日本語 2 字綴りの連想価と出現頻度の関係について*

—京都大学 今 栄 国 晴**—

学習実験における言語材料の特性として、意味性、熟知度、感情性などが、取り扱われてきたが、言語材料の母国語における出現頻度を実験変数とした実験も数多い。

従来、意味性、熟知度、出現頻度の 3 特性は、それぞれ、非常に高い相関があることが英語においてあきらかになっていた（これらについては、森川(3)、Underwood(5) がくわしく論じている）。しかし、筆者(1) は、日本語 2 字綴り (digram) 40,000 綴りの出現頻度を数えたが、それらと、梅本ら(4) の無連想価、有意味度との相関は、それぞれ、 -0.249 、 0.281 という有意ではあるが、あまり高いとは言えない結果を得た。

このように相関が低かったのは、両特性の分布が異なっている（出現頻度の分布は L 型）ことも原因であろうが、それにしても、このことは、日本語においては、意味性および出現頻度という言語材料の 2 つの特性が、英語ほどは、密接な関係にないことを示すものであろう。

本研究は、日本語 2 字綴りの出現頻度という、いわば、言語材料の非心理学的測度が、意味性という心理学的測度と、どのような関係にあるのかを研究することを目的とする。すなわち、日本語においては、頻度は、意味性と相関が低いけれども、それでも、なお、頻度は、心理学的に意味ある特性なのか、あるいは、その特性は、意味性のもつ特性から分離できるものであるかどうかを研究した。なお、本研究においては、意味性は、従来、もっとも、よく使われている梅本ら(4) の無連想価をもって代表させた（無連想価と有意味度は、非常に相関が高い [$r = -0.928$] ので、一方での結論は、他方にも適用できよう）。

本研究は、2 つの実験より成立している。第 1 実験は、近似文実験で、Miller et al(2) に始まる近似文実験は、出現頻度が主変数であったが、これに、連想価を変数として追加し、両変数の関係を調べ、第 2 実験の対連合実験では、従来、連想価が主変数であったのに対して、出現頻度を変数として導入したのである。本研究においては、この 2 つの実験は、ともに、出現頻度が、意味性とは別個に学習を規定する特性であることを示し

た。なお、実験は、1960 年 10 月から 12 月の間におこなわれた。

実 験

実験 I 近似文再生実験

目的 従来の近似文実験に用いられた近似文は、出現頻度のみを変数として構成されているが、そこに、連想価などの心理学的測度が、どのように介入しているかは、ほとんど省みられていない。そこで、今まで漠然と考えられていた近似文の中から、2 字綴りからできている 2 次近似文（ある文字の出現が、直前の 1 文字のみによって確率的に規定されている文字系列）を選び、それを頻度と連想価の高低から 4 種作製し、その直後再生を調べ、両特性の関係を調べた。

使用リスト 近似文における頻度は、文字間の遷移確率のことである。すなわち、日本語清音において、ある文字は、次にくる文字として、44 字の可能性をもち、それぞれの文字への遷移確率は異なるのであるが、その 44 個の確率を高確率順に並べて、上位第 1 四分偏差にあたる 12 綴りを高頻度 (HF) とし、第 4 四分偏差にあたる 12 綴りを低頻度 (LF) とした。なお、確率 0 はすべて、LF とした。また、連想価は、梅本ら(4) の表からとり、無連想価 0—19 の綴り（計 932 個）を高連想価綴り (HA) とし、無連想価 20 以上（計 960 個）を低連想価綴り (LA) とした。

こうして、頻度 (F)、連想価 (A) を高低に分けた後、次の 4 種（1 種につき、4 リスト）の近似文（長さ 10 字）を作った。

1. HF-HA 系列

ムシカツウスマルクレ
モタマチイシカレツヨ
クシラカイコトキミタ
ツクリカレイサキナシ

2. HF-LA 系列

ラタエルメセヨテアノ
ニカタリトエソシコノ
マシテナリニモワノイ
テヨサセンコノニネア

3. LF-HA 系列

カスフセムツネキユミ
スロラツヤミコフワマ
コメロキヨモメチワニ
コソネットソセメロキヌ

4. LF-LA 系列

シフ라우ヌナテヒヨヤ
ノスニミヒオネユシフ
ユヨヘニムホヤヒオヌ
ウノホナヘロニラオミ

たとえば、上記、HF-HA 系列のモタマチイシカレツヨという近似文は、次のように構成されている。すなわち、

遷移確率	.028	.044	.049	.071	.069	
	モ	タ	マ	チ	イ	シ………
無連想価	16	0	0	6	2	

これに対し、たとえば、LF-HA 系列は次のような構

* On the Relationship between the Association Value and the Frequency of Occurrence in Japanese Digrams.

** Kuniharu Imae (Kyoto University).

Table 1 直後再生された2字綴りの数

	HF	LF	平均
HA	6.56	4.44	5.50
LA	5.75	2.53	4.14
平均	6.13	3.49	

Table 2 Table 1 の分散分析表

	SS	df	ms	F
頻度(F)	57.11	1	57.11	70.26***
無連想価(A)	14.79	1	14.79	18.20**
被験者(S)	55.67	7	7.95	9.78*
F×A	2.39	1	2.39	2.95
F×S	14.78	7	2.11	2.60
A×S	16.10	7	2.30	2.84
誤差	5.69	7	0.81	
計	166.53	31		

*** $p < 0.001$ ** $p < 0.01$ * $p < 0.05$

成になっている。

遷移確率	0	0	0	0	0
無連想価	1	2	16	9	15

すなわち、HF-HA 系列と LF-HA 系列は、無連想価に関しては、両系列はともに 19 以下であり、高連想であるのに対し、遷移確率に関しては、HF-HA 系列では、いずれも、0 より相当大であり、LF-HA 系列では、遷移確率が 0 であるか、または、0 に非常に近いという点で両系列は異なっている。すなわち、もし、両系列に学習上の差があれば、それは、連想価の差ではなくて、遷移確率の差であると推察されるのである。同様に、HF-HA 系列と HF-LA 系列の差は、遷移確率ではなく、連想価の差であろうと考えられるであろう。

手続 メモリードラム使用。1秒に1字提示し、1リスト(10字)が終ると、直ちに、系列順序に注意して再生させる。被験者は、1回の来室において、1種4リストをおこない、計4回来室する。来室の間隔は、少なくとも、1週間とし、前経験の妨害を除くようにした。被験者は、大学生8名。

採点は、系列中の2字綴り1つが再生された場合1点とした。1リストは10字からできているから、満点は、9点である。

4種の系列の実験順序は、被験者4人を1群とするラテン方格のくり返しの形をとっておこない、各種系列の順序の偏りを除くようにした(ただし、結果の分析は、ラテン方格は考慮に入れていない)。

結果 各系列における直後再生された2字綴りの平均数を Table 1 に示す。Table 1 の見方は、たとえば、HF-LA 系列の平均再生数は 5.75 というように見る。こうしてみると、最大再生数は、HF-HA 系列であり、最小再生数は、LF-LA 系列であることがわかる。さら

に、Table 1 を分散分析したものが Table 2 である。

Table 1 および Table 2 から、頻度、連想価の2つの特性は、ともに、学習に対して有意な効果をもつことがわかった。しかも、その効果は、一方が他方を含むという——例えば、連想価の効果は頻度の効果の中に含まれて、その一部である——のではなく、連想価および頻度は、互いに、独立な特性と見るべきであろう。

また、本実験は、従来、漠然としていて、精密な学習実験での使用が困難視されていた近似度という概念も、連想価を導入することにより、一層、精密な分類が可能となることを暗示していると思われる。

実験 II 対連合学習実験

目的 従来、対連合における言語材料は、意味性の変化が主であり、特に我国においては、連想価の変化がほとんどであった。これに対して、本実験は、頻度を変数として、導入することにより、出現頻度の心理学的特性を知ることとする。

さらに、実験 I で取り扱われた頻度は、遷移確率であったが、本実験では、出現頻度を生の形で取り扱う結合確率を問題にする。

また、リストの学習速度を決定する要因は、刺激語よりも反応語の方が強いことがわかっており(これについては、森川(3)を参照)、梅本ら(5)も、それにならっている。本研究においても、条件変化は反応語についてのみおこない、刺激語は、4リストを通じて、同一アルファベットの組を用いた。

使用リスト

A. 出現頻度について

1. 低頻度綴り (LF): 0 頻度 836 綴り
2. 高頻度綴り (HF): 40,000 綴り中、出現頻度 9 以上の綴り 275 個

以上の基準で選んだ。しかし、HF については、実際に使用したのは、頻度 17 以上のもののみである。

B. 連想価について

1. 高連想価綴り (HA): 無連想価 0—4 の綴り 314 個
2. 低連想価綴り (LA): 無連想価 20—29 の綴り 285 個

次に、上述の基準で、出現頻度の高低および連想価の高低を組み合わせて作った4種のリストを Table 3 に一覧する。なお、1つのリスト内では、同一文字が2回以上出ないようにして、リスト内類似性を低めた。また、各リストは、提示時には、4つの異なった順序で各被験者に提示された。

手続 適中予言法。1回1.5秒提示。学習基準は、1回完全正答。被験者は、1回の来室において、1リストをおこなう。各来室の間隔は、1週間以上。被験者は、大学生8名。

結果 4種のリストのそれぞれ完全学習に要した平均

Table 3 使用リスト一覧

HF-HA リスト				HF-LA リスト			
対	頻 度	無連想価	10試行迄の 平均正答数	対	頻 度	無連想価	10試行迄の 平均正答数
B-ヒ ト	61	0	7.25	X-ソ レ	61	20	4.75
P-イ カ	57	1	4.00	P-コ ノ	60	24	6.50
Q-ワ タ	51	1	6.38	Q-テ イ	58	23	5.25
Y-コ ロ	43	3	4.38	V-タ シ	55	29	4.63
V-オ モ	37	3	6.13	Y-キ ニ	23	20	5.13
X-ヤ ク	25	3	5.00	Z-エ ル	20	24	5.88
Z-シ マ	24	1	6.75	B-セ ツ	17	20	3.88
平 均	42.6	1.7	5.70	平 均	42.0	22.9	5.14

HF-LA リスト				LF-LA リスト			
対	頻 度	無連想価	10試行迄の 平均正答数	対	頻 度	無連想価	10試行迄の 平均正答数
B-タ ネ	0	0	4.50	B-ス ツ	0	20	4.75
X-ヤ ミ	0	1	4.00	Z-テ ト	0	20	3.75
V-ハ ト	0	1	5.63	V-ホ エ	0	20	5.00
P-ホ ル	0	3	6.25	P-ニ ス	0	23	4.25
Y-ケ サ	0	1	5.25	Y-ナ コ	0	24	2.50
Q-ニ コ	0	4	7.13	Q-メ ラ	0	24	4.75
Z-ヌ レ	0	2	5.00	X-ケ マ	0	20	3.00
平 均	0	1.7	5.39	平 均	0	22.9	4.00

Table 4 各リストの平均試行数

	HF	LF	平 均
HA	10.50	11.13	10.81
LA	10.85	17.85	14.38
平 均	10.69	14.50	

試行数が Table 4 に記してある。これから、最小試行数で学習できたものは、HF-HA リストであり、最大試行数を要したのは、LF-LA リストであることがわかる。このことを検定したのが、Table 5 である。

Table 4, Table 5 から、次のことがあきらかである。

1. 出現頻度および連想価は、ともに、対連合において有意な効果をもつ。
2. しかし、HF-HA, LF-HA, HF-LA の3リスト間には、有意な差はない。差があるのは、LF-LA リストと他の3リストの間のみである。このことが、第5表において、 $F \times A$ の交互作用が有意になった理由である。実験 I のように、4種類の問すべてに、有意な差が生まれなかった理由は不明であるが、多分、被験者が少なかったことなども、その原因であろう。しかし、いずれにしても、少なくとも、LA においては、HF と LF によ

Table 5 Table 4 の分散分析

	SS	df	ms	F
頻 度(F)	116.28	1	116.28	9.25*
無連想価(A)	101.53	1	101.53	8.08*
被 験 者(S)	828.47	7	118.35	9.42**
$F \times A$	81.79	1	81.79	6.51*
$F \times S$	258.47	7	36.92	2.94
$A \times S$	215.22	7	30.75	2.45
誤 差	87.96	7	12.57	
計	1689.72	31		

** $p < 0.01$ * $p < 0.05$

って、試行数に大きな差があり(10.85回と17.85回)、このように、新しく導入された頻度という言語材料の特性は、あきらかに、対連合に対して、有意な効果をもつことを示しているのである。

また、試行回数 10 回までに、リストの各項目が、平均、何回、正しく連合されたかを示すものが Table 3 の「10 試行迄の平均正答数」欄である。これは、リスト内の各項目の差異を調べることを目的として算出した。もし、10 試行までに、リスト全体が、完全学習さ

れたときは、各項目は、それ以後、10 試行までをすべて正答とした。

このようにしてみると、各リストについて若干の項目については、著しく、学習に難易の差があることがわかる。HF-HA リストでは、ヒトは、イカ、コロ、に比べて、有意に正答数が多く、シマも、イカ、に比べて、正答数が多い。HF-LA リストでは、セツが、コノ、エルに比べて、有意に少なく、LF-HA リストでは、ニコがタネ、ヤミに比べて、有意に多い。また、LF-LA リストでは、項目間に有意な差はない。

このような項目による正答数の差は、頻度、連想価などと、密接な関係にあることが本研究の仮説から予想されるのであるが、しかし、これらの結果は、頻度および連想価の2つだけでは完全に説明できないのである。たとえば、HF-HA リストにおいて、ヒト ($F=61$) は、イカ ($F=57$) よりも正答数が多く、仮説に従うが、同時に、シマ ($F=24$) も、イカよりも正答数が多いというまったく逆の結果があるのである。

項目間の差が、このように、仮説に従わない原因の主たるものは、何といても、標本数が小さいことであろうが、同時に、本実験で統制されなかった他の要因、たとえば、感情性、印象性などの影響も考えられよう。すなわち、LF-HA リストにおいて、頻度はすべて0の項目間で、連想価の最も低いニコが、最も高い正答数を示しているのは、これらの未統制要因の存在を暗示していると思う。

結 論 と 考 察

従来、意味性と出現頻度は、相関が高く、同一事象の表裏であるかのように取り扱われてきた。しかし、本研究は、以上の実験Ⅰ、Ⅱの結果から、意味性と出現頻度は、同一特性ではなく、互いに、独立に操作しうる変数であることを示した。すなわち、同一連想価の綴りでも、その間に、出現頻度の高低が存在し、それによって、学習の難易が変化することがあきらかである。このことは、前述のように、出現頻度という非心理学的測度が、意味性という心理学的測度と同様に、学習を規定する特性——その強さは、たとえ、それほど、強くないと

しても——であることを意味する。

勿論、問題は残っている。すなわち、出現頻度という非心理学的測度が、如何なる過程を経て、心理学的に意味ある特性となるかという問題は、まだ、解決されていない。予想される過程は、出現頻度が、熟知度に反映することによって、心理学的に意味をもつようになるのではないかということである。Underwood(5)も、このような考えをもっているのであろうが、彼は、英語において monogram の出現頻度と熟知度の相関は高いが、digram, trigram となると、相関が低く、trigram では、 $r=.5$ 程度であるため、明瞭な結論は、出していないようである。しかし、相関は高くなくても、無相関でない限り、上述の仮説を捨てることはできない。日本語2字綴りに対する熟知度の研究の現状では、この仮説にすぐ答えられないが、言語材料の特性相互間の関係を追求するためにも、両者の関係は、是非、解決されねばならない問題である。

後記 本研究にあたって、御指導下さった京都大学 学阪良二助教授、梅本堯夫助教授の両先生に深く感謝します。また、森川弥寿雄助手にも御助言を頂いたことを深く感謝します。

文 献

- 1) 今米国晴 日本語の digram の相対頻度とその特性 心理学評論, 1960, 4, 85—100.
- 2) Miller, G. A., & Selfridge, J. A. Verbal context and recall of meaningful materials. *Amer. J. Psychol.*, 1950, 63, 178—185.
- 3) 森川弥寿雄 対連合学習の機構 哲学研究, 1961, 41, 293—360.
- 4) 梅本堯夫・森川弥寿雄・伊吹昌夫 清音2字音節の無連想価及び有意度 心研, 1955, 26, 148—155.
- 5) Underwood, B. J., & Schulz, R. W. *Meaningfulness and verbal learning*. Chicago: Lippincott, 1960.

——1961. 7. 28. 受稿——