

## 語彙の検索速度に影響する要因について

久野 雅樹

東京大学 大学院 教育学研究科

### 1 はじめに

本研究では、単語レベルに焦点をあてて、人間の言語処理システムを解明することを目指す。

言語処理の基本単位である単語の認識には多様な要因が影響しうることが想定できる。おおざっぱに見ても、単語は、綴りをもち、音声をもち、意味をになう、というように複合的なレベルから構成されており、それぞれのレベル内にその処理を規定する要因・特性が多様に存在することが推測できる。そして、これらの多様な要因が、最終的になんらかの形で単語の処理に関わっていることは明らかであろう。さて、それでは、それぞれの要因は、処理経路のどの段階で、どのように影響するのだろうか。

このような問題をめくって、実験心理学的なデータに基づいて語彙処理システムの構成が研究されてきている(概観については Balota, 1992 を参照)。ただし、多くの要因が関与すると考えられるため、実験要因を事前に十分に制御することが難しい。いくつかの要因を制御することによって意図せざる効果もたらされることが、少なからず予想される。たとえば、意味的な要因のひとつである多義性が単語のアクセスに影響するか、といった問題に一貫した結果が得られていない(e.g. 久野, 1994)背景のひとつとして、こうしたサンプリングの難しさを指摘することができるだろう(特に日本語の場合はそうである)。

そこで、本研究では、予め制御された少数の素材を用いて実験することよりも大きな標本数の単語素材について多数の特性を測定し、相関的に分析することで、単語処理のメカニズムにアプローチする。ここでは、特定の個々の要因の効果に限ってその有無の解明をねらうというよりも、そうした個々の要因が処理システム全体の中で占める重要性の度合い、程度を評価することを試みようとする。研究の段階としては、初歩的・予備的なものとも言えるが、包括的・総合的な性質をもあわせもつものである。

今回取り上げる要因として、大きく言えば、単語全体に関わる特性と、意味的な特性とがあるが、前者としては、その単語全体の使用頻度、特定表記での使用頻度、親近性、後者としては、具体性、意味の広さ、多義性、連想価、を取り上げる。

### 2 方法

#### (1) 素材

本研究では対象とする単語を3音節(かな表記で3文字)の和語動詞に限った。素材の例を挙げれば、“あおぐ”、“あきる”、“あげる”、等である。この3音節和語動詞を賀集(1966)の素材(523語;これは3音節動詞をほぼ網羅したものである)を中心として、402語選定した。その際、単語としての独立性が高いものを選ぶようにした。この基準によって、複合語としても解釈できるもの(e.g. “うめる”は、“生むことができる”の意にも解せる)は原則として除いた。また、語彙カテゴリが動詞以外にもあるものも原則として除いた。

理想的には、全ての(タイプの)語彙に対して、あらゆる要因を考えられればよいのだが、素材が膨大になり

すぎたり、要因が交絡しすぎたりすることは、現時点では好ましくない。素材の限定は、語彙としての性質が比較的均一で、表記や意味に関して語彙全体の下位集合として扱いやすいものとなるように配慮したもので、音節数や文字数が単語認知速度に影響を及ぼすことは当然予想されるが、本研究では扱わない。

## (2) 単語認識速度の測定：語彙判断課題の反応速度を指標として

被験者 大学生15人。

### 手続き

単語認識プロセスを調べるための代表的な手続きである語彙判断課題(LDT=lexical decision task; 提示された文字列が単語であるか否かできるだけ速く正確に判断する課題)の反応速度(RT)を用いた。

この課題はPC9801を用いた個人実験である。上記の402語とほぼ同数(実際には約8割=318語)の非単語(擬似単語;たとえば"あこる"のように語尾を単語群にあわせて動詞らしくしてある)を用意し、ランダムな順序で提示した。個々の試行では、まずピープ音とともに画面中央に注視点を提示し(500 msec)、続いて単語もしくは非単語を最大5秒提示した。被験者は、この文字列に対して、"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。試行間は練習セッションと、本試行のセッション間を除き自動的に進行する(試行間隔は1500 msec)。20試行からなる練習セッション後、本試行を実施した。本試行は40試行ごとの18セッション(720試行)から構成した。

## (3) 単語の諸特性の指標

### ・使用頻度

国立国語学研究所(1971)の新聞語彙における使用頻度のデータを用いた。表記に関わらない総頻度(かな表記、漢字表記の両方を含む)と、かな表記に限った頻度の2通りを取り上げたが、分布にゆがみが大きいので、それぞれ頻度値を対数変換した値を分析に用いた(頻度0のものは1として計算)。

### ・親近性(熟知度)

7段階(数値が大きいほど親近性が高い)の評定によりデータを収集した(約25人の平均)。

### ・具体性

7段階(数値が大きいほど具体的)の評定によりデータを収集した(約25人の平均)。

### ・意味の広さ

7段階(数値が大きいほど意味が広い)の評定によりデータを収集した(約25人の平均)。

### ・多義性

国立国語研究所(1971)の使用頻度表に見出しとして記載されている漢字表記の異なり数を指標として用いた(使用頻度0の場合は1とした)。このような基準を用いたのは、違う表記を使用するのは心理的に異なる語義だからであると思われるからであり、また、新聞の語彙として実際に使用されているものによって比較的代表的な表記についてはカバーできると考えられるからである。ただし、これは多義性の指標としては暫定的、便宜的

なものであり、総じて現実的な多義性より小さく評価されていると思われる。特に低頻度の多義については、見逃が多いようである。

・連想価（いずれも賀集(1996)による）

3通りのF価（“F連想=from-association；発散連想”の被験者あたりの連想頻度平均）

- ① [F] 価： “3音節動詞→動詞（音節数制限なし）”の制限連想でのF価。
- ② F 価： 上記の[F]において3音節動詞に限った場合のF価。
- ③ Fm 価： 自由連想のもとでのF価（頻度集計も動詞に限らない）。

T 価（“T連想=to-association；収束連想”の頻度平均；使用頻度や熟知価に近い特性をもつとされる。）

### 3 結果と考察

語彙判断課題の反応時間の分析にあたっては、まず正反応から外れ値を除き、各語について15人の被験者の反応時間の平均値を代表値とした。外れ値としては、2000 msec 以上の反応値と 200 msec 以下の反応値を除き、さらに各被験者の平均から全反応の  $\pm 2.5$  標準偏差 以上 隔たった反応時間を欠損値として除いた。また、刺激として各被験者に提示した単語は402語だが、有効反応が 2/3 未満（有効反応が10未満）であったものについては、今回の分析から除外した（この基準により、23語が除かれ、379語が分析の対象となった）。

反応時間と単語の諸特性について相関係数を求めたところ、表1に示すような結果が得られた。

まず、今回とりあげた単語の諸特性がそれぞれ反応時間と相関をもっていることがわかる（ $.103 \leq r \leq .431$ ）ただし、一部（[F] 価とF 価）を除いて、これらの諸特性は、反応時間との相関係数が最も高い親近性とも相関が高い。そこで、親近性の影響を除いて、単語特性と反応時間の相関（偏相関係数）を計算すると、ほとんどが無相関となる。各要因が、親近性を中心としてひとまとまりの要因を形作っていることがうかがえる。

こうしたなかで、[F] 価とF 価は、弱いながら、偏相関が有意であることが注目される（なお、Fm 価、T 価、についても要約方法によっては偏相関が有意となる）。課題遂行に意味的な要因が関連しているわけだが、類似の結果としては、連想語産出反応時間と語彙判断時間とに相関を見出した Chumbley & Balota(1984)の研究、文脈利用可能性（その語を使用する文脈の想起しやすさで評定）と語彙判断時間の相関が強いことを指摘した Schwanenflugel et al.(1988)を挙げることができる。

本研究におけるこの意味要因の効果については、いくつかの解釈可能性がある。まず、一般によくなされる解釈のように、語彙アクセス自体によるものである可能性がある。たとえば、連想価が高いと綴り表現に接続するノードが多く全体としての活性化が強まりやすいから、というようなメカニズムを考えることができる。しかし、その一方で、課題特有、素材特有の要因が関与した可能性も残る。語彙判断課題は、“判断”過程が含まれるので、語彙アクセス自体の指標としては微妙さをもっている。また、連想価が高いということは、その語が関連語を多くもっていて、“語群”を形成していることを意味する。この“語群”がなんらかの識別性をもっているために、“判断”課題の遂行が容易になった、という可能性も残る。これらの可能性の中から、どのような解釈の妥当性が強いかは、今後の分析にまらたい。

表1 語彙判断課題の反応時間と単語の諸特性の相関

	総頻度	かな頻度	親近性	具体性	意味の広さ	多義性	[F]値	F値	F m値	T値
LDTのRT	-0.247*	-0.207*	-0.431*	-0.275*	-0.222*	-0.103*	-0.189*	-0.189*	-0.324*	-0.282*
親近性	0.545*	0.318*	----	0.435*	0.601*	0.432*	0.127*	0.062	0.525*	0.540*
LDTのRT (親近性の影響を除く)	-0.062	-0.177	----	-0.165	-0.107	0.021	-0.241*	-0.239*	-0.148	-0.185

無印は有意でない。+ は5%水準で有意。\* は1%水準で有意。

#### 4 おわりに

本研究は、日本語における単語認識メカニズムを包括的に解明しようとする試みの第1歩である。

今後、本研究でとりあげたような要因の実質的な意味について明らかにする (e.g.そもそも親近性とはなんであるのか? 連想価にはどのような意味があるのか? 意味的要因の相互関係はどのようなものなのか? 等) とともに、諸要因の影響機序について明らかにしてゆく必要がある。もちろん、課題 (e.g.命名課題=読み上げ課題と比較する)、素材 (e.g.名詞を対象とする)、表記 (e.g.漢字を取り入れる) 等について、研究を拡張することが望まれるのは言うまでもない。

#### 引用文献

- Balota, D.A. 1994 Visual word recognition. In M.A. Genesbacher (Ed.), *Handbook of psycholinguistics*. pp. 303-358. Academic Press.
- Chumbley, J.I., & Balota, D.A. 1984 A word's meaning affects the decision in lexical decision. *Memory & Cognition*, 12, 590-606.
- 久野雅樹 1994 語彙判断課題の反応時間に及ぼす多義性の効果(3) 日本心理学会第58回大会発表論文集, 870.
- 賀集寛 1966 連想の機構 心理学モノグラフ No.1 東京大学出版会.
- 国立国語学研究所 1971 国立国語研究所報告38 電子計算機による新聞の語彙調査(II) 秀英出版.
- Schwanenflugel, P.J., Harnishfeger, K.K., & Stowe, R.W. 1988 Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, 27, 499-520.