

心的語彙の検索速度に影響する要因について (5)

久野 雅樹

電気通信大学

1 はじめに

本研究は、人間の語彙処理システムの特性について、語彙検索速度と単語特性の関連を通して検討する一連の研究の第 5 報である。前 4 報 (久野, 1995; 1996; 1997a; 1998a) では、日本語の典型的な単語群 (3 文字和語動詞、清音 2 字語、カタカナ表記外来語、漢字 2 字熟語) を対象として、使用頻度、親近性、多義性などの単語がもつ様々な特性と語彙の検索速度の関連を検討した。本報告では、外国語としての英単語を素材として認知実験を行なった場合、どのような処理様式が観察されるかを調べる。外国語の心的語彙がどのように構成されているかを検討することは、それ自体、一般的に重要な研究課題であると言えるが、それに加えて、英語を素材として用いるメリットとして、日本語に比して単語という単位が明快で扱いやすいこと、既に公刊されている資料が豊富であり、基準表の類を利用したり、他の研究結果と対照したりしやすいことが挙げられる。また、外国語であることから、大人を対象としても、想定しなければならない語彙の規模が小さくて済むこと、多様な学習段階の単語の存在が推定できることも、研究遂行上、大きなプラスである。

本研究は、これまでの一連の調査報告で進めてきた、心的語彙に関する様々な利用可能な情報を蓄積、統合し、データベース的な研究環境を整備する作業をさらに拡張する営みでもあり (久野, 1997b)、外国語としての英語 (あるいは日本人が学ぶ英語) における初歩的な単語について、ほぼ網羅した形で実験を行なった。具体的には中学レベル、もしくは 1000 語レベル程度の基本英単語をほぼカバーした。

語彙検索速度は、これまでと同様に、語彙判断課題 (LDT = lexical decision task ; ディスプレイに提示された綴り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) の反応時間 (RT = response time) をその指標とした。単語特性のデータは、主に MRC 言語心理学データベース (Coltheart, 1981) を参照することで用意した。なお、本研究は、久野 (1998b) の報告と内容的に重なるものだが、被験者を 2 倍に増やし、単語特性の指標を追加し、分析方法にも改善を加えている。

2 方法

(1) 素材

中学校レベルの英単語 828 を用意した。これは、文部省が定めた中学必修語 (507 語) と、中学校教科書の過半数 (7 種類のうち 4 種類以上) で使用されている単語とで構成した素材プールから、冠詞、数詞、人称代名詞、固有名詞、略語などを一部、除いたものである。中学英語教科書に出てくる単語については、小学館辞典編集部 (1993) を参照した。

この 828 語それぞれについて、1 字を置換して発音可能な非単語を作成した。そして、おのおのの単語とそれから作った非単語とが別の刺激系列に属するように振り分けて、2 つの刺激系列を作成した (各系列は、単語・非単語各 414 から構成されることになる)。

(2) 単語認識速度の測定 : 語彙判断課題

被験者 : 首都圏の大学生 46 名。2 つの刺激系列に等人数を割り当てた。

手続き：実験の制御は PC9801 (NEC 製) で行った。

個々の試行では、まず、ビープ音とともに注視点 (+) がディスプレイの中央に 500 ms 提示され、それがアルファベット文字列で置きかわる。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る。反応はテンキーの“1”と“3”を用いて行う。ターゲットは最大 5,000 ms 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 ms の試行間隔をおいて、次の試行に移り、自動的に実験は進行する。

実験に当たっては、まず 20 試行の練習を行なった。練習では、被験者のセルフペースで進行できるようにし、被験者の反応後、正解 (ターゲットが単語か非単語か) と反応の正誤を画面に示した。本試行は、18 セッション (1 セッションは 46 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

単語の特性について用意したのは、以下の 11 指標である。自身で算出した、a 文字数、d. 類似語数 と、c. 使用頻度 (ニュープロシード) 以外は、MRC 言語心理学データベース (Coltheart, 1981) による。

a. 文字数：2~13 文字で、最も長いものは“international”である。

b. 使用頻度 (Kucera & Francis)：頻度指標として、英単語の使用頻度の計量に関しては古典的であり、かつ広く用いられている Kucera & Francis (1967) を用いた。これは、1961 年時点でのアメリカ英語、約 100 万語のコーパスに基づいて集計したもので、0~69971 の値をとる (最大は“the”)。

c. 使用頻度 (ニュープロシード)：“ニュープロシード英和辞典” (長谷川・小池・島岡・竹蓋, 1993) では、延べで約 130 万語の英語テキストの分析から得られた高頻度語 5000 を“新キーワード 5000”として重要語に定め、10 段階の頻度を表示している (取りうる値は 1~10)。これは、近年の英語使用状況を広く反映し、日本人が触れる英語テキストを考慮しているという理由から参照した。

d. 類似語数 (neighborhood size)：それぞれの単語に対して、Kucera & Francis (1967) に出現する全ての語の中から、綴りが 1 字異なる語の数を算出した (取りうる値は 0~36)。

e. 親近性：主観的な親近性判断の平均値。Pavio, Yuille, & Madigan (1968) を Pavio が拡張した基準表 (未公開)、Toglia & Battig (1978) の基準表、Gilhooly & Logie (1980) の基準表を統合したもの。もともとの基準表での範囲は 1~7 だが、整数化して 100~700 の値を取りうるようにしている。

f. 具体性：主観的評定の平均値。上記の 3 基準表を統合したもので、取りうる値も同様に 100~700。

g. 心像性：主観的評定の平均値。上記の 3 基準表を統合したもので、取りうる値も同様に 100~700。

h. 有意味性：Toglia & Battig (1978) の主観的評定値を 100 倍して整数化したもの (100~700 の値を取りうる)。MRC 言語心理学データベースには、有意味性の指標として、Pavio の基準表によるものを別に示しているが、単語数も少なく、実際に RT との相関を算出しても低かったので省略した。

i. 獲得年齢：これも主観的評定によるもので、もともとなった Gilhooly & Logie (1980) の数値を 100 倍している (100~700 の値を取りうる)。

j. 品詞数：MRC 言語心理学データベースでは、*Shorter Oxford English Dictionary* に基づいて語義ごとに見出しをたてて 10 カテゴリーの品詞区分の情報を記載しているが、これは語義、品詞区分ともに細かすぎるので、Jones (1963) の *Everyman's English Pronouncing Dictionary* に基づき絞られた語義に付された 4 品詞カテゴリ (名詞、動詞、形容詞、その他) を用いて、いくつの品詞にまたがっているかをカウントした。取りうる値は 1~4 (記載がない場合は欠損) である。

k. 語義数：MRC 言語心理学データベースで上記の Jones (1963) の品詞情報が記載されている語義の数をカウントした。取りうる値は、1~14 (記載がない場合は欠損) である。

Table 語彙判断課題の反応時間と単語特性の相関

特 性	N	平 均	標準偏差	単相関	偏 相 関			
					a 除外	b 除外	d 除外	abd 除外
a. 文 字 数	828	5.01	1.72	.778	—	.763	.564	—
b. 頻度 (Kucera & Francis)	828	2.14	0.69	-.480	-.340	—	-.412	—
c. 頻度 (ニュープロシード)	823	9.02	1.48	-.444	-.327	-.237	-.361	-.174
d. 類 似 語 数	828	0.77	0.36	-.671	-.171	-.665	—	—
e. 親 近 性	654	580.10	32.82	-.257	-.278	-.130	-.273	-.187
f. 具 体 性	605	425.72	130.15	.248	.145	-.047 (ns)	.232	-.053 (ns)
g. 心 像 性	653	454.09	123.59	.180	.097	-.138	.175	-.125
h. 有 意 味 性	537	444.87	82.19	.208	.015 (ns)	-.045 (ns)	.144	-.167
i. 獲 得 年 齢	202	282.72	79.91	.597	.329	.632	.433	.375
j. 品 詞 数	793	1.58	0.66	-.210	-.131	-.235	-.110	-.142
k. 語 義 数	793	3.37	2.28	-.231	-.155	-.090 *	-.197	-.060 †

Note: 相関係数は、無印のもの $p < .01$, * $p < .05$, † $p < .10$, (ns) $p > .10$

特性の中で、頻度、類似語数は対数変換後の値に基づいて、平均、標準偏差、相関が求められている。

3 結果と考察

語彙判断課題 (LDT) の反応時間 (RT) は、単語に対する正反応のみを対象とし、外れ値 (200 ms 以下, 2,000 ms 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。RT, 2通りの頻度, 類似語数の4指標は、分布の歪みを補正し、極端な値の影響を抑えるために、対数変換 (頻度, 類似語数は、0を取りうるので特性値に1を加えて変換) した値を分析に用いた。対数変換したRTは、各単語ごとに平均して、その単語の認識速度の指標とした。なお、このRTを全単語で平均して、もとの単位に戻すと、874.67msとなる。この値は、カタカナ表記外来語 (久野, 1997), 漢字2字熟語 (久野, 1998) でのLDTのRTに同様の集計をした場合の722.45ms, 667.42msと比べると、長めであると言えるが、逆に、英語であっても、この程度の速度で認識、処理することができるという方向から見ることでもできよう。

Table にRTと単語特性との相関を示した。相関係数は、全般に、久野 (1998b) よりもわずかながら高くなっており、人数を増やしたことで、結果が安定したものとなっていると言える。

RTとの相関が最も高いのは文字数で、.778。次いで、類似語数が-.671となっている。これは文字数が少ない単語、類似語が多い単語は反応が早い、という了解しやすい結果である。文字数との相関がここまで高いのは、逐字読み的な処理がかなり行われたことを反映していると思われる。なお、文字数、類似語数の2指標は、Tableには示さなかったが相関が高い ($r = -.671$)。頻度指標では、Kucera & Francisのものが、-.480, “ニュープロシード”のものが-.444となっている。本調査の素材が高頻度語に偏っていること、および後者の指標は10段階に縮約されたものであることを考えると、低くない相関と言ってよいだろう。以上、外面的、客観的な指標、あるいは単語の基本度を示すような指標において、英語LDTのRTとの間にはっきりとした相関が見られた。

主観的評定による指標では、獲得年齢との相関が.597と高めである。これは、先の基本度のカテゴリの指標と見ることができるだろう。獲得年齢以外の評定指標は、親近性、具体性、心像性、有意味性、いずれも弱い相関にとどまっている。このうち、親近性については、素材が高親近語に偏ったために相関が低めに出たことが想像できるが、全般に、主観の部分でのずれは、先に挙げた、文字数、頻度等のクリアな指標よりもサンプルによって大きいものと推測される。なお、心像性、有意味性のように、単相関と、他の変数の影響を除外した偏相関とで相関の方向が逆転する指標も見られる。この2特性は、いずれも単独では、微弱ながらRTを速める方向で働いている。

残った品詞数、語義数は、多義性の指標と見なせる。特に品詞数は、他の影響を除いて偏相関を求めても、RT

との間に弱い負の相関が見られ、心像性、有意味性ととともに、意味の要因が単語認知に影響しうることを示す。

単語の諸特性と RT との相関関係をより単純化した形で表現するために、重回帰分析を行なった。特性によって有効な単語数が異なるため、どの特性を候補とするかで結果は異なる。有効単語数が少ない獲得年齢はまず外として、大きく分けると、主観的評定による特性群を含める場合（単語数は 600 程度となる）と含めない場合（単語数は 800 程度）という選択がありうるが、後者の決定係数が全般に高いことが確かめられた。以下に、文字数、頻度（頻度特性は Kucera & Francis のもののみとした）、類似語数、品詞数、語義数の 5 変数に対して、ステップワイズ法で変数選択を行なって得られた回帰式を示す（計算に用いた単語数は 792）。

$$RT = 0.568 \cdot \text{文字数} - 0.258 \cdot \text{頻度} - 0.142 \cdot \text{類似語数} - 0.083 \cdot \text{品詞数}$$

（標準偏回帰係数はいずれも $p < .0001$ で有意； $R^2 = .667$ ）

この決定係数は、久野（1997）がカタカナ語の RT を 4 特性（文字数、親近性、頻度、評定多義性）で予想した際の決定係数が .568 であるのと比べて高い。ただし、この結果は、文字数の規定力の強さ、つまり、逐字読みのな処理の関与に多くをよっていると考えるのが適当だろう。

4 おわりに

英単語 LDT に対する日本人大学生の反応様式を、RT と単語特性との相関から吟味したところ、英語ネイティブが示すパターン、あるいは日本人が母語である日本語による LDT で示す反応パターンと基本的に一致する結果が得られた。RT と文字数との相関の高さ、主観的評定値、意味指標との相関の低さも含めて、平均的な日本人大学生の英語語彙状態の表現として、相関の方向、程度とともに、非常に明快なパターンを示していると言ってよいだろう。このクリアな結果が、英語全体から見れば一部にすぎない基本語レベルの素材群を用いて得られたのは、外国語としての英語を用いたがゆえの成果である。

今後、英語での LDT については、語彙力のような個人差要因による反応様式の差異、学習定着度といった素材、あるいは個人内での差異、変化を視野に入れて分析を深めて行くことを考えている。こうした個人差、素材差、あるいは個人内差（変化）といった点に注目した柔軟で実際の心的語彙研究を進めるために、外国語としての英語を用いることは有効性が高いであろう。

引用文献

- Coltheart, M. 1981 The MRC psycholinguistic database. *Quarterly Journal of Experimental Psychology*, 33A, 497-505.
- Gilhooly, K.J., & Logie, R.H. 1980 Age of acquisition, imagery, concreteness, familiarity and ambiguity measures for 1944 words. *Behaviour Research Methods and Instrumentation*, 12, 395-427.
- 長谷川潔・小池生夫・島岡 丘・竹蓋幸生（編） 1994 ニューブロード英和辞典 ベネッセコーポレーション。
- 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.
- 久野雅樹 1996 心的語彙の検索速度に影響する要因について (2) 言語処理学会第 2 回年次大会発表論文集, 409-412.
- 久野雅樹 1997a 心的語彙の検索速度に影響する要因について (3) 言語処理学会第 3 回年次大会発表論文集, 181-184.
- 久野雅樹 1997b 心理実験と日本語 日本語学, 16, 44-52, 明治書院。
- 久野雅樹 1998a 心的語彙の検索速度に影響する要因について (4) 言語処理学会第 4 回年次大会発表論文集, 139-142.
- 久野雅樹 1998b 英単語の特性と語彙認知速度の関連 日本心理学会第 62 回大会発表論文集, 909.
- Jones, D. (ed.) 1963 *Everyman's English pronouncing dictionary (12th ed.)*. London: Dent.
- Kucera, H., & Francis, W.N. 1967 *Computational analysis of present-day American English*. Providence: Brown University Press.
- Pavio, A., Yuille, J.C. & Madigan, S.A. 1968 Concreteness, imagery and meaningfulness values for 925 words. *Journal of Experimental Psychology Monograph Supplement*, 76 (3, part 2).
- 小学館辞典編集部（編） 1993 中学教科書に出てくる全単語学習リスト 小学館。
- Toglia, M.P., & Battig, W.R. 1978 *Handbook of semantic word norms*. New York: Erlbaum.