

意味による漢字語彙の体系化

石川 皓 勇 (トッパン・ムーア)

1. 本稿の目的

カナ鍵盤上に、使用頻度が高い少数の漢字を配列して、これらとの熟語、類義関係により、多数の使用頻度の低い漢字を選択する入力方法を提案したことがあるが、その狙いは、常識的な漢字と熟語の知識を整理して練習するだけで、タッチタイプを可能にすることであり、人工的なコード体系を使用せずに、カナ漢字変換の同音語選択を不要とすることであった。(石川、昭和56、昭58)

現代の日本人に共通な漢字の使用頻度と、熟語、類義関係による語彙の体系が実在し、話す、書くなど実践されていることが、この方法の前提条件である。この実践を明示するのが本稿の目的である。

意味による語彙の体系化は、入力法のみでなく、機械翻訳や漢字教育など、多くの領域で参考になるであろう。

2. 語彙体系の考え方

日本語は僅か112種の音節(拍)を組み合わせ、すべての語を仮名表記できるので、英語のように多数の単語ごとに異なる音節と綴字を、個別に覚えなくてもよいのが長所である(金田一、昭56、P.56~57)。

しかし、その代償として、和語は語が長たらしいので、単音節語である古代中国語から、一字語を大量に輸入して、これに和語による読み方(字訓)を付与して、原語に近い読み方(字音)と併用して、日本語の構文の中で、日本語の語として使用し、簡潔、精細な「漢字仮名まじり文」を成立させた(中田、昭和57、P.163~)。日本語の同音異義語は、文字使用以前から存在していて、イル(射ル・入ル・豆を)炮ル・(お茶を)沃ル)の語例(大野、昭56、P.115~117)は、漢字使用以前の話し言葉としての和語の語彙が増加した時期から、既に日本語の意味が、文脈、場面に依存していたことを推測させる。

日本文の語順は、名詞の後に関係する動詞が来るのは、パントマイムにしても理解しやすいが、名詞の前に長い修飾語が来ると、理解しにくくなる(金田一、昭56、P.196~)。

これに対して、西欧語のセンテンスは、名詞の後に関係代名詞、動名詞(~ing)が導く従属節により意味が展開される。これを日本語に翻訳するには、従属節中の動詞を名詞に置き換えたり、被修飾語の名詞を動詞に置き換えるなどして、動詞中心の読みやすい日本文に翻訳した。例えばアメリカ独立宣言の中の一句、*dissolve the political bands which have connected them with another* を、福沢諭吉は、「他国の政治を離れ」と訳している(柳父、昭54、P.39~43)。

ある字(語)が名詞にもなり、動詞にもなるのは、古代中国語の特徴であって(香坂、昭46、P.40~)、上記のような翻訳文は、漢文の教養により作成されたのである。

漢文の古典から輸入した一字語を組合わせた熟語により、オランダ語、英語などの文献を翻訳して近代日本語が形成された事は周知の事実である。同時に同音語も増加したので、もともと文脈や場面に依存して理解されていた日本語の意味は、日本製の漢語の語彙体系に依存して理解されている筈である。

これについて、鈴木(昭50、P.73~75)は、日本語では同音語を避けるどころか、

積極的に活用している事実を述べて、

① 上位概念と下位概念の間

例、科学者と化学者、消失と焼失、(憲法の)全文と前文

② 類似概念

例、私立と市立、民俗学と民族学、水星と彗星、廃水と排水

③ 対立または反対概念

例、排外と排外(思想)、給水車と吸水車、荒天と好天

④ ある限定された思考あるいは対象領域

例、重症と重傷、幹事と監事、研修員と検収員、想像(力)と創造(力)、予言と預言などが、しばしば取り違えられ、誤解されても使用されていることを指摘し、更に二つの概念の関連性を示すために、まず同一の音形を使い、次にその差異を文字で示す造語法(辞典と事典、名答と迷答など)を示した。

このように漢字が日本語の中心で機能しているにもかかわらず、漢字が国語辞典の見出し語に採用されたのは、昭和38年の「岩波国語辞典」(西尾、岩淵、水谷)からである。常用漢字の日本語としての意味の成立を解説したものは、実に今年(昭和60年)の「字義字訓辞典」(佐藤、角川小辞典4)、JIS第一水準漢字などを含む「漢字情報辞典」(金田一、高橋、(株)語研)が初めてのものである。従って、日本人は、常用漢字の意味を未整理のまま使用している場合が多いようである。

昭和48年に告示された「送り仮名の付け方」は、漢字一字で表記して音読または訓読する語を「単独の語」と呼び、漢字二字以上で表記して、音または訓を複合させている語を「複合の語」と定義している。

送り仮名は、漢字と仮名の表記上の接続手段であり、漢字かなまじり文にとって最も本質的な問題であるだけに、解決困難な問題であった。この告示(原案は、熊沢竜 作成)は、中国語と日本語の比較を語形変化、即ち「活用」の有無で促えて、従来の品詞別の整理法に代えた。また、従来の文法学で単語を「単純語」と「複合語」に二分してきたが、その内容には学者によって意見の相違があり、共通理解が不可能であった。これを子供でも理解できる漢字の数で整理したことによって、送り仮名法が初めて国民全体のものになった。(島田、昭59、P.133~148)

語彙は世界像を反映した目録であるとする立場からは、実体(独立した実在)の概念を表現しない「てにをは」や、「~的」「~人」などの要素が軽視されがちであるが、これらの要素こそ使用頻度が最も高い事が知られている。実体を指示する働らき(指示的意味)だけが意味ではなく、連想を喚起し、文脈を展開させる働らき(喚起的意味、内包的意味)

もまた意味である。オグデンの Basic English が、動詞の節約に成功したのは、汎用の動詞と前置詞を組合せた連語を認めたからである(玉村、昭59、P.59)。

タイピングにおいては、話す、書くなどの言語行動と同様に、言葉を時間と共に展開する「連系」として扱う。連系としての扱いは、言葉を表現過程として見る立場に載りやすく(水谷、昭45、P.74~76)、この立場の文法(時枝文法)では、「暖か」、「甘」、「大きい」、「赤さ」、「商人」、「的」、「御」など、名詞とするにはふさわしくないが、或る観念を表現し、かつ語形変化をしないものを体言としている(時枝、昭25、P.70~71)。

言葉として表現し得る、即ち連系になり得る意味の空間を、語彙により展開するのは、日本語の場合、上記のような短い要素を出発点とするのが適当である。

その理由は、話す単位は音節であって、単語は音節により構成され、語彙は、最少の音節から成る単語を中心に成り立っている事であり、これは、どの言語にも共通な一般的な事実である。

以上を要約すると、言語の単位は「語」であるが、語彙即ち語の集合として促えた場合、中心的に使用されるのは一音節語であるので、話される言葉は、音節ごとに聞き取られて理解される。音節を表現する仮名、および、元来一音節語であった漢字は、当然、理解されやすく、漢字仮名まじり文はこの長所を生かしたものであり、カナ鍵盤による入力法も、この長所を生かすのが当然であると言える。

3. 語長と使用率

昭和41年の新聞から抽出した延べ94万語の80%を占める頻度順位2,000までの語の語長別の内訳を、第1図に示す(国研報告37、P. 139～、度数順短単位表による)。

「語」を次の例のように区切っている(図説日本語、P. 583～) 親子 | 関係 | 不 | 存在 | 認知 | 取り消し | 要求 | など | の

国立 | 国語 | 研究 | 所

スペイン | 風 | トマトソース

ひらがな、数字などの一字語が、200位までに累計45.1%あり、ひらがななどの二字語は、300位まで累計4.3%あり、以下には稀である。

漢字の一字語は1,000位までの累計13.0%、2,000位まで14.2%漢字を含む二字語は1,000位で累計8.5%、2,000位で12.6%

三字以上の語は2,000位までで、累計1.8%である。

下位に向って増加傾向が大きいのは、漢字を含む二字語と三字以上の語である。

1位から50位までの累計46.4%の内訳は、下記の通り。数字、記号の他は、主として仮名の一字語であり、これらが、日本語の「語」の知覚の核心になっている。

数字15.0% (0, 1, 2, 5, 3, 4, 6, 8, 7, 9)

記号12.7% (, , . , 諸記号 , . [,] , - , ' , ' , ")

仮名1字14.4% (の、を、に、は、が、て、と、で、た、し、の(助詞)、た(助動)、も、に、な、れ)

仮名2字1.3% (いる、から、ある、こと、ない)

漢字1字3.0% (一、二、三、万、五、〇、日、円)

漢字の一字語は200位までで8.6%、300位までで9.8%に達しており、これらの漢字を中心にして、漢字二字語が形成されて、下位に向って多様な熟語が形成されると推測される。

4. 意味空間の展開モデル

上記の日本語の特徴を生かして、作文即ち意味展開が可能なように、鍵盤上に漢字を配列する方法を第2図に示した。

4.1 全体的特性

① 連糸0→1→2の順に意味が分かれる。

連糸0は、テニヲハによる構文と基本的な漢字の意味カテゴリーの接点。

連系1は、鍵盤上の基本的な漢字、(鍵盤内漢字)

連系2は、上記の他にその他の多数の漢字(鍵盤外漢字)を含む

- ② 階層的。使用頻度が高い一字語でグループの意味を代表させる。
- ③ 一義的。多義の高頻度漢字では中立的な意味で代表させる。(後述する)
- ④ 連系の同一レベル内で結合(網構造)。
- ⑤ 他グループの異なったレベルの文字との連結を許容する。

4.2 連系0 (文脈空間)

- ① カテゴリは数種類に限られる。これは人間が同時に考え得るのは数個に限られるからである。
- ② 格。「てにをは」は主語、述語、客語などの文中の役割を表わすので、これらと類似したカテゴリにより漢字の意味を分類すれば、文脈に調和する。
- ③ 指示的意味。高頻度漢字の意味から歸納した
主体(人、団体、土地)
事柄(政治、社会、学術)
自然、言動(人名、地名を含む)
により、「てにをは」に類似させたが、この他に形式的認識(時間、空間、存在)を示す漢字が多いので、これを中心に、鍵盤の各段に配置する。

4.3 連系1 (一字語空間、鍵盤)

- ① 高頻度の漢字を選ぶ。
- ② 多くの二字語を作る漢字を選ぶ。
- ③ 一義的、中立的に(原義と比喻の中立点を選んで)配列する。例えば「長」は「長い」「長所」「社長」の意味があるが、共通の属性を表わす「長安」という句で配列する。
- ④ 類義の漢字は、なるべく同一キーまたは隣接キーに集める。
- ⑤ 各指の分担するキー(鍵盤上の縦の配列)の喚起的意味を統一する。ホームポジションのキーに、このためのキーワードとなる漢字を配置する。(第3図で∩で囲んだ2字)。
- ⑥ 各キーの漢字が四字句を構成するようにし、頭文字にはカナキーと同じ読みのものを、なるべく選ぶ。

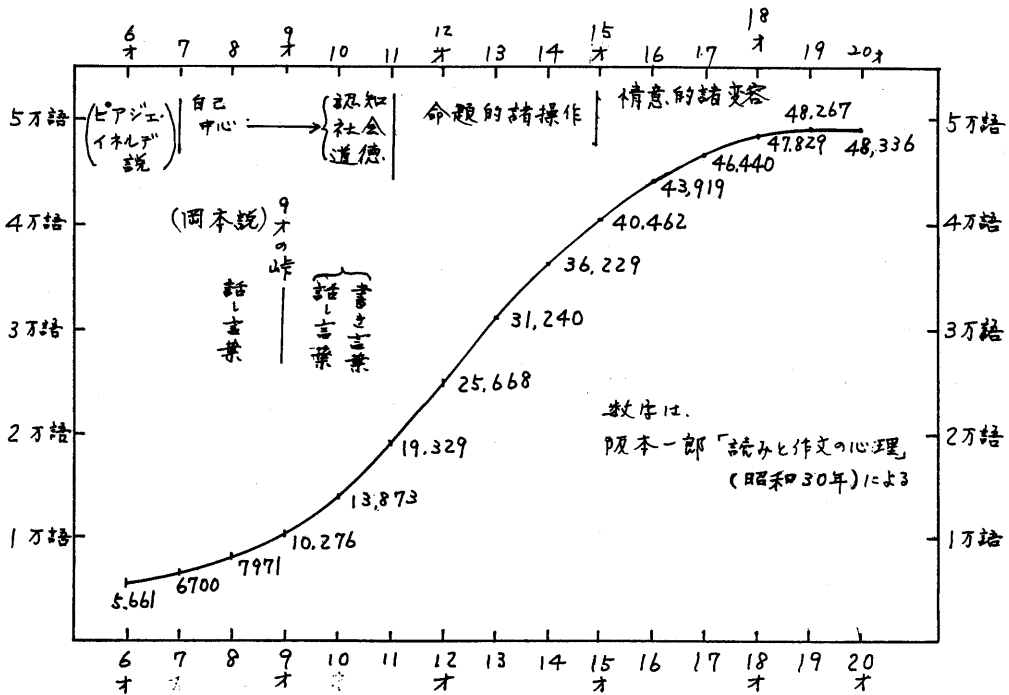
4.4 連系2 (二字語空間、辞書)

- ① 一字で意味が明白な漢字は、なるべく類義のものを同じ場所に集めて、グループを形成する。
- ② 一字で意味が不明瞭な漢字は、熟語関係で配置する。
- ③ 二字語は種々の結合関係で成立し得るので、検索しやすいように中立点を求めることが特に重要になる場合がある。例えば漢字「満」の訓は「ミチル」であるが、満水、満月、満期、満場、満身、満年令、満潮、満足、満開、円満、不満、満州、などの語に用いられる。このうち、最も広い共通の意味は、時間の経過であるから、「満期」を選び、鍵盤上の漢字「期」の下位区分として、「初」「待」と共に所属させた。(第6図の上から3段目、左から4列目の「ひ秋夜早期」の欄)。

第4回 意味カテゴリーと学年配当

	顔軍 ^{E4} 忍撃 ^B	法主 ^{E3} 政任 ^{E5}	不病 ^{E3} 租苦 ^{E2}			演進 ^{E3} 習世 ^{E3}	身衣 ^{C4} 体食 ^{C2}	屋童 ^{C3} 火守 ^{C2}		
送化 ^{C3} 受業 ^{C3}	法生 ^{E1} 戦死 ^{E3}	会同 ^{E2} 経性 ^{E5}	自何 ^{B2} 当者 ^{B3}	良子 ^{B4} 象女 ^{B2}	通学 ^{E1} 交術 ^{B2}	運御 ^{E3} 天土 ^{E1}	野吉 ^{C2} 火木 ^{C1}	海川 ^{C1} 山水 ^{C1}	理潔 ^{C2} 徳空 ^{C2}	色葉 ^{C3} 花光 ^{C2}
今立 ^{B1} 様東 ^{B2}	変終 ^{A3} 別閑 ^{A4}	来果 ^{B3} 市名 ^{B1}	秋早 ^{A1} 夜期 ^{A2}	用要 ^{B4} 和利 ^{A3}	正本 ^{A1} 見原 ^{A2}	平円 ^{A1} 面形 ^{A2}	間算 ^{A2} 次教 ^{A2}	解度 ^{A3} 明回 ^{B2}	有現 ^{A5} 事実 ^{A3}	
外分 ^{A2} 力分 ^{B1}	会利 ^{A4} 場殿 ^{B2}	的地 ^{A2} 所方 ^{A2}	年日 ^{A1} 月時 ^{A2}	対国 ^{B2} 人民 ^{B4}	入中 ^{A1} 金右 ^{A1}	大上 ^{A1} 小下 ^{A1}	一合 ^{A2} 二計 ^{A2}	意思 ^{B2} 表知 ^{B2}	教情 ^{B5} 行感 ^{B2}	成強 ^{B2} 心然 ^{B4}
出底 ^{D5} 屋動 ^{D4}	口手 ^{D1} 目足 ^{D1}	苦物 ^{D3} 加業 ^{D4}	作戸 ^{D2} 美料 ^{D3}	新前 ^{A2} 高開 ^{A3}	慮長 ^{A2} 間安 ^{A2}	派系 ^{D1} 持機 ^{D3}	論証 ^{D2} 集紙 ^{D2}	止定 ^{D3} 悪品 ^{D3}	文言 ^{D1} 書切 ^{D2}	

第5回 理解語彙の発達



- ④ 開いている空間である。(3字語以上の空間を許容する。)
- ⑤ キー配列との関係。類義の鍵盤外漢字は、同一キー内の連続した配列の鍵盤内漢字(鍵盤上の漢字)に所属させる。

4.5 外部的要因

以上の方法で漢字を配置すれば、常識的な漢字の知識を整理して練習するだけで、タッチタイプが可能であるが、打鍵を混乱させる漢字体系以外からの影響を避ける必要がある。

カナキーの配列は、漢字の体系と直接には無関係であるが、各キーに四字句として配置した漢字の頭文字の読み方をカナキーと一致させて、暗記しやすくしている。カナキーの配列が変わると漢字の配列も変わる。多義の漢字の中立的な意味は、漢字句の中の前後関係により変化するので、カナキーの配列を変えた場合は鍵盤内漢字の配列を十分に練り直さなければならない。鍵盤外漢字への影響は間接的なもので、鍵盤内漢字に所属するグループ単位での移動が可能である。

マルチストロークの打鍵法は、なるべく単純な方法が良い。各字の先頭に所定の漢字シフトキーを打鍵して、

- ① 一字の先頭を示す記号
- ② カナ、鍵盤内漢字、鍵盤外漢字、語句の区別を示す信号
- ③ 所属する鍵盤内漢字のシフト段数(キーに配置した四字句の何字目であるか)を示す信号を兼ねさせるのが最も単純なようである。

鍵盤内漢字は漢字シフトキー、文字キーの順に打鍵し、鍵盤外漢字は更に読み方を打鍵する。

同一の鍵盤内漢字に所属する鍵盤外漢字に同音異字が生じる場合がある。これは短縮打鍵(漢字音の第2音節のイ、ウ、キ、ク、チ、ツ、ンの省略、濁点省略など)、同音漢字が多い拗音(シャ、シュ、ショ)の処理にも関係がある。これらを考慮して、鍵盤外漢字の配置を変更すると、打鍵に際してその漢字の所在が想起し難くなる。従って、あくまで想起しやすい中立的な意味体系で配置することを主とし、同音のコードが生じた場合は、コード付けと打鍵法のレベルで解決すべきである。

5. 漢字の学習過程

以上のように展開した漢字の意味体系と、児童、生徒の漢字学習の過程との相関を考察する。

5.1 鍵盤内漢字の性格

第3図のカナ鍵盤の配列は、工業技術院の新JIS案を基準にしているが、ホームポジションにある、し、く、う、い、のキーに、テ、ニ、ヲ、ノ、を入れ替えている。新JIS案はカナ漢字変換方式用であるのに対し、第3図は、てにをは類を中心とする配列だからである。(朝日新聞、昭和60)

各キーの上下に四字づつ配置した漢字は、雑誌90種に高頻度の漢字から選び、和文タイプライターの一級文字なども参考にした。(国研報告22、昭和37。竹内、昭和49。)

第4図に各漢字の右下にその意味カテゴリーと、小学校の学年別に学習すべき漢字の配当を示す。(小学校学習指導要領、昭26～52)(意味カテゴリー)

A形式的認識。B主体。C自然。D言動。E事件。(学年別漢字配当)

1…1年生。 2…2年生。 3…3年生。

(左手)

第6回 丁IS第一水準漢字の体系化	上、シフト	あ	ほ	こ	か
	上校	い	せ	し(て)	じ
	中、シフト	え	そ	す	じ
	中校	は	な	に(り)	に
下校	ち	ひ	ふ	へ	ほ

上付数字1、5、6は教育漢字の学年配当、(5)、(6)、(7)は百位制への頻度、2、7、16は100%頻度、20は100%以下、常用漢字を示す、
 () 内の漢字は、重複登録または三重登録を示す、
 上付数字は、1、5、6は教育漢字の学年配当、(5)、(6)、(7)は百位制への頻度、2、7、16は100%頻度、20は100%以下、常用漢字を示す、
 頻度、100%以下、常用漢字を示す、
 頻度、100%以下、常用漢字を示す、

例えば「大A1」は「大」という漢字の意味カテゴリーが形式的認識であり、小学校1年生で学習されることを示す。

第4図を閲覧すると、1年生は形式的認識、2年生は主体、自然、3年生は言動、3年生以上から高学年にわたり事件のカテゴリの漢字が学習される傾向が見える。

ほゞ3年生(9才)までに、これらの基本的な観念を習得していると推定される。

5.2 理解語彙の発達

第5図において、9才で書き言葉の生活を始めると、11才までに理解語彙が倍増し、以後15才まで毎年5000語程度語彙を増している。11才(5年生)は漢字を書く成績が最低になり、同音異字の使い分けと、抽象的な意味の漢字が頭の中で十分に整理し切れない時期である。一方読みの成績は下がらず、学習していない漢字を読めるようになる。(海保、山田、昭和59、P.62、P.139)。

この事実は、小学校高学年から、中学校卒業(15才)までの間に於て、漢字の意味を体系的に整理することの重要性を示している。大学を卒業しても不明瞭な文章を書くのは、漢字の字体を知らないからではなく、その概念と語法を正確に理解していないからである。小学校高学年から中学校にかけての認知、命題的諸操作の発達の時期に、これが行われないと、理解力と思考力の限界が生じる。(岡本、昭60。ピアジェ、イネルデ、昭41、邦訳 P.130~P.133、第4章の結論と第5章の序論)(参考文献)

石川 情報処理学会(昭56後期)全国大会予稿集P.965

石川 日本文入力方式研資12-1(昭58年11月)

金田一春彦 日本語の特質 日本放送出版協会

中田祝夫 日本の漢字 中央公論社

大野晋 日本語とタミル語 新潮社

柳父章 比較日本語論 日本翻訳家養成センター

香坂順一 中国語学の基礎知識 光生館

鈴木孝夫 閉された言語・日本語の世界 新潮社

西尾実、岩淵悦太郎、水谷静夫 岩波国語辞典 岩波書店

佐藤喜代治 字義字訓辞典 角川書店

金田一春彦監修、高橋秀治編著 漢字情報辞典 (株)語研

島田昌彦 日本語の再生 桜楓社

玉村二郎 語彙の研究と教育(上) 国立国語研究所

水谷静夫 言語と数学 森北出版

時枝誠記 日本文法口語篇 岩波全書

国立国語研究所報告37 電子計算機による新聞の語彙調査

林大、他 図説日本語 角川書店

朝日新聞 昭60.3.27、3面 ワープロの文字配列

国立国語研究所報告22、現代雑誌90種の用語用字(2)漢字表

竹内甚一 和文タイプ 金園社

大矢武師・野地潤家 国語教育必携 教育出版センター

海保博之編 漢字を科学する 有斐閣選書

第5図の阪本の統計は、上記の玉村のP.98に所載

岡本夏木 ことばと発達 岩波新書

ピアジェ&イネルデ 新らしい児童心理学 白水社、文庫クセジュ