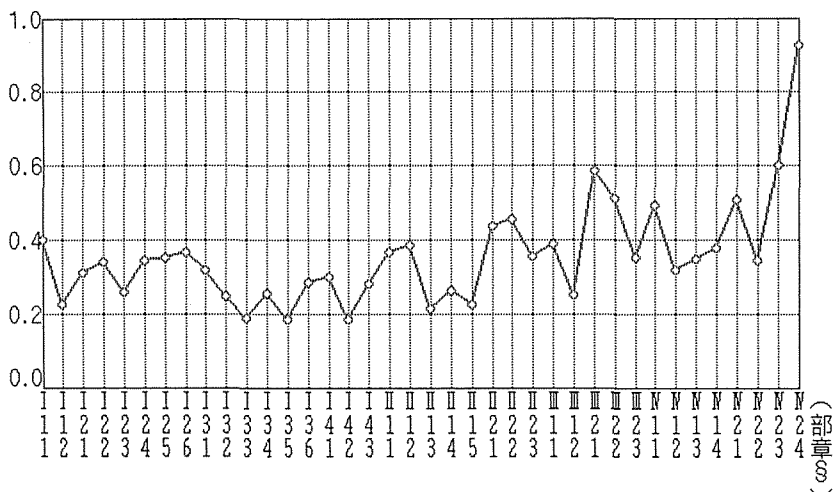
§ 1 陰極線と電子

§ 3 原子

§ 2 電気素量

§ 4 物質探究の足跡

〔図2〕は、『物理』の各セッションごとに、延べ語数に対する頻度1の語の比率を求め、それを目次に従って並べることにより、その変動をみたものである。これによれば、先頭からII部1章までの比率に比べて、それ以降の比率が高くなり、とくにIV部2章後半の比率がかなり高くなっている。



〔図2〕『物理』における頻度1の語の比率の変動

『物理』のように、部・章・セッションが並べられた形式の文章では、それぞれの間の関連性が大きければ、全体として、頻度1の語の出現可能性は小さくなり、逆に、相互の関連性が小さければ、頻度1の語の出現可能性は大きくなるのではないかと考えられる。この関連性（のなさ）を内容上の独立性とおきかえれば、それは、語彙の類似度によっておしはかることができる。〔表7〕は、『物理』の各章間の類似度を「宮島の類似度」（宮島達夫(1970)）によって計算したものである。

[表7] 『物理』各章の語彙の類似度

部-章		0.650以上									
1-1											
0.661	1-2										
.683	.718	1-3									
.646	.689	.684	1-4								
.629	.703	.687	.654	II-1							
.572	.631	.614	.602	.648	II-2						
.604	.612	.624	.615	.604	.598	III-1					
.574	.612	.613	.593	.602	.589	.665	III-2				
.597	.658	.635	.637	.629	.614	.617	.612	IV-1			
.576	.627	.618	.599	.600	.620	.596	.601	.640	IV-2		

これによれば、I部1章からII部1章までの各章間の類似度が他に比べて高いことがよみとれる。この傾向は、[図2]における傾向——I部1章からII部1章までの「頻度1の語の比率」が小さい——と符合するものであり、類似度と頻度1の語の比率との間に負の相関があること、すなわち、内容の独立性が高い文章（の部分）ほど頻度1の語の出現する可能性が高い、ということを示唆するものである（[表1]の第1位に「理論」、第4位に「研究」という語があるのも、前者では12例中11例、後者では12例すべてがIV部で用いられていることによるものと考えられる）。

3.2 文体の相違

IV部2章の後半、とくに「§4 物質探求の足跡」は、また、その内容がそれまでの各章・セクションと異なっているとともに、記述の文体も歴史教科書風であるように思われる。そして、そのような文体の違いが頻度1の語を生み出す背景となっているようにも思われる。たとえば、以下の例における「みうけられる」（1877）、「いよいよ」（1880）、「助けられて」（1883）などの語は、ここでのそのような文体に依存する性格が強いものではないだろうか。

このことは、ある文章ないしその部分の文体が、他の「大部分」の文章ないしその部分のそれと、大きく異なっていれば、当該の文章ないし部分に頻度1の語が現れやすくなる、ということを示唆するものと考えられる。

(1876) §4 物質探究の足跡

- (1877) 物質が、ごくわずかな種類の粒子によって組み立てられているのではないかという考え方は、遠くギリシアの時代にもみうけられる。
- (1878) しかしそれは、当時の、単に観念の上の産物であり、実証的に裏付けられた原子論は、他の科学と同じように、17世紀以後に誕生した。
- (1879) 化学変化の過程を追求した結果から、19世紀の初めにはドルトン、アボガドロ、ベルセリウスたちの力で、ほぼ今日の考えに近い原子を物質の基本単位とする考えがまとめ上げられた。
- (1880) さらに19世紀の半ばになり、原子の性質がメンデレーフの周期表の形で整理されてみると、原子の存在は、いよいよ 確実と思われるようになった。
- (1881) 物理学的な研究の面でも、物体の運動や物質の状態変化というような外に現われた現象の法則性を追求する研究方法に対して、物質内部を作る分子あるいは原子を予想して、その性質や振舞いから現象を説明しようとする方法もとられはじめた。
- (1882) 熱現象を分子や原子の運動で説明するのもその一例である。
- (1883) そして19世紀末から20世紀にはいると、単に理論上のものとして仮想するだけではなく、直接そのような構成粒子を取り出して研究しようとする方向が、理論や実験技術の進歩に助けられて 進みだした。
- (1884) 前の節で述べた陰極線の研究からとらえられた電子は、このような成果の最初のものといえることができる。

一つの文章において、ある部分（文章全体から見るとごく一部分）の内容や文体が他の大部分と異なっている場合、そこに頻度1の語が現れやすいのだとすれば、このことは、内容・文体を異にする別々の文章の間にもいえることであろう。言い換えれば、『物理』において、IV部2章（後半）は別の文章というべきなのかもしれない。しかしながら、内容・文体が独立していればすべて異なる文章というべきかといえ、かならずしも、そうともいえないであろう。何を「一つの文章」と考えるべきか、その認定基準とはどのよ

うなものであるべきか、文章における頻度1の語および低頻度語の使用を追究する上で、明らかにしなければならない課題である。

[注]

- 1 田中章夫(1978), p. 34。「低頻度語」「高頻度語」という用語もこれによる。
- 2 キーワードについては、かならずしも高頻度語に多くないという指摘がある。水谷静夫(1983), p. 169 など参照。
- 3 田島毓堂(1992)は、「語彙分析」における「最低頻度語の扱い」について言及している(pp. 5-6)。
- 4 土屋信一(1989)は、同調査が全数調査を採用したことについて、次のように述べる。

これは専門知識体系を記述するための語彙を把握するという観点から採用されたものである。従来のように、ブロックサンプリング法で、ランダムにデータを取り出したのでは、知識体系の記述状態を解明することはできないのである。(中略)

全数調査を採った理由は他にもある。文章の流れの中で語彙の時系列分析をするため、および全数データに対してさまざまなサンプリング法でデータを取り出して、サンプリング法の在り方を検討するためである。

最後に、この教科書データが日本語研究のためのデータベースのひとつとして、今後十分に活用されることを考えた。高校理科・社会科教科書の文章が、9冊分そっくり、等質の単位に切られ、読み仮名・品詞情報・意味番号等が添えられているということは、今後さまざまな実験に堪えるデータたり得ることを示している。(p. 2)

- 5 データ中の“◆○式”“◆○○”は、それぞれ、“式”“JIS規格にない記号”を表す。国語研(1984), p. 27 参照。
- 6 上田尚一(1981)によれば、「特化係数」とは、全体についての構成比を基準にして各語についての構成比をその何倍にあたるかという形で表現したもの(p. 33)である。『物理』において頻度2以上であった語全体についてみれば、頻度1の語と共起するものが延べ14841語、共起しないものが延べ13292語であり、両者の構成比は0.53対0.47である。したがって、頻度2以上の各語の特化係数は、各語ごとに全体と同様の構成比を求め、頻度1の語と共起する場合の比率を全体の比率0.53で除したものの、また、共起しない場合の比率を全体の比率0.47で除し

たものを求めればよい。ただし、この場合は、共起するかしないかの二通りしかないので、一方の特化係数で各語の全体からの隔たりを代表させることができる。小稿の各表では、共起する場合の特化係数のみを示した。

- 7 『物理』に用いられた助辞は延べ14844であり(異なりは36)、頻度1の語に下接する場合と下接しない場合との構成比は0.08対0.92である。
- 8 「つまり」「したがって」には、《いいかえ》の用法もある。

(0079)このグラフが直線になることから、図I-6のような斜面上の物体の運動では、速度の変化する割合、つまり加速度の大きさは一定であることがわかる。

(1210)このとき、両側の媒質の性質がきまれば、それぞれの側での波動の速さ、したがって波長もきまり、境界面の両側での波長の比も一定になる。

- 9 「臨時一語」という名称は林四郎(1982)によるが、その定義は小稿で述べたものよりも広いので注意されたい。
- 10 臨時一語の認定基準は石井(1993)による。
- 11 石井(1993)は、臨時一語の出現が文章の種類によって異なることを計量的に確かめている。

[文献]

- 石井正彦(1993)「臨時一語と文章の凝縮」(『国語学』173)
- 上田尚一(1981)『統計データの見方・使い方』(朝倉書店)
- 金水敏・田窪行則(1990)「談話管理理論からみた日本語の指示詞」(『認知科学の発展』3)
- 国語研(1962)『現代雑誌九十種の用語用字(第一分冊)総記及び語彙表』(秀英出版)
- 国語研(1984)『高校教科書の語彙調査II』(秀英出版)
- 正保勇(1981)『「コソア」の体系』(『日本語教育指導参考書8・日本語の指示詞』, 国立国語研究所)
- 田島毓堂(1992)「語彙論的語の単位——意味単位と分類単位と——」(『文化言語学その提言と建設』, 三省堂)
- 田中章夫(1978)『国語語彙論』(明治書院)
- 玉村文郎(1984)『日本語教育指導参考書12 語彙の研究と教育(上)』(国立国語研究所)
- 土屋信一(1989)「語彙調査——全体的な見通しとねらい——」(国立国語研究所『高

校中学校教科書の語彙調査 分析編』, 秀英出版)

林四郎 (1982) 「臨時一語の構造」(『国語学』131, 林四郎『漢字・語彙・文章の研究へ』(明治書院)に再録)

水谷静夫 (1983) 『朝倉日本語新講座 2 語彙』(朝倉書店)

宮島達夫 (1970) 「語いの類似度」(『国語学』82)

[付記]

小稿を成すにあたり, 青葉ことばの会(第41回, 1994年7月23日), 東北大学国語学研究会(第273回, 1995年11月10日), および, 所内研究部会議(1995年11月22日)において発表し, 多くの助言を得た。ここに記して感謝申し上げる。