

心的語彙の検索速度に影響する要因について (2)

久野 雅樹

東京大学大学院 教育学研究科

1 はじめに

心的語彙 (mental lexicon) の運用にあたっては、言語素材のもつ多面的な特性が影響をおよぼすと考えられる。例えば、頻度、親近性などの経験的な要因、連想のしやすさや多義性といった意味的な要因などである。言語素材を用いた心理実験を行う際、こうした諸要因については "コントロール" し、ごく限られた要因を対象とした実験が行われることが一般的だが、本研究では、多様な要因を広く参照する方向で人間の言語処理システムについての分析を試みるという主旨でデータの蓄積をおこなう。

前報 (久野, 1995) では、ひらがな表記した3文字動詞 (和語動詞; e.g. "あおぐ") を素材として、その処理速度 (語彙判断課題 = lexical decision task で測定。これは提示された語が単語であるか否かをできるだけ速く正確に判断する課題である) に影響をおよぼす要因について検討を加えた。引き続いて、今回は第2報として、カタカナ表記した清音2字綴り (e.g. "アイ") において、言語素材の特性と、語彙判断課題の反応時間の関係について分析を行う。清音2字綴りは、総数 1,892 で素材プールとして体系的に扱うのに手頃であり、以下に示すように、心理学的な記憶研究において長年にわたって蓄積されてきた知見が適用できるという利点がある。

2 方法

(1) 素材

清音2字綴り (2音節) 1,892。これは同音が連続する場合、"ン" を含む場合を除いて日本語の清音2字綴りのすべてである (4₄P₄₃)。

語彙判断課題では、刺激として単語と非単語が必要であるが、清音2字綴りには、単語も非単語も含まれているので、そのまま全綴りを用いた。分析に当っては、各綴りが単語であるか否かについての基準を用意する必要があり (これによって反応の正誤を決定する)、今回は、『分類語彙表』 (国立国語研究所, 1964) にその読み出し語が見出しとなっている綴りを単語とした。この基準により、後述の分析の対象となる単語は 744 となった。

この素材を連想価 (林, 1975; 後述) の分布がほぼ均等になるように2つ (946 ずつ) に分けることで、2系統の刺激を用意した。

(2) 単語認識速度の測定: 語彙判断課題

被験者 東京都内の大学生 36 名。2つの刺激系列に 17人ずつ割り当てた。

手続き

実験の制御は P C 9801 (N E C 製) で行った。

個々の試行では、まず、ピーブ音とともに注視点 (+) が画面に 500 msec 提示され、それがカタカナの 2 字綴りで置きかわる。被験者は、このターゲット刺激に対して、単語－非単語 の判断をできるだけ速く正確に行うことを求められた。ターゲットは最大 5000 msec 提示され、被験者の反応とともに画面から消去される。そして、1,500 msec の試行間隔において、自動的に実験は進行する。

実験に当たっては、まず練習を 20 試行行った。練習では、被験者の反応後、ターゲットが単語か否かを画面に示すとともに、被験者のセルフペースで進行できるようにした。本試行は、22 セッション (1 セッションは 43 試行) で構成され、セッション間では休息をとれるようにした。

(3) 単語の諸特性の指標

以下の 19 の指標を参照あるいは算出した。なお、今栄 (1968)、秋田 (1975)、浮田ほか (1993) については、清音 2 字綴り全体をカバーしているわけではなく、今回の分析対象となる綴りに対応するデータが存在するのは、それぞれ、68、204、184 であった。

・出現頻度

各綴りの音を読みとしてもつ語の頻度を国立国語研究所 (1970) の短単位 (意味を担っている最小単位) の頻度表によって求めた。1 つの綴りに対して複数の単語が対応する場合がありますので、すべての単語の頻度を合計したもの ("a. 総合頻度") と、最も頻度が高い単語の頻度 ("b. 第 1 頻度") を用いた。

・ネイバーフッドサイズ (neighborhood size)

音韻、綴りを共有する (e.g. hate に対して late) 近似単語がどの程度あるかが単語認識速度に影響することが指摘されている (e.g. Andrews, 1989)。本研究では、『分類語彙表』 (国立国語研究所, 1964) の索引で何度見出しとして出現するかをカウントして、近似語の集合サイズを求めた (同一単語であると思われるものも、複数の分類に入るならば別語としてカウントしている)。今回の素材は、2 文字なので、前後いずれか 1 文字を共有していれば近似単語ということになるので、"c. 全近似語"、"d. 第 1 字共有語"、"e. 第 2 字共有語" の 3 通りのネイバーフッドサイズを求めた。

また、関連して、各綴りに対応する索引の見出しの数もカウントした ("f. 見出し数")。

なお、実際の作業には、『分類語彙表』の索引を電子化したファイルである *sakuin.dat* (国立国語学研究所, 1994) を用いた。

・連想・評定による指標 (出典を示していないものは、今栄, 1968 による)

g. 連想価 : 5 秒間に連想語を書けた被験者のパーセント (林, 1976)。

h. 有意味度 : 10 秒間の平均連想語数 (梅本ほか, 1955)

i. 1 字連想 : 清音 1 字から 20 秒間にその 1 字が語頭にくるような単語を連想するように教示したときに想された単語群の語頭部分の 2 字音節の出現頻度。

j. 学習容易性 : 7 点尺度 (憶え易い－憶え難い ; "非常に憶え易い" が 7 で、"非常に憶え難い" が 1 ; 数が大きいほど憶え易い) の評定値。

k. 評定有意味度 : 7 点尺度 (有意味な－無意味な ; "非常に有意味な" が 7 = 数が大きいほど有意味) の評定値。

表 1 語彙判断課題の反応時間と単語の諸特性の相関
および、諸特性の因子分析の結果 (VARIMAX 回転後の因子負荷量 等)

	R T との相関	第 1 因子	第 2 因子	第 3 因子	第 4 因子	第 5 因子	h ²
j. 学習容易性	-0.690 *	0.951	0.165	-0.063	0.157	-0.019	0.960
k. 評定有意味度	-0.648 *	0.913	0.165	0.013	0.124	-0.055	0.880
l. 発音性	-0.572 *	0.868	0.257	0.082	-0.028	0.069	0.831
h. 有意味度	-0.570 *	0.844	0.082	0.198	-0.031	-0.039	0.761
g. 連想価	-0.444 *	0.803	0.095	0.230	0.131	0.080	0.730
n. 感情性	-0.426 *	0.796	0.103	0.122	-0.387	0.109	0.821
o. 品詞集中度	-0.593 *	0.777	-0.126	-0.217	0.180	-0.107	0.710
m. 熟知度	-0.482 *	0.771	0.437	0.199	-0.216	-0.116	0.886
-. L D T の R T	-----	-0.665	-0.170	0.013	-0.524	-0.052	0.748
a. 総合頻度	-0.222 *	0.250	0.889	0.118	-0.082	-0.043	0.875
b. 第 1 頻度	-0.248 *	0.311	0.870	0.091	-0.107	-0.040	0.875
p. 多義性	-0.175 *	-0.007	0.632	0.316	-0.125	0.370	0.652
f. 見出し数	-0.165 *	0.101	0.582	0.567	-0.012	0.166	0.698
c. 全近似語	-0.049	0.041	0.201	0.856	-0.030	0.346	0.896
e. 第 2 字共有語	-0.052	0.069	0.200	0.855	0.015	-0.232	0.830
s. 主観的カタカナ表記頻度	-0.586 *	0.208	-0.248	0.044	0.871	0.064	0.870
r. 主観的ひらがな表記頻度	-0.391 *	0.155	-0.043	-0.071	0.618	-0.068	0.417
q. 主観的漢字表記頻度	0.102	0.252	0.027	-0.033	-0.709	0.031	0.568
d. 第 1 字共有語	-0.055	0.107	0.058	0.101	-0.044	0.909	0.854
i. 1 字連想	-0.235+	0.391	-0.047	0.500	-0.032	-0.553	0.712
寄与		6.593	2.819	2.380	2.243	1.540	
寄与率 (%)		32.97	14.10	11.90	11.22	7.70	
累積寄与率 (%)		32.97	47.06	58.96	70.18	77.88	

note: 相関係数の数値の右に + とあるものは 5% 水準で有意。* は 1% 水準で有意。無印は有意でない。

- l. 発音性 : 7点尺度 (発音しやすい-発音しにくい; "非常に発音しやすい" が 7 = 数が大きいほど発音しやすい) の評定値。
- m. 熟知度 (親近性) : 7点尺度 (非常に多い-めったにない; "非常に多い" が 7 = 数が多いほどよく知っている) の評定値。
- n. 感情性 : 7点尺度 (感じのよい-感じの悪い; "非常に感じがよい" が 7 = 数が多いほど感じがよい)。
- o. 品詞集中度 : 強制的に分類させた場合に、どの程度、分類結果がひとつの品詞に集中するかの指標。それぞれの綴りについて回答が最も集中した品詞に分類した人数。
(大学生111人の回答は 18~110 に分布)。
- p. 多義性 : 同音異義語の数の評定値 (評定時間は各綴り約 5 秒) (秋田, 1975)。
- q. 主観的漢字表記頻度 :
- r. 主観的ひらがな表記頻度 :
- s. 主観的カタカナ表記頻度 :
- それぞれの単語を特定の表記形態 (漢字、ひらがな、カタカナ) でどの程度目にする事が多いかについての 3 段階で評定し、主観的な表記頻度をパーセントで表示 (浮田ほか, 1993)。

3 結果と考察

反応時間は、単語に対する正反応のみを対象とし、外れ値（200 msec 以下、2000 msec以上、各個人内で全反応の平均から ± 2.5 標準偏差 以上隔たった反応）は予め除いた。また、個人間で反応時間のベースラインの差がかなり大きかったので、各被験者ごとに全反応の平均と標準偏差を用いて標準化し、その値を17人で平均することで各単語の平均反応時間の指標とした。また、出現頻度とネイバーフッドサイズの諸指標については、分布のゆがみが大きいので、それぞれ対数変換した値を分析に用いた。

R Tの指標と素材の諸特性の相関を見ると 表1 に示すような結果となった。全般に、前報よりも高い相関の値が見られているのは、素材（e.g.前回は動詞=活用語 であったのに対して今回は名詞が中心）が、語彙判断と各種評定・連想の双方の手続きになじみやすいものであったことを反映していると思われる。

さらに、R Tを含む 20×20 の相関行列をもとにして、因子分析をおこなったところ（主成分解、バリマックス回転）、固有値、累積寄与率等を参考にして、因子数5において、解釈しやすい結果を得た（表1 参照）。すなわち、第1因子は各種の主観的な有意性・親近性の要因、第2因子は客観的頻度に関わる要因（あるいは客観的な親近性と言ってもよいだろう）、第3因子はネイバーフッドサイズの要因、第4因子は表記に限定した親近性の要因、第5因子は1文字目の特性の要因、と考えることができる。そして、第1、第2、第4因子が語彙判断課題の反応時間に主に影響していることが想定できる。主観的・客観的な親近性、そして表記、と多面的な要因が影響しているのである。

4 おわりに

本報告は、単語処理に影響する要因を広く検討することをねらった一連の研究の一部である。素材の拡張という方向で、引き続き、(1)外来語を中心とするカタカナ語（正書法が安定した仮名単語として独特である）、(2)2文字の漢字熟語（漢字表記語、漢語についての検討が必要）、および、(3)単独の漢字（漢字は文字であると同時に単語としての性質をもつ）、についてのデータの収集・分析が進行中である。

引用文献

- 秋田 清 1975 評定法によるひらがな2字音節有意語の性状印象の関与程度 その1 同志社大学 人文学,128, 1-57.
- Andrews, S 1989 Frequency and neighborhood size effects on lexical access: Activation or search? *Journal of Experimental Psychology: Learning, Memory, and Cognition*,15,802-814.
- 林 貞子 1976 ノンセンスシラブル新規準表 東海大学出版会.
- 久野雅樹 1995 語彙の検索速度に影響する要因について 言語処理学会第1回年次大会発表論文集,361-364.
- 今栄国晴 1968 日本語2字音節の単語性と抽象性 心理学研究,39(4),200-211.
- 国立国語研究所 1964 分類語彙表 秀英出版.
- 国立国語研究所 1970 電子計算機による新聞の語彙調査 秀英出版.
- 国立国語研究所 1994 [フロッピー版] 分類語彙表 秀英出版.
- 浮田 潤・杉島一郎・皆川 直凡・賀集 寛 1993 日本語の表記形態に関する研究(1) 日本心理学会第57回大会発表論文集,434.
- 梅本 亮夫・森川 弥寿雄・伊吹 昌夫 1995 清音2字音節の無連想価及び有意度 心理学研究,26(3),148-155.