

専門用語研究

Journal of the Japan Terminology Association

No.17 1999-02

文字シンポジウムから

文字概説	矢島 文夫	1
日本語表記	石井 久雄	7
漢字の構造を考える——コンピュータの入力・検索の視点から——	松岡 榮志	15
脳と文字	杉下 守弘	22
漢字コードについて	宮澤 彰	25

専門用語集の編集と活用——用語集超活用ソフトHT³(エイチティースリー)

の利用効果——	貝島 良太	31
外字の作成事例：『広辞苑』を中心に	上野 真志	38
国内刊行用語辞典リスト(1998 後半)		42
JIS用語規格制定の動き 1998		48
編集後記	編集後記	52

専門用語研究会
Japan Terminology Association

文字シンポジウム

1998年11月6日、日本学術会議講堂にて、約70名の参加者を得て、「文字シンポジウム」（専門用語研究会と情報知識学会との共催）を開催した。

今回は、文字の機能を視点として、漢字に配慮し、各方面の研究の概要と動向を、客観的に捉えようと試みた。文字については、様々な意見が様々な場面で討議されている現状を鑑み、序論的な講演を5氏よりいただいた。

本号では、各講師に修正・加筆いただいた2件（石井・宮沢）と予稿集からの転載3件（矢島、松岡、杉下）を掲載する。

～プログラムおよび講師～

総合司会	石川 徹也	図書館情報大学教授
挨拶	柴田 武	専門用語研究会会長
趣旨説明	仲本秀四郎	IRIS
講演セッション		
司会	千原 秀昭	化学情報協会副会長
1. 文字概説	矢島 文夫	アジア・アフリカ図書館長
2. 日本語の表記	石井 久雄	武蔵大学教授
3. 漢字の「構造」を考えるーコンピュータの入力・検索の視点からー	松岡 榮志	東京学芸大学教授
4. 脳と文字	杉下 守弘	東京大学医学部教授
5. 漢字コードについて	宮沢 彰	学術情報センター教授

～文字シンポジウムから～

文字概説

矢島 文夫* YAJIMA Fumio

- 1 文字の出現とその系譜
- 2 古代オリエントの文字体系
- 3 古代中国起源の文字体系
- 4 アルファベット体系(1) セム系子音文字
- 5 アルファベット体系(2) 「完全アルファベット」
- 6 アルファベット体系(3) 準音節文字への発展
- 7 その他の文字体系
- 8 現代世界の文字問題

1. 文字の出現とその系譜

「文字」という用語を人間の音声言語に対応し、多少とも表音性を示すものと定義して使用する限り人類史のなかで文字という情報伝達・保存の手段が出現したのは、ごく最近の出来事にすぎない。人類の歴史を近年の研究に従って約600万年とするならば、「文字」の成立は多めに見積もっても今からせいぜい6000年以内のことであるから、それは全人類史の長さの1000分の1か、それ以下ということになる。

ただしそれ以前の時代には情報伝達（時には保存）の方法がなかったというのではない。音声言語はごく早い時期から使われていたであろうし、口笛や太鼓や煙を使う情報伝達はかなり古くから行なわれていたと思われる（これらのなかで煙火を使うのろし台は近代まで使われていた例もある）。しかし音声や煙火を使うものは情報の保存には適していない。文字の創造に役立ったのは、旧石器時代にすでに生じ、徐々に発達した人類の造形能力、なかでも絵画と彫刻の能力であったと思われる。原始絵画の代表としてはアルタミラやラスコーの洞窟絵画、サハラ奥地のタッシリ岩壁絵画などがあり、また原始彫刻としては旧大陸各地で数百

体発見されている地母神像（後代のアフロディテ＝ヴィーナスの原型とされているもの）などがあるが、文字の出現はこれらを生み出したリアリズムの表現能力とともに、次には事物を特定の図形により標準化する抽象能力が必要であった。とりわけ初期の文字体系の一部として生じた数字は、すでにかなり発達した数概念と計算能力を示しており、旧石器時代末期から新石器時代（前15000～5000年頃）の人類の知的能力が急速に発達したことを示している。

今日知られている最古の文字遺物は、メソポタミア（イラク）中部のウルク遺跡出土の粘土書板に記されたウルク文字（一種の象形文字）で、前3100年頃のものと思われる。この文字体系は、次にはいわゆる楔形文字体系に発展し、この地域で豊富に入手できる粘土を利用した書板上に各種の記録・情報が記され、初期都市文明の発展に貢献したメソポタミアにおける楔形文字の出現と普及は古代エジプト人にも刺激を及ぼしたが、ここでは方式の異なる聖刻文字（ヒエログリフ）体系を出現させ、両者ともに3000年にわたって使用された。前2000～1500年にはここからアルファベット文字体系が出現し、その発展形態は今日まで続き、現代文明の成立に寄与してきた。オリエントにおける文字の出現は東アジアにまで影響を及ぼし、中国では前2000～1500年に漢字の祖形が生まれた。これらの文字体系から多くの派生文字が出現したが、これらについては別図を参照されたい。また、上記の主要文字体系については、下記の個別的記述を見られたい。

2. 古代オリエントの文字体系

古代オリエントの文字体系はメソポタミアの楔形文字体系は古代エジプトの聖刻文字（ヒエログ

* アジア・アフリカ図書館長

リフ)体系に大別される。前者のうち前述のウルク文字(楔形文字の前身)はまだ十分には解読されていないが、シュメール語が使われていると思われる。文字記号は1000~1500数えられており、その多くは単語文字(1文字が1単語を表わす)から成っているが、その後急速に複雑な用法が発達した。例えば「頭」を表わす文字記号の口の部分にグーヌー記号と呼ばれる斜線を付して「食べる」とか「言葉」を表わすような方式が生じ、これが別個の文字に発展した。他方、音節(日本の仮名文字の「か=ka」「さ=sa」のようなもの。他にak, asなど、さらにはkak, sasなども含む)を表わす表音文字(音節文字)が生じ、この音声表記が進むと、文字記号の総数は減少するようになった。このシュメール文字はまもなくセム語族アッカド語(バビロニア・アッシリア語)の使用者たちに借用され、アッカド語による多種の文書が作成された。ここでは600ほど(後期には350ほど)の文字が数えられるが、多くのシュメール語単語がそのままアッカド語の文脈に借用され、複雑な使われ方が行なわれた(この点は、日本文字体系に取り入れられた漢字の複雑な用法と似ている)。

古代エジプトの聖刻文字(ヒエログリフ)は、前3000年頃にかなり突然出現したが、これはメソポタミアにおける文字出現から刺激を受けた結果と考えられる。しかしここでは、子音のみを表記する方式(ただし、24個の「アルファベット」、すなわち1音表記文字以外に、2音あるいは3音を表わす「複音文字」も含む)を用いており、メソポタミア楔形文字の「音節文字」方式とは構造上まったく異なっている。この「子音文字」方式は、のちのセム子音文字(フェニキア・アラム・ヘブライ・アラビア各文字)に受け継がれ、今日まで主としてアラビア文字、ヘブライ文字の特徴となっているものである。しかし古代エジプト文字は、多くの単語文字(表意+表音文字)および限定符(単語の意味を限定する符号)を用いたため、文字数は800~1500(時代による)と、かなりの数になった。

古代エジプトの聖刻文字の絵文字(ピクトグラム)的構造は、たぶんヒッタイトおよび古代地中

海の象形文字群に間接的影響を及ぼしたが、エジプト聖刻文字から直接に派生したのはヌビアのメロエ文字のみであった(別図参照)。

3. 古代中国起源の文字

漢字および日本文字の細部については専門家の方々の発表もあるので、ここでは原・漢字の出現期(前1500~1000年と推定)における先行文字(古代オリエントの文字群)との比較に論議を限定することにしたい。古代オリエント文字、とりわけ楔形文字の先行形態の出現は前3000年か、その少し前と考えられる。とりわけシュメール文字と親近性のあるエラム文字(西イラン)が注目されるが、これらの文字使用についての情報は、きわめて緩慢にユーラシア大陸の東方へと伝えられたように思われる。前1500年ころに、今はまだ知られていない原・甲骨文字が中国西部で作られた可能性がある。

これまでに初期漢字を古代オリエント文字(特にエジプト聖刻文字)と同起源のものとする論説があったが、両者の直接の関係を論証することは難しい。両者ともに絵文字(ピクトグラム)的要素をもつために、太陽、月、山、川、木などを表わす文字に形態上の共通性が生ずることがあるが、これは影響関係の結果とは思えず、人類の絵画能力の共通性とみなすべきである。

漢字は(若干の非表音的文字を別として)単語文字的性格をもっているが、この文字記号を借用した日本文字では、純粋表音文字である仮名文字を派生させ、混合文字体系として使用されている。この点は、シュメール文字を借用しつつ独自の音節文字(純粋表音文字)との混合文字体系を使っていたアッカド語楔形文字体系、さらにこの2語の用法を取り込んだヒッタイト(ほかにフルリ、ウラルトゥ)楔形文字体系と構造上きわめて似ており、さらなる比較研究が求められる。

漢字体系は巨大なものとなり、多くの書体・字体を生むとともに、周辺世界に西夏文字、ベトナム漢字その他の支流を出現させた。他方、現代になり、政治的に革新の道を歩んだ漢字のふるさと中国では新字体(簡体)が一般化しつつあるのに対し、伝統性を維持する台湾や周辺の華僑世界で

は旧字体（繁体）の使用が続くという事態（これはメソポタミア楔形文字体系におけるアッシリア字体とパピロニア字体の対立に相当する）が生じているが、今後これはどうなるだろうか。

4. アルファベット体系(1)

セム系「子音文字」

紀元前 2000～1500 年（たぶん前 1700 年頃）に東地中海沿岸で最初のアルファベット文字（単音文字）が生まれた。今日知られている最古のものは今世紀はじめに発見されたシナイ文字で、22 個の子音文字をもつフェニキア文字の前身と考えられる。言語音のうち子音だけを表記する方式（m-l-k＝「王」など）は古代エジプト人がすでに使っている（n-f-r＝「美しい」）ので、これを踏襲しているにすぎないが、文字記号は独自に案出している。こうして第 1 字は「牛（の頭）」（ヘブライ語でアールフ）の形、第 2 字は「家」（ヘブライ語でベート）の形が作られ、ここからアルファベットという呼称が生じた。22 個の子音文字方式は今日のヘブライ語に受け継がれているほか、ナバテア文字を介して生じたアラビア文字（基礎文字は 28 個）および、その派生形ペルシア文字、ウルドゥー文字などが今日広く使われている。アラビア文字は古くにはオスマン・トルコ語やマレー語、スワヒリ語などの表記にも用いられたが、子音文字群を基礎的な意味素とするセム語と異なる上記の語群では、表記が十分できないといううらみがあり、今日では漸次ラテン文字に置き換えられている。ヘブライ文字はユダヤ教、アラビア文字はイスラム教の伝統と密接にかかわっているためもあり、文字改革の動きは少ない。

5. アルファベット体系(2)

「完全アルファベット」

22 個の子音文字から成るフェニキア文字は前 1000＝500 年頃に地中海各地に伝えられ、前 900～800 年にギリシア人が自己の言語をこれで表記することを思いついた。しかし子音文字だけの表記では母音の豊富なギリシア語を十分に書き表わすことができないので、いくつかの原文字記号に手を加えたり文字を追加し、基礎文字 24 個

（うち母音文字 7 個）のギリシア文字を生み出した。ここから初期にはコプト文字、ゴート文字など、そして次にはスラヴ文字が、いずれもキリスト教布教のために作られた。スラヴ文字は今日のロシア文字とそのいくつかの支流へと発展した。他方、ギリシア文字はエトルリア文字を介してラテン文字（ローマ字）を出現させた（字母数は後代に付け加えられた J, W を加えて 26 個）。これらは当該言語をかなり忠実に表記するところから、しばしば「完全アルファベット」と呼ばれるが、言語音は多様であり、時とともに変化するから、必ずしも完全に表記しているわけではない。セム語における子音表記の用法はエトルリア語などではかなり残っており、後代の各国語の表記にも見られる（例えばチェコ語 srp「鎌」）。言語音の変化に伴い、無声化（ラテン語 et に対するフランス語 et）、同音の別表記（英語の night と knight）などの問題も生じ、「完全アルファベット」使用圏では現実の言語音と正書法による表記とのずれをどうするかがつねに問われている。

6. アルファベット体系(3)

準音節文字への発展

フェニキア文字は他方で 2 つの準音節文字、すなわちインド系譜文字とエチオピア文字を生んだ。準音節文字というのは、インド系譜文字の代表であるデーバナーガリ文字で見ると、子音 k と音節 ka を同一文字で表わすが、kaa, ki, kii, ku, kri (ri は母音としてあつかう、kee, kai, koo, kau は k/ka 文字に付加要素を加えて表わす（すなわち全部で 10 種の字母が生ずる）というように、単音文字と音節文字（複音文字）の混合形態のことであり、歴史的には古代ペルシア文字（非メソポタミア系楔形文字）がこれであった。インド系文字の場合、結果として 200～300 個の字母が生じており、「完全アルファベット」群とは対照的に、学習および読解に複雑さが生じていることは否定できない。しかし母音を独立文字とせず付加要素で処理することにより、表記スペースを節約できるという大きなメリットがある。それに対して「完全アルファベット」表記は、ときには長いスペリングの単語を生む（フィンランド語

harvalukuinen「少数」)。その点で、準音節文字群は、かなりよく出来た文字体系といえることができる。

7. その他の文字体系

オリエント文字体系(楔形文字系および聖刻文字系)、漢字文字体系およびアルファベット文字体系(これは数十種もの支流を生んだ)以外にも歴史上現われたいくつかの文字体系がある。古代文字の領域に入るものとしては、ほとんど未解読のインダス文字と地中海のいくつかの文字群(「ファイストスの円盤」はその代表)、解読がやや進んでいる中央アメリカのマヤ・アルテカ文字がある。小規模のものとしては、南海イースター島の「コハウ・ロンゴ・ロンゴ」の絵文字、中国南部のいくつかの絵文字群(トンパ文字など)があるが、これらは表音性の乏しい表意文字とみなされる。

他方、近世以降に作られたものに、朝鮮半島で使われているハングルがある。これは子音、母音、半母音を示す文字要素を組み合わせた「完全音節文字」ともいうべきもので、 $\begin{matrix} K & J & a \\ & & K \end{matrix} = kak$ のように子音文字+母音文字の下部に子音文字をおく構造は、前記の準音節文字と類似の音節文字群(約2000個)を生み出しており、表記スペースの節約につながっている。

独立的文字体系としては、これらの他に、文字をもたない言語使用者に対して新たに(たいていは既存文字をモデルにして)文字を作成した例が10~20数えられる。これらのなかには、活字が作られて印刷物が発行されているものもあるが、その使用範囲は限られていることが多く、使用中止になっている例も見られる。

8. 現代世界の文字問題

今までは主として歴史上の文字を見てきたが、現代世界とのかかわりとしては、(1) 政治と文字(インターナショナルリズムとナショナルリズム)、(2)

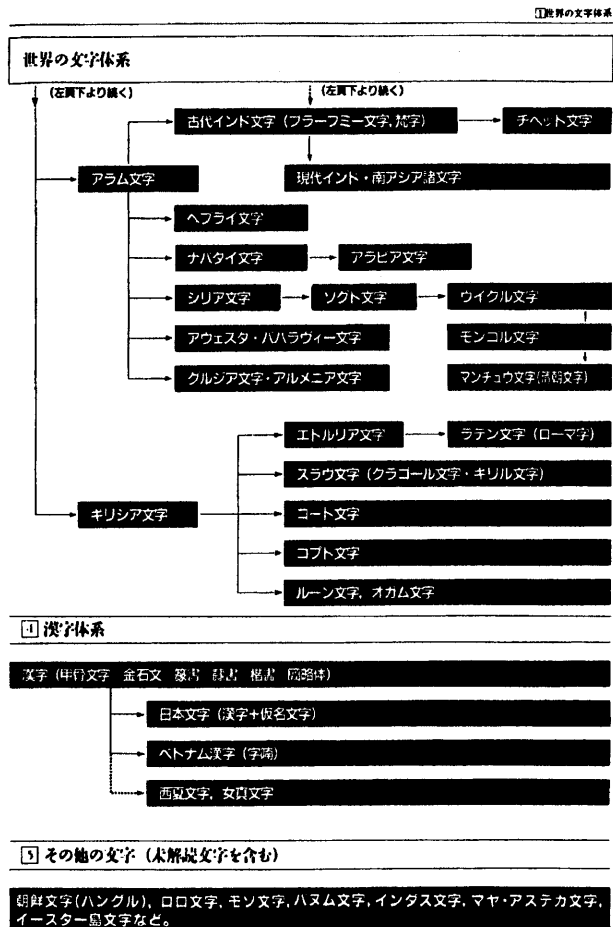
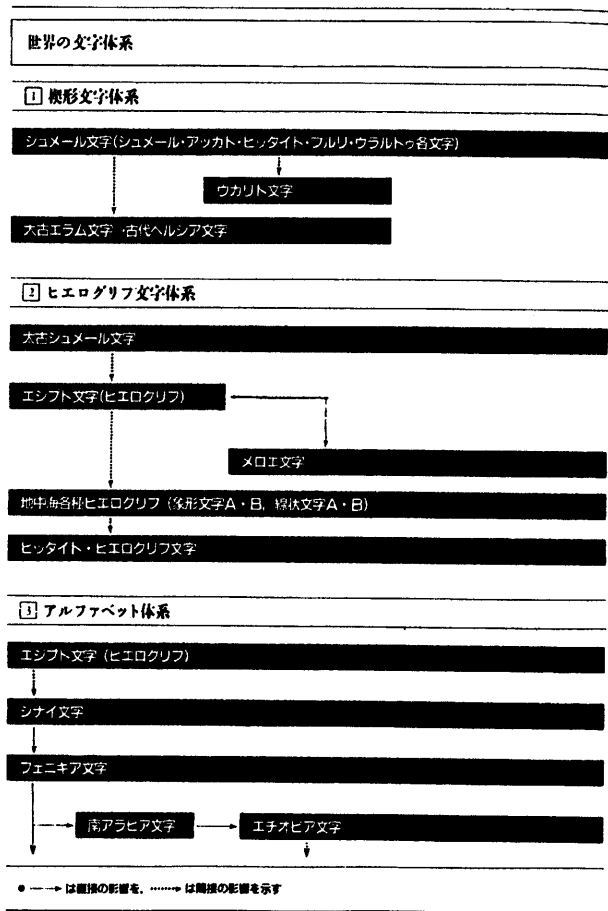
教育と文字(文字教育の意義の再検討)、(3) 科学技術の発達と文字(電子文字化の将来)などの問題がある。(1)と(2)はあまりにも大きな問題なのでここでは扱わず、(3)に関係のある技術論的問

題のいくつかを提示するにとどめたい。それらは(a) 書字方向、(b) 文字記号の「誤差範囲」、(c) 特殊記号問題、である。

(a) 書字方向は歴史上現われたブストロフェドン(「牛の歩み」、1行ごとに左>右、右>左)や渦巻き方式(「ファイストスの円盤」。これは今日のレコードやCDで再使用されている)を除けば、一般にはⅠ 左>右(ギリシア・ラテン系アルファベット体系)、Ⅱ 右>左(ヘブライ文字、アラビア文字など)およびⅢ 縦書き(縦組み)の上>下(伝統的漢字体系と、それに従う日本文字体系)の3種となる(縦書きの下>上は北アメリカのティフナグ文字で見られたが、今はすたれた)。Ⅲは若干他にやされ気味だが(縦書きの本場の中国でも文書の多くがⅠとなった)、Ⅱ(アルファベット文字の本来の書字方向)は今後も保持されるだろう。

横書きと縦書きの起源は、それぞれの書字材料と書字道具に関係している。竹簡・木簡から始まった漢字が縦書き用に作られたのに対し、メソポタミアで粘土が書字材料に選ばれたとき、最初は葦ペンで「縦に」(右行から左行に)書かれたが、ここでは粘土書板がすべて落ちてやすいので、100年ほどのちに横書き(左>右)になったのだと考えられている(この際、字形は90度回転したために、当初の象形性が弱まった)。古代エジプトでは左>右、右>左(文字記号が裏返しになる)自由であったが、後代には右>左が一般的になった。たぶんこの書字方向はフェニキア人、そしてヘブライ人、アラビア人に引き継がれ、今日にいたっている。しかしラテン文字の広範囲な普及と、欧米の政治的・文化的進展とともに、欧米語の書字方向(上記のⅠ)は全世界的スタンダードとなり、今日の電子文字化とも相まって、日本でも一般化しつつある(日常の書き文字の世界ではまだためらいがある)。

(b) 文字記号の「誤差範囲」というのは、文字の書体・字体が許容範囲の問題である。書き文字時代には弁別の特徴が認められる限り変種が許されたが、電子文字時代になると、機械が正しく読み取ってくれるかどうかの問題となる。それ以外にも、しばらく前に論争があったように、ハは



Aと同じか(アメリカのN/S\に発する問題)のような、歴史的意味合いを含む論議も生じている。活版印刷の時代には、かなり一定した字体が使われていたのに対し、ドットで構成された電子文字がほとんど自由に作られる時代になると、書き文字時代に逆行する面もないではない。

(c) 特殊記号というのは、基礎文字に対して付加された独立単位の各種の文字記号で、いわゆる句読点(日本のマル、テン。英文のピリオド、コンマ、コロンの、セミコロンなど)、カッコ、その他?、!などで、それらの用法が言語によって異なることが多い(英語とフランス語ではカッコの用法が違うし、?はスペイン語や古期ギリシア語では使い方が異なる)。日本文字は縦書きにも

横書きにも「同じ向き」のまま使えて便利だが、一(オンビキ)だけは縦書きでは|となるために、ワープロではこれだけのために1セットの変換装置を必要とするのだから厄介である。上記のコロンやセミコロンは、日本では対応するものがないので、正確に翻訳するためにはこれをどうしたらよいかは翻訳者たちの悩みの種だが、これは文字論というより翻訳文化論になるので、これ以上論じないことにする。

文字の世界史6000年には、まだまだ多くの話題があるが、日本文字の一(オンビキ)という小さくて大きな問題に焦点がしぼられたところで、この発表を終わることにしたい。

～文字シンポジウムから～

日本語の表記 The Japanese Writing System

石井 久雄* ISII Hisao

概要：日本語を書き記すには、漢字と出会ったときから現在に至るまで、漢字と仮名とをいかに折衷させるかという問題があった。どのような文章に漢字を用い、仮名を用いるか、混用するときにはどのように組み合わせるか、といった問題である。今日では、その問題の解決は、政策的に図られている。

日本語の表記について、現代の原理に立ち入ることを目標として、概要を述べます。現代に先立って歴史に触れ、その桎梏を脱しようとして打ち立てられた原理を述べる、という筋で、話を進めることにします。

1. 日本語の表記の歴史

1.1 日本は漢語・漢字の摂取に努めた

日本が初めて出会った文字は漢字であり、その文字が漢語（漢民族の言語）を記していたことは、疑いを入れる余地がありません。漢籍も仏典も、大陸から直接に伝えられるとともに、朝鮮半島を経由したものも多かるべく、古代日本における漢字の普及にも、朝鮮から帰化した人びとの力は大きかったでしょう。なお、漢字でない文字であって古代日本が触れたものに、仏典のうちの悉曇（古代インドの文字）があります。

漢語・漢字で書き記された文章を通して日本は漢文化を深く取り入れました。政治の運営などに用いる公式の文書などを、極力漢字で書き記そうとし、その志向は、以後、明治維新に至るまで、長く続くことになります。このことは、銘記しておいてよいことです。そのような状態を通して漢語・漢字から日本語が受けた影響も、甚大でした。

漢語に翻訳して漢字で日本のことを書き記しつつ、しかし、次のようなものは、漢語に翻訳し難いところがありました。

- ・人名・地名。
- ・助詞・助動詞。動詞・形容詞等の変化部分。
- ・和歌。擬音語・擬態語。

人名・地名の記録は、支配している人民・領地を掌握するための基礎であり、助詞・助動詞等は、命令をきちんと残すために記さなければならず、和歌は、韻律を明らかにしておく必要があった、というように、これらには、元の日本語をそのまま残す理由がありした。

漢字を用いて日本語をそのまま書き記すための工夫は、奈良時代をさかのぼって、早くから始められています。初めは、漢字を用いるしかありませんでしたが、しかし、その読みを重視し、対応している意味をほとんど無視しました。例えば、助詞「も」のために、モの音をもつ漢字「毛」を用いましたが、その漢字が対応している毛（け）の意味は無視しました。助詞「て」のためにテと読む漢字「手」を用いても、手（て）の意味は無視しています。

説明は省きますが、例を一つ挙げます。このようなものであるという印象だけをご承知ください。振り仮名は、後世のものです。

籠毛与 美籠母乳 布久思毛与 美夫君志持
このかまなつとみいのみまふとみとみ
 此岳爾 菜採須兒 家吉閑 名告沙根 虚見津
このかたのくにはおしなべつ家吉閑なつさねのつみ
 山跡乃国者 押奈戸手 吾許會居師 告名倍手
やまあとのかくに おしなべつ われつをりしつがし
 吾己曾座 我許者 背齒 告目 家呼毛名雄母
われこころし われこころし せなはに つばし いはちをりなつとみ

（万葉集巻第一、西本願寺本）

朝鮮でも、朝鮮語の助詞・助動詞や活用部分を記す工夫がなされ、吏読「りとう」または「りと」と言われます。日本で、日本語をそのまま残して

* 武蔵大学人文学部教授

書き記すための工夫が、吏読とどのように関係しているかは分かっていません。しかし、次の段階で、日本は朝鮮と分かれます。すなわち、この、日本語を書き記すために流用した漢字から、一部を取り出したり全体を崩したりして、片仮名・平仮名を生み出し、それを今日まで残します。これに対応する動きが朝鮮には見られません。

——朝鮮では、時代が下がって、ハングルを発明した。ハングルは、幾つかの要素を組み合わせて1音節を1字とする点で、漢字の影響を受けているが、要素の形が、発音のしかたを図示している趣である点で、漢字の影響を脱している。

片仮名・平仮名を生み出したところで、日本は、漢字を放棄するという道を進むこともできたはずですが。実際にそうしなかったことは、漢字による文章を志向したとして上に触れたとおりです。

なお、日本が漢語・漢字をどれほど摂取しようとしたかということについては、公式の文書などのほかに、さらに数点、挙げることができます。

- ・漢字一字一字について、字音を引き継ぎながら、字訓をほとんど与えた。
- ・漢語語彙を大量に借用し、また、それに準じて大量の造語をした。
- ・漢文訓読という、漢語に密着した翻訳様式を確立した。

これらについては、時間の余裕がありませんので、立ち入ることはしません。

漢語・漢字は、現代に至るまでの歴史を通して日本語に広く深く入り込んでしまっていて、いかにも日本語に適切な語彙・文字であるように見えます。しかし、日本語が漢語・漢字に出会ったのは、単なる偶然であると考えべきです。すなわち、漢語・漢字に支えられた強大な文明が、たまたま日本の近隣にあったということです。

歴史研究では嫌われますが、仮定を持ち込みますと、もし日本語がセイロン島にあったならば、サンスクリット系統の語彙とブラーフミー系統の文字とを取り入れ、イベリア半島にあったならば、ラテン系統とアラビア系統との語彙・文字を取り入れていた、のではないかと想像します。

1.2 漢字・片仮名と平仮名とはそれぞれに生きた

仮名には、万葉仮名と片仮名・平仮名とがあります。万葉仮名は、上で、日本語を書き記すために、意味と切り離して流用したとした漢字です。平安時代にも用いられました。その万葉仮名から、片仮名と平仮名とが、平安時代の初期に生み出され、万葉仮名に取って代わってゆきます。そうして、片仮名のほうは、漢字に寄り添う姿をはっきり見せ、平仮名のほうは、漢字と時に交じり合いながらも、むしろ独立して生きてきました。

漢字に寄り添う片仮名は、学問・政治や漢文芸など漢文化の世界のものです。それに対して、漢字から独立している平仮名は、和歌・物語など日本文化の世界に属していました。同じく日記と言われながら、公家のは漢字・片仮名で書き記され、女流のものは平仮名で書き記された、というようなこともあります。説話や戦記は、片仮名の世界に属します。

ここでも印象をお持ちいただくためということで、説明を省いて例を挙げます。次は片仮名の例であり、本文のものはここに記したよりは小さく、振り仮名のものは大きく、実際には両者ほぼ同様の大きさで書かれています。平仮名および濁点・句読点・数字は、補ったものです。

霜シモノ曉アカトキ雪ユキノノタニハ、坐イマセバチ父コ公シシ、不ズ。
 念おも二物モノモ、降フリ雨アメ、吹ふく風カゼ、時トキニモ、坐イマセバ、
 母はは氏ウヂシ、无な二愛うモレ。万マンノ物モノノ子コ、雖いへども、
 愛むツミガタシトト、父コ公シハ、不ズ宣のたま二愛はむツミ
 ガタシトトモ、无なキ一ノ利こどもモ子コ友トモ等トヲコソ、
 宣のたまケレ二愛ありとトハ、千チノ珍タカラ、雖ありと有あり一、母はは氏ウヂハ、
 不ズ宣のたま二珍たかくトモ、愚おろかに癡かたくなナル子コ友トモ等トヲコソ、
 宣のたまケレ二珍たかくトハ、云々。(9世紀前半、

東大寺諷誦文稿。中田祝夫の訓みによる) 次は平仮名です。濁点・句読点は補いました。

なほ、あかずやあらん、また、かくなん。
 みしひとの まつのちとせに みましかば
 とほくかなしき わかれせましや
 わすれがたく、ゝちをしきこと、おほかれど、
 えつくさず、とまれかうまれ、とくやりてん。

(10世紀前半成、土佐日記)

ふるとしに春の立ける日 在原元方
 としのうちに はるはきにけり ひととせを

こぞとやいはむ ことしとやいはむ

はるたちける日よめる 紀貫之
そでひちて んすびしみづの こほれるを
はるたつけふの かぜやとくらん

(1120 元永 3 年写, 古今和歌集元永本)

片仮名と平仮名とは、用いられる世界の対立もありますが、機能と言いましょか、音に対応するか語に対応するかという違いがあります。そのことを端的に表しているのは、仮名遣いとかかわりです。仮名遣いは、ある語をどう書き記すかということの問題とし、特に、平仮名で書き記される文芸などで、古典的にあるべき姿として扱われました。片仮名は、そういった問題からは解放され、一音一音を表すように用いられました。

—今日、外来語を記すのに片仮名を用いるのは、片仮名の、音と関係する性格による。

平仮名と片仮名とは、形そのものも対立的です。片仮名は直線的であり、平仮名は曲線的で、また、片仮名は、例えば「ア」については一通りであるなど、形が限られていて、早くから現代のものに近い姿を見せています。平仮名は、「あ」について、漢字「安」に由来するもの、「阿」に由来するものなどがあり、また、それぞれに崩しかたの違いによるものがある、といったように、いわゆる変体仮名ですが、多様です。平仮名は、連綿という続け書きの形も発達して、創作・鑑賞つまり書道の対象ともなりました。

—つい先ごろまで、電報やコンピュータで用いたのは、片仮名である。直線的な形による。

漢字・片仮名の世界と平仮名の世界とは、基本的には、江戸時代まで続きます。次に、平仮名の世界に属する江戸時代の読み物から、例を挙げます。漢字が多くはありません。濁点・句読点を補います。

今はむかし、かたいなかに金むらや金兵へといふ者ありけり。むまれつき心ゆふにして、うき世のたのしみをつくさんとおもへども、いたつてまづしくしてところにまかせず。

(1775 安永 4 年刊, 黄表紙金々先生栄花夢)

なお、平仮名の世界で書き記される言葉は、平安時代以来ほとんど変えられず、いわゆる文語の成立を見ることとなります。

しかし、漢字・片仮名と平仮名とは、用いられる世界について、相互に領域を侵し合いました。平安時代にも、片仮名で書き記された日本文芸があって、片仮名が平仮名のように独立的に用いられる端緒を開き、また、逆に、平仮名で書き記された文書があります。

江戸時代には、印刷が広く行われるようになり、様ざまな文章が世に出ました。社会全体で読み書きされる文字の総量が飛躍的に大きくなって、前の時代と一線を画していると考えられます。平仮名の世界に漢字をふんだんに取り入れることは、振り仮名を施した江戸時代の印刷と、関係が深いかもしれません。次がその例です。濁点を付けるのも、印刷と関係するようです。

けした所が恋のはじまり
桜もちるに敷き、月はかぎりありて、入佐
山、爰に但馬の国、かねほる里の辺に、
浮世の事を外になして、色道ふたつに、
寝ても覚ても、夢介と、かえなよばれて、

(1982 天和 2 年刊, 浮世草子好色一代男)

清和天皇七世の皇孫。鎮守府將軍陸奥守
義家朝臣の嫡孫。六条 判官 為義の八男。
冠者 為朝と聞えしは。智勇無双にして身の
丈七尺。材の目猿の臂。臂力 人に勝れて。
よく九石の弓を曳。矢継早の手鍛練なり。

(1807 文化 4 年刊, 椿説弓張月)

漢字・片仮名によって支えられる表記は、明治時代以降でも、法令や新聞について、われわれの知識にあるところです。今日、平仮名が漢字と絡み合った表記が、表記全般の基盤となり、公式の文書にも採られて、それに何の不思議も感じないとしても、ここ 50 年余りのことに過ぎません。

平仮名と漢字とによる表記は、漢字に関係する仮名が入れ代わったに過ぎないのではなく、むしろ、平仮名の世界に音読みの漢字が入ってくることを意味しました。それは、日本語のなかに漢語が入り込むということであり、あるいは日常の言葉に政治や学問の言葉が入り込むということでもあります。また、触れる余裕がありませんでした。送り仮名や振り仮名というものを、あらわに表に引き出すことにもなりました。

2. 現代日本における文字・表記

2.1 現代の方針は明治時代に議論された

現在の日本語の表記は、その基盤にある思想が、明治時代あるいは第二次世界大戦前に議論されています。その議論はその時代には必ずしも決着せず、改新しようとする主張に沿ったものが一般に実行されたとも言い難いのですが、第二次世界大戦後に一気に決着させられることとなります。

以下、述べることは、政策の解説が中心になります。しかし、それによって政策を擁護するといった意図はありません。ただ、理念とかの水準でなく、現在の実情の水準において、その実情は、政策が実現していると把握して、基本的に妥当であるのです。政策に強い反対意見もあり続けてきていることは、ここで立ち入りはしませんが、もとより心に留めておきましょう。

さて、明治に入る直前から、日本語を書き記す文字をどのようにするかということが、言い出されています。漢字を使用しないで仮名のみですませようという意見です。明治に入って、ラテン文字にしようという意見が加わります。

——前島密(1866)漢字御廃止之議、福沢諭吉(1873)文字之教、清水卯三郎(1874)平仮名ノ説、南部義壽(1872)文字ヲ改換スル議、西周(1874)洋字ヲ以テ国語ヲ書スルノ論、矢田部良吉(1882)羅馬字ヲ以テ日本語ヲ綴ルノ説、田中館愛橘(1885)本会雑誌ヲ羅馬字ニテ発兌スルノ發議及ヒ羅馬字用法意見、など。

文部省の方針は、漢字は使用を制限し、仮名は発音に沿って用いる、というものでした。帝国教育会の請願などもあり、文字・言語の改良のための調査を始めます。1902明治35年には、官制で国語調査委員会を發布し、この委員会は、発足3箇月余で、調査方針4箇条をまとめました。「調査」委員会であり、その調査方針ですが、調査を支える前提を含んでいます。

- ・文字は音韻文字「フォノグラム」を採用することとし、仮名・羅馬字の特質を調査すること。

- ・文章は言文一致体を採用することとし、之に関する調査を為すこと。
- ・国語の音韻組織を調査すること。
- ・方言を調査して標準語を選定すること。

つまり、文字については、仮名とラテン文字とのいずれかを採用することが、既定方針としてあったこととなります。

——帝国教育会(1900)国字国語国文ノ改良ニ関スル請願書、国語調査委員会(1902)調査方針ニ関スル報告、加藤弘之(1902)国語調査に就て。

この国語調査委員会は、1913大正2年に廃止になります。しかし、1921大正10年には、同様の趣旨で臨時国語調査会が官制で発布され、1934昭和9年には、文部大臣の諮問機関として国語審議会が発布されます。この国語審議会は、漢字および仮名遣いについて審議を重ね、1945昭和20年の終戦以降に答申をまとめあげます。1949昭和24年に建議機関として改組され、影響力を保ちます。

なお、明治時代に決着した問題もあります。片仮名・平仮名の字体が一つずつに定められたのは、1900明治33年、「小学校施行令規則」によります。書き記す言葉を口語に近づけるといふ言文一致運動は、1906明治39年、尾崎紅葉『多情多恨』が「である」体を完成させたことで、成功したことになります。

2.2 現代日本の表記は政策が支配している

現代日本における文字・表記について、顕著なことがらには、政策の課題として取り上げられ、決定した結果が、教育やマスメディアを通して、社会的に広く速やかに行われるに至るということです。具体的には、次のような内閣告示・訓令がありました。

- ・「当用漢字表」(1946年)
- 「常用漢字表」(1981年)
- ・「現代かなづかい」(1946年)
- 「現代仮名遣い」(1986年)
- 「外来語の表記」(1991年)
- ・「送りがなのつけ方」(1959年)

「送り仮名の付け方」(1973年)

以下、これらの要点を述べます。それによって、現代日本における漢字・仮名の実情の概要をとらえることになるでしょう。

第1に、「当用漢字表」「常用漢字表」は、一般社会で使用する漢字を限定し、義務教育で扱うものもその範囲のうちとすることになりました。それによって、古典漢語の強い影響から日本語を解放することになります。制限的色彩が、前者で強く、後者で弱いという違いはありますが、基本は変わっていないと見るのが穏当かもしれません。国語審議会答申「常用漢字表」の「前文」の「常用漢字表作成の経緯」に、次のようにあります。

当用漢字表等の内容に急激な変化を与えて混

イ. この表の漢字で書きあらわせないことばは、別のことばにかえるか、または、かな書きにする。

ロ. 代名詞・副詞・接続詞・感動詞・助動詞・助詞は、なるべくかな書きにする。

ハ. 外国の地名・人名は、かな書きにする。

ニ. 外来語は、かな書きにする。

ホ. 動植物の名称は、かな書きにする。

ヘ. あて字は、かな書きにする。

ト. ふりがなは、原則として使わない。

チ. 専門用語については、この表を基準として、整理することが望ましい。

乱を来すことのないよう留意しながら、今後とも一般の社会生活において相互の伝達や理解を円滑にするため、当用漢字表等に代わる一層適切な漢字表を新たに作成する必要があると考え、ここに常用漢字表を作成するに至った。

しかし、違いは大きいと見ることもできます。「当用漢字表」に、「使用上の注意事項」というのがあり、次の左側一ページの項目があります。また、国語審議会答申「常用漢字表」の「前文」には、字種と音訓の選定の主な考え方として、次の1-7の条条があります。後者は、前者との対比のために、順序を変えました。

- 1 使用度や機能度の高いものを取り上げる。
なお、使用分野の広さも参考にする。
- 2 使用度や機能度がさほど高くなくても、概念の表現という点から考えた場合に、仮名書きでは分かりにくく、特に必要と思われるものは取り上げる。
- 6 異字同訓はなるべく避けるが、漢字の使い分けのできるもの及び漢字で書く習慣の強いものは取り上げる。
- 5 代名詞・副詞・接続詞のためのものは広く使用されるものを取り上げる。
- 4 感動詞・助動詞・助詞のためのものは取り上げない。
- 3 地名・人名など、主として固有名詞として用いられるものは取り上げない。
- 7 いわゆる当て字や熟字訓のうち、慣用の広く久しいものは取り上げる。

「当用漢字表」の項目イについては、内閣官房長官から各省庁次官にあてた依命通知「公用文作成の要領」(1952年)に、具体例が示されました。「哺育」に対する「保育」,「洗職」に対する「汚職」などが、本来の漢字とは異なる漢字で普及し

たものです。

その項目イを含めてへまでは、漢字と仮名とのいずれで書き記すかという問題に、仮名で書き記す語を列挙するようにして、規定しています。それまでは、漢字・片仮名の世界にあるか平仮名の

世界にあるかで、同じ語でも別に書き記されていましたが、同じ語には一つの書き記しかたを与えようとしていることになります。その項目に対して、国語審議会答申「常用漢字表」では、漢字を用いる、または判断の対象としない、という方向を採っています。

漢字について、振り仮名を必要とするようなものは用いない、というのが、「当用漢字表」の項目トが言うところです。振り仮名を用いないことは、「当用漢字表」に結実した言語運動のスローガンでもありました。それに対して、国語審議会答申「常用漢字表」の「前文」は、次のように振り仮名に触れています。

常用漢字表は、上述のように一般の社会生活における漢字使用の目安となることを目指すものであるから、表に掲げられた漢字だけを用いて文章を書かなければならないというものではなく、運用に当たって、個々の事情に応じて適切な考慮を加える余地のあるものである。読みにくいと思われるような場合は、必要に応じて振り仮名を用いるような配慮をするのも一つの方法であろう。

この引用の前半を、「当用漢字表」の「使用上の注意事項」の項目イと比べてみるならば、使用する漢字の範囲を「当用漢字表」から解放することが、国語審議会答申「常用漢字表」の眼目であると知られます。極端な見かたをするならば、国語審議会答申「常用漢字表」のそのくだりは、「常用漢字表」自体からの解放をもうたっていることになります。

「当用漢字表」の「使用上の注意事項」の項目チに対しては、内閣告示・訓令「常用漢字表」の「前書き」に、次のようにあります。

この表は、科学、技術、芸術その他の各種専門分野や個人々の表記にまで及ぼそうとするものではない。

今日では、恐らく、「常用漢字表」の範囲を越えて、漢字は用いられているものと思われまふ。多くの方が、漢字「誰」「頃」を、「常用漢字表」にないとは知らずに用いているのではないでしょう。ワードプロセッサが普及したために、日本工業規格による漢字の制約のほうが、「常用漢字

表」の制約よりも強く意識されているであろうことは、近ごろの字種や字体の論議で感じられるところです。日本工業規格による漢字の制約とは言っても、常用漢字表を完全に包み込み、それに数倍する範囲のものであることは、言うまでもありません。

ところで、筋から言うならば、現実の使用を制約するための漢字表を定めるには、基礎的な語彙を先に見極め、後にその表記に必要である漢字を選定しなければならないところです。しかし、「当用漢字表」「常用漢字表」いずれにせよ、また日本工業規格のものにせよ、漢字選定では、直接に漢字そのものを対象としてきました。漢字がもっている表意性のために、そのような手続きでも実効性が見込まれることになったのかかもしれません。

なお、ついでをもって言うようなことではありませんが、仮名を平仮名によると規定したのは、「公用文作成の要領」です。法令の書きかたも含んでいて、次のような条条があります。

- ・かなは、ひらがなを用いることとする。かたかなは特殊な場合に用いる。
- ・文語体・かたかな書きを用いている法令を改正する場合は、改正の部分が一つのまとまった形をしているときは、その部分は、口語体を用い、ひらがな書きにする。

現代日本の文字・表記についての政策の第2に、「現代かなづかい」「現代仮名遣い」は、述べる形式が異なりますが、結果としての仮名の遣いかたには違いがありません。前者は、現代語の発音を参照しながら、歴史的仮名遣いをどう改めるか、という形式を採っています。改まらない部分については、触れていません。後者は、現代語でこう発音しているものをどのように書き記すか、という形式を採っています。

「現代かなづかい」「現代仮名遣い」は、和語および漢語借用語を対象としています。英語などから大量に入ってきている外来語には、必ずしも対応し切ることができていません。「外来語の表記」は、特に、和語・漢語になかった発音「ファ」や、外来語のために作られた仮名「ヴ」に留意して、定められました。

「ヴ、ヴァ」などを「ブ、バ」などと聞き分け

ることは、日本人にはできません。したがって、仮名「ヴ」を使おうとする限りは、それを含む語を覚えなければなりません。特定の語のために特別の仮名を用意する点で、助詞「を」のために仮名「を」を用いるのと同様であると見えますが、新たな語が加わり続ける点で、異なります。仮名遣いとしては、歴史的仮名遣いと同じところに戻ったと言うことができるでしょう。

現代日本の文字・表記についての政策の第3に、「送りがなのつけ方」「送り仮名の付け方」は、日本の表記のうちで恐らく最も難しい問題を扱ったものです。平安時代以来、漢字・片仮名の世界では総体的に送り仮名が少なく、平仮名の世界で用いられた漢字も、例えば「申」一字で「まうし」でも「まうす」でもあるように、動詞の連用形や終止形は送り仮名を伴わないのが普通である、といったことがありました。一語については、漢字の部分の読みを共通させる、というのが、現代の原理です。「申」は「まう」にのみ対応させることになります。

先行する「送りがなのつけ方」と後行する「送り仮名の付け方」との間では、例えば「行なう・行う」「表わす・表す」のような違いがあり、後行するほうで送り仮名が少なくなります。先行するほうでは、別の語に読み誤ることがないようにすることを重視し、後行するほうでは、別の語については別に考えるという方針を採っています。

ラテン文字については次にも触れるところがありますが、ローマ字の綴りかたについて一言します。現在、学校教育では、「シ、ツ」を“si, tu”と綴るような訓令式を中心に教えています。しかし、社会的には、“shi, tsu”と綴るようなヘボン式が通用しています。第二次世界大戦前、鉄道省は、訓令式で統一しました。戦後、ヘボン式を採っているのは、連合国占領軍総司令部の指令によって改めてからです。

2.3 これからのために考えるべき問題もある

現代ないしこれからの日本における文字・表記の問題として、考えるべきことに、二三触れます。

まず、文字の種類として、現代日本には、漢字・

仮名のほかに、ラテン文字が加わっていることを指摘すべきでしょう。ラテン文字は、日本語をローマ字で記す、例えば電車の駅名のみでなく、符号や、外国語を記すものとしても、多く見ることができます。例えば、

列記	a, b, c
符牒	1年A組 X氏
略号	JIS NHK
略語	etc. bldg.
語句	商標 店舗名 スポーツのユニフォームの人名
詞章	ニューソング・広告等の詞章

今後、こうしたものを無視することはできないでしょう。漢語・漢字に出会ってそれを吸収したように、日本語は、これから、ラテン文字を通して、英語等を摂取し続けることになるのではないかと思います。

次に、現代の文字が古代のものと異なる、その著しい点として、符号に補助されながら記されることがあります。濁点・半濁点は、文字そのものに加えられ、句読点が文に与えられ、段落を示す空白が置かれます。括弧も用いられます。このような符号と文字とをどのように共存させるか、考えることも、これからの一つの課題としてよいかもしれません。日本語では、漢字の連なりの始めが語あるいは文節の始めを示し、ラテン文字の文章の分かち書きと対等のことが実現していると、よく言われますが、仮名書きをする語が多くなると、仮名の連なりの中にも多分に語の切れ目がかかることになります。

縦書きにするか横書きにするかという問題、着色したり変形したりするといったデザインの問題も、文字にはあります。

最後に、固定されない文字ということについて、一言します。次つぎに発されては消えてゆく音声言語を、文字は固定する。文字自体が紙や石に固定されつつ、言語を固定する。というのが、文字についての一般的な了解事項です。しかしながら、今や、文字にも、流動的な使用法が広がっています。映画やテレビのスーパー、テレビのフリップ、文字放送、コンピュータのモニタ、街頭の電光掲示など、極めて多様に、とどまっていない文字を

見つけることができます。そこには、声に出してはできない通報の効果があります。このような文字には、読み取りのしやすさが求められます。

終わりに

日本では、仮名を生み出して以来、漢字の世界と仮名の世界とをそれぞれに築いてきた。現代に至って二つの世界を融合させることになり、種々の問題に直面することになった。というのが、ここに述べたところです。歴史と現代とをつなぐのがうまくいっていませんが、お許しください。

このお話の趣旨は、日本語を書き記す文字について、漢字を特に意識しながら、概要を述べることにありましたので、漢語借用語については、ほとんど触れることをしませんでした。しかし、文字をどのように理解するかということは、語彙の理解なくしては、困難です。日本の文字・表記の歴史をたどりながら、「漢語・漢字」と並べるのを繰り返したのには、その意図がありました。

全体にわたっていろいろと省いたことがありますが、このお話の水準で参考となるような書目を添えます。

参考書目

- 網野 善彦 (1988) 日本の文字社会の特質。
= 網野 (1993) pp.353-404。
—— (1993) 日本論の視座——列島の社会と国家。小学館ライブラリー。
- 池上 禎造 (1982) 表記の歴史から見た現代語表記法。森岡 pp.1-18。
- 岩淵 匡 (1982) 片仮名の機能の歴史。森岡 pp.112-134。
- 遠藤 邦基 (1982) 平仮名の機能の歴史。森岡 pp.84-111。
- 遠藤 好英 (1982) 漢字仮名交じり文の流れ。森岡 pp.159-185。
- 大坪 併治 (1977) 片仮名・平仮名。大野ほか pp.249-299。
- 大野 晋・柴田 武 (1977) 文字。岩波講座日本語 8。林, 大坪, 鶴を含む。
- 樺島 忠夫 (1979) 日本の文字——表記体系を考える——。岩波新書。

- 亀井 孝 (1989) 日本語の歴史。
= 亀井ほか (1997) pp.107-153。
- 亀井 孝・河野 六郎・千野 栄一 (1997) 言語学大辞典セレクション 日本列島の言語。三省堂。亀井, 小松, 佐竹を含む。
- 小松 英雄 (1989) 日本語の歴史 書記。
= 亀井ほか (1997) pp.172-178。
- 佐竹 秀雄 (1989) 現代日本語 文字・表記。
= 亀井ほか (1997) pp.235-242。
- 築島 裕 (1969) 平安時代語新論。東京大学出版会, 東大人文学研究叢書。
—— (1981) 仮名。中央公論社, 日本語の世界 5。
—— (1987) 平安時代の国語。東京堂出版, 国語学叢書 3。
- 鶴 久 (1977) 万葉仮名。大野ほか pp.209-248。
- 中田 祝夫 (1982) 日本の漢字。中央公論社, 日本語の世界 4。
- 野村 雅昭 (1982) 漢字の機能の歴史。森岡 pp.135-158。
—— (1988) 漢字の未来。筑摩書房。
- 浜田 啓介 (1982) 出版と文字の歴史。森岡 pp.208-227。
- 林 史典 (1977) 日本における漢字。大野ほか pp.159-208。
- 文化 庁 (1992) 公用文の書き表し方の基準 (資料集) 増補版。第一法規出版。
- 馬淵 和夫 (1971) 日本語の表記体系。
= 馬淵 (1996) pp.33-50。
—— (1975) 表記の変遷。
= 馬淵 (1996) pp.51-71。
—— (1996) 国語史叢書。笠間叢書。
- 森岡 健二 (1982) 現代表記との史的対照。明治書院, 講座日本語学 6。池上, 遠藤邦基, 岩淵, 野村, 遠藤好英, 浜田を含む。
- 森岡 健二・柴田 武・山田 俊雄・樺島 忠夫・野村 雅昭 (1975) 日本語の文字。学生社, シンポジウム日本語 4。

～文字シンポジウムから～

漢字の「構造」を考える ——コンピュータの入力・検索の視点から

松岡 榮志・ MATSUOKA Eishi

0. はじめに

1. 縦軸から考える——「六書」・「部首」・「筆画」／「筆順」
2. 横軸から考える——コンピュータの入力・文字合成
3. おわりに

0. はじめに

漢字は、「形・音・義」を持っています。

ここで、漢字の「構造」について考えるのですが、主に「形」の部分について話を進めることになります。ただ、あらかじめ注意していただきたいことがあります。

漢字は、たしかに数千年前に中国で作られ、日本をはじめアジアの各国に伝えられました。そして、その国のことばを書き表す手段として長く使われてきました。そこで、中国ではどちらかというと、「音」が意味や用法の中心であったのにひきかえ、日本では「形」に重きが置かれてきました。そこで、今日、漢字について議論する場合、双方の姿勢にかなりの違いがあります。

私のこれからのお話でも、どちらかという中国古典学の見方が前提になります。それは、あるいは現代の普通の日本人の感覚からは若干ずれるかも知れません。その点をあらかじめお断りしておきます。

1. 縦軸から考える

——「六書」・「部首」・「筆順」

1. 1 「六書」

漢字の「構造」について考えると、まず頭に浮

かぶのは「六書」です。

これは、後漢の許慎がA.D.100年頃に著した『説文解字』に見られるものです。

- ①象形 「日」「月」「馬」「鳥」のような線で事物のイメージを表す。(約4%)
- ②指事 「一」「上」「本」のように抽象的な記号で表す。(約1.1%)
- ③会意 「休」「鳴」「武」のように、①②などを組み合わせて意味を表す。(約12.4%)
- ④形声 「湖」「詩」のように、「義(意)符」と「声(音)符」を組み合わせて意味を表す。(約82%)
- ⑤転注 語義の転化で、注釈する方法。
- ⑥仮借 「豆」のように、同音の当て字として使う。

()の中の数字は、『説文解字』に収められているパーセントですが、「形声」が大部分であることがわかります。ただ、①②③④は造字法ですが、⑤⑥はその転用法にすぎません。

「形声」文字を考える上で、いくつか注意すべき点があります。

- a) 声符は、必ずしも字全体の音と対応しない。
- b) 声符が意味の核となることがある。

まず、a) についていえば、二千年前にはそれなりの対応があったと見られますが、現代の中国語では、ほとんど対応していません。『現代漢字形声字字彙』(倪海曙編, 文字改革出版社, 1975)

* 東京学芸大学教授

☆ISO/IEC JTC1/SC2およびWG2/1RG委員

☆☆ISO/IEC 10646-1の制定・改訂に参加

に収められる5990字をもとにした調査では、対応する漢字は26.3%にすぎません。つまり、声符から類推しても、26.3%しか当たらないのです。日本語の漢音の場合は、かなりの確率で一致すると思われそうですが、現代中国語音ではかなり低い確率になってしまいます。^{*1}

たとえば、長江の「江」を見てみましょう。日本語では、声符が「工(こう)」で字そのものの音も「こう」です。ところが、中国語では、「工(gong)」が「江(jiang)」になってしまいます。

次に、b)です。「包」という声符を持ついくつかの字「飽」「抱」「泡」などには、意味の上でも共通性があります。それは、「包」そのものにも見られる「中にたっぷりと含んだ」イメージです。「飽」は、腹いっぱい食べて膨らんだようす。「抱」は、両手でふんわりと抱きかかえること。「泡」は、シャボン玉のようにふくらんだ泡です。このように、声符はしばしば意味を担っています。漢字は、よく「表意文字」だと言われますが、それは当たってはいません。むしろ「表語文字」とでも呼ぶべきでしょう。

1.2 「部首」

次に、部首について考えてみましょう。

今日、私たちが漢和辞典などで慣れ親しんでいるのは、『康熙字典』に代表される214部首です。(図①, P.18)ただ、歴史的に見ると、部首の数は一定ではありません。先ず挙げた『説文解字』では、540部。梁代の顧野王(519-581)の『玉篇』では、542部。明代の『字彙』になって214部となり、『康熙字典』に引き継がれました。わが国の『大漢和字辞典』をはじめ、ほとんどの漢和辞典は子の214部によって漢字を分類しています。ただ、現代の中国では『新華字典』は189部、『漢語大字典』は200部に整理しています。

この部首による分類は、漢字を大きく二つに分けて、部首の部分とその他の部分とするものです。何よりも検索のために使われますが、漢字構成の考え方の一つと言ってよいでしょう。ついでに言えば、「偏旁」という考え方があります。これは、

あくまでの漢字を左の部分(「偏」)と右の部分(「旁」)に分けて考えるものです。

1.3 「筆画」／「筆順」

「筆画」とは、漢字を点や線に分解したものを言います。それらを組み合わせる順序が、「筆順」です。「筆順」は、左から右へ、上から下へ、が原則になっています。「筆順」を重視するのは、まず漢字を美しく書くためですが、同時に漢字を検索するためにも便利なためです。

2. 横軸から考える

——コンピュータの入力・文字合成

2.1 入力法

現代の漢字入力には、大きく言って二つの方法があります。

a) 音音入力

b) 字形入力

もちろん、それらを組み合わせたものもあり、多種多様ですが、ここでは上の二つを基本として考えておきます。ちなみに、「コード入力」もあるのではないかと、という声が聞こえてきそうですが、これもやはりコードを何によって振るのかを考えれば、上の二つに行き着いてしまいます。たとえば、日本のJISの「第一、第二水準」は、「第一水準」が音訓(a)順、「第二水準」が部首(b)順に並べられています。これから見ても、やはりこの二つが軸になります。

(b)の字形入力を、もう少し細かく見てみると、

①部首

②部件(component)

③筆画

に分けられます。従来の漢和辞典は、部首に画数を加味して、すべての収録文字を配列しています。しかし、索引となると、「部首一覧」「音訓索引」「絵画索引」をつけなければなりません。日本では、たいてい「音訓索引」が最もよく使われています。JISの場合、「第二水準」が部首順であるため、その区点を探すための「ワープロ漢字辞典」なるものが多く作られ、よく利用されました。

今日では、むしろ②の部件(component)を組

*1 張静賢『現代中国漢字学講座』(三省堂, 1997.8) p.80

み合わせて検索することが一般的になってきました。そこで、「ワープロ漢字辞典」の出番が少なくなっています。「部件」という用語は、ちょっと聞き慣れないでしょうが、部首に残りの偏旁を加えたもの、と考えて下さい。

これは、文字を分解する時に、いくつかのモデルを立てるか、にかかわります。さらに、一万字以内の漢字を扱うのか、五万字の漢字を扱うのかによって、大きな違いが生じます。

たとえば、五万字の漢字を、上下・左右などの二分画法で扱おうとすれば、部件の数は数千になります。とくに、組み合わせたときの文字としてのバランスをも要求するとすれば、七、八千個になってしまうでしょう。部件がこんなに多くては、分解する意味がありません。図②は、現在、私たちがIRGで検討しているモデルの一つです。これは、主に中国が提案しているものですが、13のモデルと約二千個の部件で、五万字が表示、検索できます。ただ、やっかいなのは、文字の分解のしかたが人によってしばしば異なることです。したがって、これも最終的な結論をみてはいません。

③は、筆画によって入力する方法です。いわゆる「手書き入力」というのも、この方法の応用と考えてよいでしょう。②の部件をもっと細かくして、このレベルで行おうという考えもあります。その場合、組み合わせのルールがさらに複雑にな

ります。

実際のところ、この方法をコンピュータの文字検索に使うのは、かなり困難です。これは、中国や日本のように漢字を現在でも使っている地域では、下から上へ、左から右へという筆順が小、中学校で教えられているため、かなり有効でしょうが、その他の地域では相当な困難が予想されます。

3. おわりに

今日、漢字を考える上で是非とも注意しておきたいことが、いくつかあります。

- 1) 漢字をアジアや世界の中で考えること
- 2) 構造を縦軸と横軸でとらえる（どちらかに偏重しないこと）
- 3) 使い道に応じて、categorizeされた文字セットを使うこと
- 4) なるべくゆったりと落ち着いて議論すること

(参考文献)

- 1) 林・松岡『日本の漢字・中国の漢字』（三省堂、1995.7）
- 2) 石川・松岡『漢字とコンピュータ』（大修館書店、1997.6）
- 3) 松岡榮志「漢字の危機は杞憂に過ぎない」（『中央公論』1998.3）


康熙字典部首索引

几	ノ	一	冂	八	入	儿	人	亠	二	丨	乙	ノ	、	一	部首	畫數	總頁	
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	
六	九	九	九	九	九	九	九	九	九	九	九	九	九	九	三	三	三	
又	士	士	口	口	又	厶	冂	冂	卜	十	匕	匕	力	刀	冂	冂	冂	
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	
七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	
广	宀	干	巾	己	工	山	山	尸	尤	小	寸	子	女	夕	夕	夕	夕	
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	
七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	七	
女	支	支	手	戶	戈	心	心	卩	彳	彳	彳	彳	彳	彳	彳	彳	彳	
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	
四	九	九	九	九	九	九	九	九	九	九	九	九	九	九	九	九	九	
爻	父	爪	火	水	气	氏	毛	比	母	爻	止	欠	木	月	日	日	无	
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六
正	田	用	生	甘	瓦	瓜	玉	女	女	王	王	王	王	王	王	王	王	
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六



糸	米	竹	立	穴	禾	内	冂	石	矢	矛	目	皿	皮	白	冂	冂
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六	六
虫	虺	艸	色	艮	舟	舛	舌	白	至	自	臣	肉	聿	耳	来	而
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
一〇〇	一〇〇	九	九	九	九	九	九	九	九	九	九	九	九	九	九	九
邑	疋	辰	辛	車	身	足	走	赤	貝	豸	豕	豆	谷	言	角	見
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
二	二	二	二	二	二	二	二	二	二	二	二	二	二	二	二	二
飛	風	頁	音	韭	韋	革	面	非	青	雨	隹	隹	隹	隹	隹	隹
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
三	三	三	三	三	三	三	三	三	三	三	三	三	三	三	三	三
黍	黃	麻	麥	鹿	鹵	鳥	魚	鬼	厲	鬲	門	影	高	骨	馬	香
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇	一〇
考	備	楠	俞	龜	龍	齒	齊	鼻	鼠	鼓	鼎	鼎	鼎	鼎	鼎	鼎
一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一	一
一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六	一六

圖 1



1 単独字構造

構造図形は 。字例：一，人，厂，七，又，工，月，马，子。

2 左右構造




- 2.1 左右構造：構造図形は 。字例：代，江，吹，张，妈，经，捉，松，轮，骄。
- 2.2 左中右構造：構造図形は 。字例：湖，滩，脚，谢，涨，绑，撕，糊，撒。

3 上下構造




- 3.1 上下構造：構造図形は 。字例：男，雷，雪，集，蛋，家，竟，罢，费，香，点，昏，爸，昌，苗，朵。
- 3.2 上中下構造：構造図形は 。字例：莫，宽，葱，鼻，裹，哀，舅，葬。


4 包む構造

4.1 二面を包む構造










- 4.1.1 上と左を包む構造：構造図形は 。字例：反，压，庆，虎，尾，虑，厠，庙，店，厘，仄，病，疾，疯，唐，居，属，痰，愿，鹰，扇，层。
- 4.1.2 上と右を包む構造：構造図形は 。字例：包，司，氦，式，武，或，戒，句，旬。
- 4.1.3 左と下を包む構造：構造図形は 。字例：边，过，达，进，运，还，道，起，赵，超，毯，翅，处，建，延。

4.2 三面を包む構造

- 4.2.1 左上右を包む構造：構造図形は 。字例：同，冈，网，周，问，间，闻，闹，阅，闪。
- 4.2.2 上左下を包む構造：構造図形は 。字例：区，巨，匠，医，匹，匪。
- 4.2.3 左下右を包む構造：構造図形は 。字例：凶，画，函，击。

4.3 四面を包む構造：構造図形は 。字例：困，因，回，园，围，困，国，囚，圆，固。

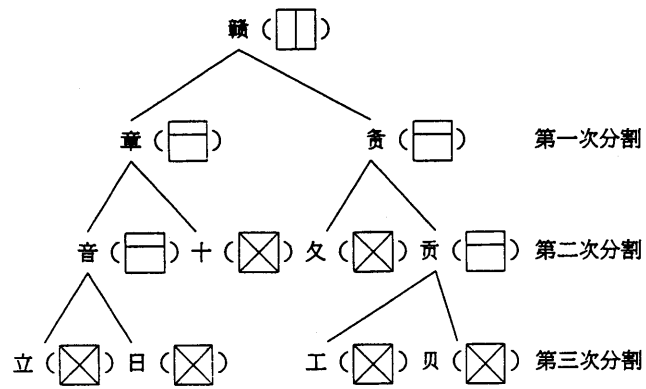
5 特殊構造 字例：坐，乘，噩，爽など。

ここにあげた5類13種の構造形式は、現代漢字の字形構造の基本的なモデルを示したに過ぎない。もっと細かく示すなら、さらに細かい分類をすることが可能である。たとえば上下構造は、 (想)， (霜)， (率)， (器)などに細分され、左右構造も (部)， (婚)， (修)， (衙)， (毀)などに細分される。

張静賢『現代中国漢字学講義』

PP.57-58

図2



1983年1月から1984年5月まで、『辞海』（1979年版）に収められる16296字と、『辞海』では未収だがGB 2312-80 国家標準「信息交換用漢字編碼字符集・基本集」に収められる43字の、計16339字について、中国文字改革委員会と武漢大学が共同して、コンピュータによる部件の分析を行った。そのなかの11834字の簡化字と未簡化字について段階ごとに行った部件の一覧表は、以下のとおりである。

分割の段階	部件総数
第一段階	2556
第二段階	1149
第三段階	477
第四段階	168
第五段階	39
第六段階	9
第七段階	2

張静賢『現代中国漢字学講義』
PP.59-60

〔参考〕

DATE:1996-06-27

WG2 N

ISO-IEC JTC1/SC2/WG2

IRG N409A

TITLE: Ideographic Structure Characters
 SOURCE: IRG
 ACTION: For Inclusion in BMP by WG2

NO.	SYMBOL	CHARACTER NAME	SHORT	CODED EXAMPLE	NOT Coded EXAMPLE	
1		IDEOGRAPHIC STRUCTURE CHARACTER LEFT-TO-RIGHT	LTR	休明	傅= 冂イ 母	
2		IDEOGRAPHIC STRUCTURE CHARACTER ABOVE-TO-BELOW	ATB	思杏	矣= 冂八天	
3		IDEOGRAPHIC STRUCTURE CHARACTER LEFT-MIDDLE-RIGHT	LMR	弼班	衙= 冂彳言于	
4		IDEOGRAPHIC STRUCTURE CHARACTER ABOVE-MIDDLE-BELOW	AMB	裏奚	咎= 冂从从日	
5		IDEOGRAPHIC STRUCTURE CHARACTER OVERALL-AROUND	OVA	OTI - Outside_To_Inside	囿= 冂口巷	
6		IDEOGRAPHIC STRUCTURE CHARACTER DOWN-TO-ENCOMPASS	DTE		同冈	閑= 冂門卞
7		IDEOGRAPHIC STRUCTURE CHARACTER UP-TO-ENCOMPASS	UTE		凶函	由= 冂口士
8		IDEOGRAPHIC STRUCTURE CHARACTER RIGHT-TO-ENCOMPASS	RTE		匠区	匯= 冂口虎
9		IDEOGRAPHIC STRUCTURE CHARACTER RIGHT-DOWN ENCOMPASS	RDE		病原	靡= 冂广舞
A		IDEOGRAPHIC STRUCTURE CHARACTER LEFT-DOWN ENCOMPASS	LDE		司句	甸= 冂勺去
B		IDEOGRAPHIC STRUCTURE CHARACTER RIGHT-UP ENCOMPASS	RUE		毯建	这= 冂辶交
C		IDEOGRAPHIC STRUCTURE CHARACTER ENBEDMENT	EMB		夹巫	

～文字シンボジウムから～

脳 と 文 字

杉下 守弘* SUGISHITA Morihito

文字には種々のものがあるが、ここでは日本語に特有の漢字と仮名を取り上げ、「脳と漢字・仮名」について論を進めたい。「脳と漢字・仮名」に関する研究は3つに分類できる。第1は、脳損傷で生ずる失語症の研究である。脳が脳卒中などのため特定の部分が破壊されると漢字と仮名の解離が生ずる。漢字と仮名の解離とは、漢字の方が仮名よりも障害されたり、その逆に、仮名の方が漢字より障害されることはないかといった方面の研究である。第2は、脳が損傷されていない通常の人で、漢字を使用しているときに脳のどの部分が活動するのか、また、仮名の場合はどうかなどの研究である。これは機能的脳画像による研究と呼ばれる。第3は、漢字あるいは仮名の学習や使用が脳に特別な影響を与えているかどうかを検討するものである。この種の研究は、漢字教育と関連のあるため、研究の推進が望まれる研究である。はじめに脳損傷による漢字と仮名の解離からのべる。

1. 脳損傷による漢字と仮名の解離

脳損傷による漢字と仮名の解離は、音読、読解、単語の指示、書字などの際に調べることができる。単語の指示とは、検査者が言った言葉に相当する単語（文字で書かれた単語）をいくつかの単語の中から選ぶことである。

脳損傷による漢字と仮名の解離は、音読について最も研究されており、読解や単語の指示では研究が少ない。ここでは音読における漢字・仮名の解離についてのみ論ずる。脳損傷は左大脳半球損傷による失語症と脳深損傷による左視野失語をとるあつかう。

1.1 漢字及び仮名の音読における解離

——左半球損傷による失語症の場合——

漢字と仮名の音読における解離については、明治以来多くの研究がなされて、多くの説が述べられてきた。たとえば、左半球の後方の損傷で、漢字の音読が障害され、左半球の前方の損傷で仮名の音読が障害されるという説 (Benson, 1985) がある。また、左半球の側頭葉後下部の損傷で「漢字の音読」が障害され、左半球の角回の損傷で「仮名の音読」が障害されるという説もある (岩田 1992)。

たしかに、これらの説を支持すると考えられる研究はあるが、それらの研究のほとんどは、検査として使用されている漢字や仮名がどのようなものであったのか述べられていなかったり、各症例の設答数が述べられていなかったり、あるいは統計処理に問題があったりといった不備が多い (Sugishita 1992, 杉下 1993)。そのうえほとんどの研究が1例か2例の症例に関する研究であり、その研究結果は漢字・仮名と大脳損傷部位の関係について一般的な結論を出すのに十分とは言い難い。

我々は多数例を対象とし、検査する漢字・仮名も適当な基準で選び、十分な統計処理をする研究を行った (Sugishita et al. 1992)。

この研究では、46文字の清音を表わす仮名と46文字の漢字を用いた。46文字の漢字のうち18文字漢字は小学校で教えられ国定教科書の第2版 (1910-1911) から第五版 (1940-1950) までに行っているものである。他に28文字の漢字は上述の国定教科書の2年生で教えられた77の漢字の中からランダムに選ばれた。失語症患者の脳損傷部位はコンピューター断層撮影法で固定した。

1. 左半球前方の損傷で生じた Broca 失語 23例

* 東京大学医学部研究科 認知・言語神経科学教授

を検討したところ、仮名の音読が主に障害されていたのは4例あったが、漢字の音読の方が主に障害されている例が3例みられた。他の16例は漢字と仮名で差がなかった。

左半球の後方の損傷で生じた7例のウェルニッケ失語では、漢字の音読が仮名よりも障害されている例が1例あったが、他の6例では漢字と仮名で差がみられなかった。したがって Benson の説が支持されなかった。

左側頭葉後下部の損傷が漢字の音読が障害されると説を検討したところ、我々のデータでは左側頭葉後下部損傷例2例は、2例とも漢字・仮名の音読の差を認めなかった。左角回損傷5例のうち1例のみが仮名の音読が漢字の音読より障害されていた。残りの4例は漢字と仮名の差は認められなかった。したがって、漢字の音読を障害するのは、左側頭葉後下部とする説や仮名の音読を障害するのは左角回損傷とする説は問題がある。正しくは、左側頭葉後下部損傷では、漢字の音読が仮名の音読よりも障害される例 (Mochizuki & Ohtomo 1988 の1例, 坂井・他 1992 の1例) もあるが、漢字と仮名の音読で差がない例 (Sugishita et al 1992 の2例) も存在するということであろう。左角回損傷では仮名の音読が漢字の音読より障害される例 (Yamadori 1979 の1例, Sugishita et al 1992 の1例) もあるが、漢字と仮名の音読障害で差がないもの (Sugishita et al 1992 の4例) の方が多いと思われる。

1.2 脳梁損傷例における漢字と仮名の音読

大脳は右半球と左半球から成り立っており、この2つの半球は初級線維束である脳梁によって連結されている。脳梁の後部を切断すると、左視野失語すなわち左視野の文字が読めなくなる症状が生ずる。これはどのようにして生ずるかという点、左視野の文字は右半球に伝達されたあと、音読するには言語優位半球である左半球に伝えられなければならない。ところが、脳梁後部の損傷のため、右半球に伝達された文字情報が右半球に伝えられない。その結果、左視野の文字が読めなくなる。この左視野失語では仮名の音読が主に障害される (Sugishita 1978, 1986)。この結果から、漢

字は仮名よりも右半球の後部損傷より前部を介して右半球に伝達されることがわかる。これは漢字が仮名よりも混同しにくいからと考えられる (Tanaka 1991)。もう1つの理由は、漢字は仮名よりも右半球で読解されやすく、このため仮名よりも漢字は右半球をより広く活動させるためと思われる。

2. 機能的脳画像法による漢字仮名の研究

健康な人が漢字や仮名を処理する時、脳のどの部分が活動するのかを研究できるようになったのは、1980年代に入って、陽電子放出断層撮影法 (略称 PET) が登場してからである。その後1990年代後半に入ると、安全でより精度の高い方法として、機能的磁気共鳴画像法 (略称機能的 MRI) が導入された。我々は漢字で1字 (視覚性複雑度7, 8画) と仮名 (7, 7画) で2字の教育漢字39語を用いて、漢字と仮名の音読の際に、1, 5テスターの装置により機能的MRIを測定した。仮名の音読で賦治され、漢字の音読では賦治されない部位が4例中3例にみられた。その部位は右側頭葉であった。この結果がどう解釈されるべきものか、今後検討を加えていきたい。機能的脳画像の研究は記憶などの分野では大きな成果をあげている。機能的脳画像法による漢字・仮名の研究はまだ端初についたばかりであるが、今後大きな発展が期待されている。

3. 漢字学習が脳に及ぼす影響

日本人は小学校、中学校及び高等学校で約2000の漢字を学習する。漢字はそれぞれ形が異なるので、2000の形を記憶しなければならず、大変な努力が払われる。こうして、憶えた漢字は新聞などの印刷物を通じ、毎日我々の目にふれる。ちなみに、日本は世界中で最も新聞の普及率が高く58%にのぼる。米国では20%である。このような漢字の学習と使用は、日本人の脳に何らかの特別な影響を与えているのであろうか？ 漢字の学習と使用の経験から考えると、視覚性の記憶に何らかの影響があり、記憶に関連する脳の部分に影響が及び可能性がある。

我々は1992年より日本人の記憶を16才から74才まで300人以上にわたり、検査し、米国人の記憶を比較することを試みた。

その結果、日本人は米国人と数唱、論理的記憶、言語性対連合、視野性対連合、視覚性記憶範囲などではあまり差がないにもかかわらず、専門学的図形の想起において、著しく成績（1SD以上平均値が高い）がよいことが明らかになった。

読字能力の高い児童は幾何学的図形の記憶がよいことと考え合せると、上述の結果は漢字の習得の効果による可能性が高いと推定している。また、以上の結果から、幾何学的図形の想起により脳のどの部分が活動するか、脳部分は日本人は他の民族にくらべ大きいかなどの問題を現在検討中である。

参考文献

- BENSON DF(1985) Alexia In: Handbook of Clinical Neurology, Volume 45. Edited by P.J Vinken, G.W. Bruyn and H.L. Klawans, Amsterdam: Elsevier, pp.433-455
- 岩田誠(1992): 神経文字学, 脳における日本語の読み書き 科学, G2, 498-509.
- Sugishita, M., Otomo, K., Kabe, S. et al. (1992): A critical appraisal of neuropsychological correlates of Japanese Ideogram (Kanji) and phonogram(Kana) reading. *Brain*, 115, 1563-1585.
- 杉下守弘, 他訳 1995, 漢字と仮名の読みと脳との関係に関する批判的総説 *imago* 6, 260-283.
- SUGISHITA M. IWATA M. TOYOKURA Y, YAMADA R (1978) Reading of ideograms and phonograms in Japanese pwlients after parcial conamisrotomy. *Neuropsychologia*, 16, 417-426.
- Sugishita M. YOSHIOKA M, KAWAMURA M(1986) Recovery from hemialexin, *Brain and Language*, 29, 106-118.
- 杉下守弘, 幕内 充, 神長達郎: 機能的磁気共鳴画像法による漢字と仮名の音読の比較, 第21回日本失語症学会総会, 1997. 12, 11-12, 東京

～ 文字シンポジウムから ～

漢字コードについて

On the “Kanji” codes

宮澤 彰・ MIYAZAWA Akira

抄録：日本語，中国語等のコンピュータ処理で使用される漢字を含む文字コードについて概説する。はじめに文字コードに関する枠組みを示し，次いで交換用文字コードの国際的な枠組みになっている ISO 2022 を紹介する。日本のほとんどのコンピュータが使用している交換コードである JIS X 0208，及び，その他の JIS コードについてもふれる。さらに，中国，韓国等の「漢字」コード，及び新しく制定されて議論を呼んでいる ISO 10646 UCS (Unicode) についても紹介する。最後に，漢字コードのかかえる問題点を整理し，また，現在進んでいる動きについて概観する。

Abstract : This paper is a survey of “kanji” codes, or character codes, including Chinese ideographs used in Japanese, Chinese and other languages. First, a framework for character set discussions is presented, then ISO 2022, which is commonly accepted frame for international character codes, is introduced. Then, JIS X 0208, widely used Japanese character code, and other JIS standards are described. Descriptions of Chinese and Korean and other character codes with ideograph follows. Then, newly established and controversial character code ISO 10646 UCS (Unicode) is reported. Finally, issues on “kanji” codes are sorted, with current various movements of “kanji” code.

和文キーワード：漢字コード，漢字，文字コード，外字，字体，文字化け，EUC，シフト JIS，UCS コード，ユニコード

English keywords: character codes, ideograph, JIS-code, EUC, Shift-JIS, UCS, Unicode

1. はじめに：文字コードをめぐる4つの枠組み

コンピュータで書き言葉を表現するには，文字コードという方式でコード化して扱う。漢字に限らず，どんなスクリプトでも，このコード化の方式や使用方法には，いろいろな問題を抱えている。これらの問題は，言語の問題であったり，コード化の約束上の問題であったりするわけであるが，しばしば錯綜した議論となって見通しが悪くなる。最初に，問題点の整理のために4つの枠組みを示す。

書き言葉の世界は，概念的に，文字として筆記

印刷されている図形としての文字の集合を考える。

スクリプトの世界は，書かれた言語の世界の文字を，ラテンアルファベット，漢字，かな，などの集合(スクリプト)にまとめ，スクリプトの中の個々の文字を集めた集合である。書かれた言語の世界の1文字は，スクリプトの世界のどれかの文字と対応させることによって，文字として認識される。

文字コードの世界は，交換コードの世界と内部コードの世界に分けて考えた方がわかりやすい。交換コードの世界は，JIS (日本工業規格) や，ISO (国際標準機構) で定められたある枠組みとスクリプトの世界の文字とを対応させた表の集合である。枠組みとして代表的なものは，ISO 2022

* 学術情報センター研究開発部教授

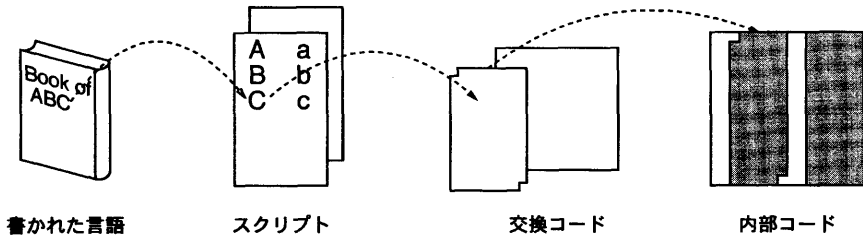


図-1 文字コードをめぐる4つの枠組み

(JIS X0202)に規定されている94文字、または94×94文字の枠組みで、7bitパターン16進21~7eの枠となっている。ISO 646IRV(ASCII)はその代表的なものである。また、JIS X0208はこの枠2つを重ねた8836文字分の枠をつかったマルチバイト交換コードの代表例である。2022と異なる枠組みとしては、IBMメインフレーム系のOSで使われているEBCDICの8bitパターン16進40~feの枠組み、後述するUCSの16bitや32ビットの枠組みなどがある。

交換コードの枠組みは一般に流通しているオペレーティングシステム(OS)で、そのまま使われているわけではない。このため、交換コードと内部コードとを分けて考える必要がある。シフトJISとか、EUCとか呼ばれるのは、内部コードの例で、同じ交換コードJIS X0208を異なる内部コードで使っている。内部コードはオペレーティングシステム(またはあるソフトウェア体系)で図形文字として扱われるビットパターン(コードポイント)の集合と、それに対する交換コードとの対応付けである。

文字コードに関する問題点のほとんどは、上記4つの世界の中が十分整理されていないことや、それらの間の関係付けが明確でないという点に起因する

2. ISO 2022

ISO 2022(JIS X0202 コードの拡張法情報交換用符号の拡張法)は交換コードの枠組みとして広く使われている。2022に従う交換コードは基本的には94文字、96文字、又は、これらのべき乗(94×94=8836文字など)で、ひとまとまりとして作られる。このまとまりとして、例えば、

ASCIIとよばれるANSI X3.4とか、JIS X 0208「7ビット及び8ビットの2バイト情報交換用符号か漢字集合」といった名前、番号が付けられており、これを国際的に登録する制度がある。(現在登録機関は日本の情報処理学会情報規格調査会にある)。登録されるとこのコードを識別するための終端文字が与えられる。JIS X0208-1997の終端文字は“B”である。

2022は、7ビット又は8ビットの中で、これらの複数の文字コードを切り替えながら使用する方式も定めている。この方式は指示(designation)と呼出し(invocation)という2段階を持つ複雑なものであるが、簡単に言えば、ある特殊な制御文字列が来るとそこから先がJIS X0208になったり、また別のコードになったりする、というものである。しかし、通常のオペレーティングシステムの中では、この制御文字列をテキストファイルの中に書いておいてもコードの切り替えと解釈してもらえない。解釈してもらえない場合でも、ある一部に限られていたり、使用できる交換コードの制限があったりするのが普通である。

94文字や8836文字といった枠は、文字コードの作り方に影響してくる。一部の記号だけが異なる

多くの文字コード(例えば、ASCIIとJIS X0201の逆斜線と円記号)は、94文字という制限の中で多くの要求を満たすために生まれたと思える。また、JIS X0208の6千数百字という漢字数も8836文字の枠が大きく影響しているように思われる。

3. JIS 漢字コード

JIS 漢字コードといわれるのは通常JIS X0208のことである。1978年の制定以来、83年、90年、

97年と改訂されてきた。「符号化漢字集合」という名称にもかかわらず、漢字以外に仮名、数字、ラテンアルファベット、ギリシャ、キリルアルファベット、その他の記号を含んだものとなっている。83年版は、常用漢字表の制定を考慮し、78年版に対し非漢字の追加、及び漢字の「字体」の入れ替えを行った。常用漢字表の制定を考慮に入れたものである。(この結果として漢字に4文字が追加となった)。90年版は83年版と漢字2文字の追加を除いて等しい。この2文字は法務省での人名用漢字表の追加によるものである。97年版は文字セットとしては90年版と変わっていない。漢字の典拠や、字体の「包摂」概念の導入など、言語学的基礎を整備したものとなっている。

これらの改訂に対してはいろいろ議論があった。少なくとも、現在では一般的に83年の改訂のような文字の入れ替えをすることは、混乱を引き起こすという認識が一般的になっていると思われる。97年改訂では文字コード表は変更されていないし、90年の改訂でも、2文字の追加は技術的な理由によるものではなかった。

JIS X0208について、いろいろ議論はあるものの、現在日本のほとんどすべてのコンピュータが使っているものであり、実用的なものであることはこの使用実績からも十分証明されていると考えられる。

JIS X0208以外に、JISの漢字コードとしてはJIS X0212「情報交換用漢字符号一補助漢字」がある。これには漢字5801字の他、ダイアクリティカルマーク付きのアルファベットなどが含まれている。漢字は印刷産業連合会、コンピュータメーカ、いくつかの団体などのいわゆる外字表を中心に選定されたものであるが、JIS X0208の83年改訂の結果なくなった字体も、文字として選定されている。この補助漢字符号は、若干の実装例はあるものの、現実としてはほとんど使用されていないといえる。

もう一つの漢字を含むJISコードであるJIS X0221「国際符号化文字集合(UCS)」は後述するISO 10646 UCSのJIS化である。詳細はUCSの節に述べる。

4. 日本以外の漢字コード

中国ではGB2312が主として使用されている。JIS X0208と似た構成を持つ94×94文字コードで、6千数百の漢字を含む。これは簡化字体であり、いわゆる繁体字は含まない。GB 2312に対応してGB 12345という文字コードがあり、こちらはGB 2312の簡体字部分を、対応する繁体字に置き換えた文字コードとなっている。これらの他にGB 7589、GB 7590などの拡張文字コードがあるが、現実的にはほとんど使用されていない。

台湾ではCNS 11643という国家標準がある。94×94文字を7面もつ大きなものである。最初は2面約1万3千漢字であったが、最近大幅に拡張された。ただ、現実的にはBig5と呼ばれる文字コードが、一般的に使用されている。この他に、CCCIIという94×94×94文字の枠組みを使用するコードがあるが、部分的にしか使用されていない。

韓国ではKS C5601という標準コードがある。これもJIS X0208とよく似た構成を持ち、2千数百のハングル(完成型)、と漢字約5千を含む。なお、KS C5601には組合型のハングルコードも規定されており、こちらを使用すると漢字は使用できない。さらに、韓国にも拡張文字コードKS C5657があるが、現実的には使用されていない。

米国にも、漢字コードが標準として存在する。ANSI Z39.64 EACC(East Asian Character Code Set for Bibliographic Use)で、台湾のCCCIIをもとにした94×94×94文字セットで、図書館目録データベースで使用されている。

ベトナムでも、字喃といわれる漢字?の文字コードが標準化されている。

5. 内部コード

日本では内部コードとして、PCの世界で主として使用されているシフトJISと、UNIXの世界で主として使用されているEUC-JPとが主流である。さらに、ISO2022jp(RFC1468)とか、JIS X0202の指示シーケンスを使ったJISコードと呼ばれるものとかが、メールやWorld Wide Web(WWW)、あるいはmuleと呼ばれるエディタなどのいくつかのアプリケーションで共通に使用さ

れている。また、メインフレームの OS では IBM 漢字コード、富士通の JEF などそれぞれ独自の内部コードが使われてきた。

これらの内部コードは、いずれも JIS X0208 及び JIS X0201 (いわゆる半角の英数字、記号及びカナ) をマッピングして使用している。(これらの他に独自の部分を持つ内部コードもあるが)。文字コードとして使用するコードポイントの集合が異なることと、それに伴って交換コードからのマッピングの方法が異なるだけである。従って、上記の交換コードに対応している限りでは、内部コード相互の変換は容易である。

現在、UNIX や Windows などの OS はロケール(locale)という概念で国際化を行なっている。この方式では例えば、ja というロケールを選ぶと、内部コード、メッセージや日付の表示方法などが、日本語用の文字コード日本語の表示に切り替わる。(もちろん、プログラムがそのように作ってあれば、また表示用のフォントが適切に切り換えられればなどの条件下であるが)。

この意味で、これら国際化された OS は複数の内部コードをサポートしている。(場合によっては日本語用でシフト JIS と EUC の両方の内部コードをサポートしているものもある)。

6. UCS

ISO 10646 Universal Multiple-Octet Coded Character Set (UCS) は、93 年に成立した新しい文字コードで、ISO 2022 の 94 文字の枠と異なり、16 ビット約 6 万 5 千字、又は 32 ビット約 40 億字の枠をもつ。現在の 16 ビット枠の基本多言語面中では、ラテンアルファベット、漢字をはじめ、ギリシャ、キリル、アラブ、ヘブライ等のアルファベット、インドや東南アジアの文字など世界中のかなりのスクリプトが含まれている。日本では JIS X0221 として、JIS 化されている。もともと、言語によって文字コードを切り替えることなく一つの文字コードで処理できることを目指したものである。

この国際標準の成立をめぐるには、ISO/IEC の委員会案と、米国の Unicode コンソーシアムの案とで激しい争いがあり、結局は Unicode コ

ンソーシアムの案が国際標準となったといういきさつがある。(これも、もともとは米国内での調整不足、あるいは ISO 活動への無理解が原因であるが)。日本は委員会案を支持し、最終的には負けることになった。このことが、いまだに Unicode に対する感情的なしこりとなる傾向も見うけられる。

技術的な面で、委員会案と Unicode 案との大きな差異は 3 点あった。一つは、コードの構成で、2022 の 94 文字枠との互換性を重視した委員会案と、16 ビットの空間をすべて使う Unicode 案、2 番目は、ダイアクリティカルマーク類について、合成済みの文字のみを認める委員会案と、要素を文字として合成を行なう Unicode 案、3 番目が、漢字について日中韓の基本文字コードのみを並列に埋め込む委員会案と、これらを統合して漢字全体で一つのまとまりとする Unicode 案、という 3 点である。

ISO での情勢がほぼ Unicode 側に傾きかけた時点で、日本は、漢字部分の統合という案を受け容れ、CJK-JRG (China, Japan, Korea - Joint Research Group) というグループの立ち上げを進め、このグループで統合作業をおこなった。ソースは、それまでに発表されていた日中韓台の標準コード(補助漢字符号等の拡張コードを含む)及び若干で、統合のための規則を定め、これに従った。統合規則は JIS X0221 では解説に示してある。結果は図-2 のように、4 カラムに中台日韓の字形とコードをならべた 20902 字となった。

この 4 カラム表示を、各国での字体の差を示すものとして捉える考え方があるが、CJK-JRG での編集時の意図としては、統合規則とともに字体の包摂を例示するものとして考えたものである。各カラムは、ソースでのコード及び表示字形を示すもので、決して各言語を示すものではない。

統合規則には原規格分離という第一規則がある。これは、ソースとなった標準の中で別のコードポイントとなっているものは、統合レポートリ上でも別のコードポイントを与えるという規則である。このため、他の統合規則と矛盾することがある。この規則を第一に置いたのは、UCS と各国の既存標準との相互変換を可能にしたいという、各国

Row/Cell	C	J	K
Hex code	G - Hanzi - T	Kanji	Hanja
078/013 4E0D	不 不	不 不	不 不
	0 - 323B 1 - 4462 0 1827 1 3666	0 - 4954 0 - 4152	0 5C74 0 6084
078/014 4E0E	与 与	与 与	与 与
	0 - 536B 2 - 212F 0 - 5175 2 - 0115	0 - 4D3F 0 - 4531	
078/015 4E0F	丐 丐	丐 丐	
	3 - 3021 2 - 212D 3 - 1601 2 - 0113		

図-2 ISO 10646 UCS CJK 統合漢字部分

からの要求によるものであった。

UCSの実装は、Windows NTなどで進んでいる。またUnixでもUCSをロケールの一つとしてサポートするという計画があると聞く。とはいえ、成立後5年にもなるのに、広く普及しているとは言い難い。最大の原因は流通しているOSでのサポートの問題と思われるが、UCSを本当に必要とするような多言語の本格的な応用というのが、あまり世の中にないせいともいえよう。

7. 漢字コードをめぐる問題点

7.1 外字

ここで外字と呼ぶのは使用している交換コードにマッピングされない文字のことである。何が問題かという、その文字を含むテキストをコード化できないわけで、ユーザ側からいえばコンピュータでその字が使えない、ということになる。これは、文字コードに対する批判としてもっとも単純明快なものである。また、漢字の場合とくに、漢字制限論批判とも結びつきやすい。

外字の問題のもうひとつとして、利用者定義文

字の交換問題がある。多くのシステムは内部コードのうち交換コードに対応する枠の外に（ある場合には交換コードの未定義文字のコードポイントに）利用者定義文字を割り当て、自分で文字パターンを作成すれば表示、印刷できるようになっている。あるシステムで使っているこの利用者定義文字をはかのシステムで使いたいと思っても、使えない。これはよく考えれば当然のことであるが、どれが交換用コードにある文字で、どれが利用者定義文字かということは一般利用者が知らないことも多いため、この問題が生じる。

これらの外字問題は94×94文字コードにおさまる範囲でしか扱えないというコンピュータシステム側の責任もあるが、もう一方で漢字というスクリプトに、文字がどれだけあるかということを示せない言語学側の責任でもある。

文字コードはある時点では有限個の文字しか持てない。一方、漢字というスクリプトでは、文字はいつまでも増え続けるかもしれない。この事を考えると、外字問題の本質的な解決はありえない。結局、どの程度で実用的に問題がないか、という

点に絞られるが、これも応用分野に依存することからであり、結論を出すのは難しいだろう。康熙字典も大漢和辞典も全ての漢字を尽くしているわけではない。これだけあれば十分と認められるような漢字スクリプトの表、あるいは字典、ができない限りは、継続的な拡張、あるいは登録制を続ける以外に方策は考えられない。

7.2 字体

字体の問題は、具体的な場面では、二つの点に問題として現れる。一つは、自分のコンピュータで作った文書を、他のコンピュータに送ってプリントしたら、別の字体になってしまった、というものである。もう一つは、字体を使い分けたい、例えば一つのコードポイントに包摂されている一点しんにょうと、二点しんにょうとを書き分けたい、というような場合に、一般的に通用する方法が無いことである。どちらも結局、字体を表現するための標準的な方法が無いことが原因である。

字体に関して、漢字の文字コードがいかにあるべきかという点で、すっかりした結論は出ていない。JIS X0208は97年版から、交換コードの一つのコードポイントは複数の字体に対応するという立場をとった。これに対し、交換コードの一つのコードポイントは一つの字体に対応すべきであるという極端な意見もある。また、コードポイントは一つの文字に対応すべきであるという立場もある。この場合は、漢字における文字とは何で、字体との関係はどうなっているかを、何らかの方法ではっきりさせる必要がある。

7.3 文字化け

文字化け問題は、例えばe-mailやWWWのブラウザで経験した人は多いと思う。一般には「\$BD9\$ \$%a!<%k\$G」といったような意味不明の文字列になる。これは、交換コードと内部コードのマッピングの方法が複数あり、それらの間の混乱が原因となっている。

もちろん、このようなことは起こらないことが望ましい。そのための標準、規則類もいろいろある。また、これらの応用プログラムは複数の内部コードを自動判別するなどの方法で対応してもい

るのであるが、実際上の文字化けはなくなる。文字化けの理由の技術的細部はさまざまで、いちいち詳述してもしかたないが、要は内部コードが複数あり、さらに、送信の際に使われるエンコーディング方法も複数あるという複雑さが、原因となっている。

8. 漢字コードをめぐる動向

漢字コードの拡張は、現在も標準化の世界で続いている。その一つはJISのいわゆる第3、第4水準の計画である。これは現代日本語の表記に使用される漢字、その他の記号類で、現在のJIS X208にないものを約5千字収録する予定で進んでいる^{1),2)}。

一方、これとは別にISO 10646 UCSの拡張として、現在の約2万漢字から、6千字ほどの拡張が提出されており、さらに、superCJKという呼び名で、康熙字典等を始め、現在集められる漢字をすべて集めた、拡張を行なう計画も進んでいる。

これら、最近の標準化の動き以外に、コンピュータでより多くの漢字を処理したいという活動として、京大人文研勝村氏らのe-漢字³⁾、東大グループの6万4千漢字などがある。

引用文献

- 1) 芝野耕司. 7ビット及び8ビットの2バイト情報交換用符号化漢字集合
——第3水準及び第4水準. s.l: s.n, [1998年12月4日参照]. ウェブページ: <<http://www.tiu.ac.jp/JCS/>>
- 2) 芝野耕司. 「7ビット及び8ビットの2バイト情報交換用符号化拡張漢字集合(案)」の公開レビューのお知らせ. [東京]: 日本規格協会, 1998-11-07, [1998年12月4日参照]. ウェブページ: <<http://jcs.aa.tufs.ac.jp/pubrev/index.htm>>
- 3) 勝村哲也, 丹羽正之. eKanji. s.l: s.n, [1998年12月4日参照]. ウェブページ: <<http://www.zinbun.kyoto-u.ac.jp/~ekanji/>>

専門用語集の編集と活用

一用語集超活用ソフト HT³(エイチティースリー)の利用効果一

Compiling the terminology list and practical application
Effect of HT³, the ultra-useful software for terminology list utilization

貝島 良太* KAIJIMA Ryota

抄録：いかに労作の用語集でも、書籍、電子版を問わず実際はさほど利用されていない。これは作り手にとっても、利用者にとっても残念なことである。書籍も電子版も共に「引かなくてはならない」というごく当然のことが災いしているからである。(株)日立国際ビジネスが開発した用語集超活用ソフト HT³を利用すれば、日/日英/英日/英英に、完全に用語集を活用できるので、作り手の苦勞が報われるようになる。HT³には用語集を制作するときに便利な機能もある。

キーワード：用語集 標準表記/同義語/異表記 訳語付与 用語統一 辞書引き

ABSTRACT: Terminology lists are rarely used sufficiently, whether they are in written or electronic form. This is because users have to consult the lists every time they encounter a term they need to know. The Hitachi Terminology Transformation Tool (HT³) eliminates the need to consult electronic terminology lists, automating the process of verifying and standardizing terms in translations from Japanese to English and vice versa, as well as writing original documents in either Japanese or English. By using HT³ the hard work required to compile terminology lists is substantially rewarded. In addition, HT³ provides a unique function to elegantly streamline the creation of electronic terminology lists.

1. はじめに

1.1 用語集の目的

用語集の作り手と利用者の双方にとり、用語集の目的は「用語/訳語の標準化(表記の統一化も含む)」を図ることだ。説明書、報告書、規則書などで、相手に誤解や迷いを与えないよう、ひとつのものごとにひとつの用語を対応させるためでもある。このため、訳語が用語集どおりに使用されたか、翻訳の検収にも用語集は利用される。

1.2 用語集使用上の矛盾

用語集では、良いものを編纂することが最重要だ。そのため多くの専門家が携わっている。しか

し、1.3に述べるように用語集は、実際は残念ながら活用されていない。用語集や辞書/辞典の根本を、いちいち引くことに置いているからである。用語集が書籍であれ電子版(EPWING形式)であれ、通常利用者は既知の用語はまず引かない。結果的に用語集のものとは違う用語/訳語を自信を持って使い、その違いに気が付かないこともある。

一般的に、用語集/辞書を引き、目的の語が見つからないときは手間が無駄になる。その確率が高いと、イライラするし用語集自体を頼りなく感じ、引かなくなる。いわゆる、用語集を飾っておく状態だ。優れた用語集は網羅用語が多い(即ち用語集が厚い)ので余計検索に手間がかかる。用語集活用で矛盾を感じるころである。

* 株式会社 日立国際ビジネス 開発推進センタ(DE)
Development Promotion Center, Hitachi Business
International, Ltd.

1.3 用語集不活用の原因

検索文字列に出会えない主な原因は、①用語集の語彙不足、②検索すること自体が面倒、③同義語や異表記の存在、④読みから引く場合の異読/誤読、⑤文字列の切り出し違い、である。紙(書籍)の用語集は物理的に引く必要があるので完璧な活用と言うのは所詮不可能だから、ここではもっぱら電子版の用語集/辞書について論じる。

なお、「異表記」とは、「コンピューターとコンピュータ」「βカロチンとベータカロチン」「一次コイルと1次コイル」「かぜと風邪」のように、人や場合により表記が異なるが意味は同じものである。英語でも英/米つづり違い、複合語の語間のハイフンの有無、スペルアウト/省略形などたくさんある。筆者は異表記のことを「言葉あそび」とも呼んでいる。同義語同様にこの「言葉あそび」による異表記の存在にも対応すると、用語集は利用者にとり格段に便利になる。

1.4 用語集活用の対策

前述①の「語彙の収集不足」は用語集の本質として当然なので、ここでは論じない。

②の「引くのが面倒」が用語集活用の最大の敵だ。これの解決が用語集活用の鍵となる。目指す用語がその用語集や辞書/辞典に出ているかどうか不明の状態で「引く(検索する)」のは、無駄を覚悟で行うものである。特に大量文書の業務翻訳の場合などは、できれば代替りの人に、用語集にある訳語を文書中の当該語の横に全部書き出してもらいたいと誰しも思うだろう。筆者はこの点の解決を主目的としてHT3を開発した。詳細は5.に後述する。

次の問題は、「③同義語や異表記」での検索で、目指す用語に出会えない場合だ。ほとんどの用語集は標準表記と若干の(有名な)同義語を見出し語として収録している。現在(平成11年1月)のところ市販の電子版用語集/辞書で、異表記からも検索できるものは、筆者の知る限りHT3対応のもの[[CD-ROM版 JIS 工業用語大辞典(財団法人日本規格協会より平成10年1月発売)]と「医薬T辞書 [HT3版] (日本データベース開発株式会社より平成10年9月発売)]」だけである。例えば、

CD-ROM版 JIS 工業用語大辞典では、JISにおける標準表記の「フレキシブルディスク」という語は、当該「フレキシブルディスク」と同義語の「フロッピーディスク」以外に、「フロッピー」「フロッピ」「フロッピ」「アーカイブディスク」「アーカイヴディスク」「アーカイブディスク」「FD」「F/D」などからも検索できる。また、機械用語の「しんだし」(センタリングの意)という語は、「心出し」が正しい表記だが、90%以上の日本人が書く「芯出し」という文字では、通常は用語検索に失敗する。しかし、同辞典では「芯出し」からでも引け、さらに「心出し」への置換もできる。この同義語や異表記から検索ができるのも用語集活用の決め手になる。

自由入力文/文字列に対するコンピュータ文字列検索では、(a)出現可能と予測される同義語や異表記を用語集に登録しておくか、(b)標準表記から異表記をプログラムで自動生成する方法を取らない限り、検索中のキーワードと同一表記の文字列にしか検索できない。HT³では(a)を採用した。(b)は用語集への異表記の登録は不要/最少になろうが、誤検索のない常識的な異表記のみを自動生成するソフトが簡単にはできそうになかったからだ。HT³用語集の編集は、用語集への異表記登録には手間はかかるが、努力は必ず報われるのである。このことについて澤田(1997)は、『...通常FDと略記される電子媒体は、JISではフレキシブルディスクとされているが、多くの人々はこれをフロッピーディスクとして理解し、それで検索している。そうした検索にもヒットさせるため、同義語や異表記を数多く付加することにしたのである。この部分は、(株)日立国際ビジネスの貴重なノウハウを盛り込むことにより実現したものである。いわゆる曖昧検索を実現するために、ソフト側をいじるのではなくファクト側に工夫を加えるという方法を我々はとった。...』(傍点筆者)と述べている。

④は用語集を「読みで引く」か「文字列で引く」かのことだ。業務で用語集を利用する場合、目の前にある説明書や論文などの文字列の用語を扱うわけで、直接文字列を用語集と照合するほうが、読みを介するよりも早い。検索文字列を文書中から

コピーし、辞書ビューワの検索文字列欄にペーストすることは普通に行われている。その方が読み違いによる誤検索が防止できる。HT³は、「文字列で引く」方式を採用している。

⑤「文字列の切り出し違い」により用語集の用語(文字列)が検索できないこともある。その点HT³が採用した「前方最長一致」方式の検索は最適であろう。

以上により、電子版用語集活用の条件は次の4点と言えよう。

- 自動検索 (用語集が文書に飛び込んでくる)
- 異表記/同義語からも検索可能
- (読みでなく)文字列での検索
- 用語集の多目的利用

2. HT³の特長

HT³の最大の特長は用語集を活用する点だ。因みにHT³は、このソフトが翻訳用に訳語付与したり、異表記の用語を標準表記に置換したりすることから、「Hitachi Terminology Transformation Tool」とし、その頭文字からつけたものである。「用語と訳語の統一用ソフト」「用語集超活用ソフト」「飛び込んでくる辞書引きソフト」などツールとしての多面性も特長だ。(機械翻訳ではないので誤解を防ぐため Translation とせず Transformation とした)

2.1 基本機能

以下にHT³の基本機能を紹介する。

- 用語集を Excel で編集し、Word 上の文書で検索

一般の機械翻訳や文章チェックソフトは辞書の一覧性がないものが多く、用語集作りには不便だ。その点HT³はExcelで用語集を作成するので、勿論一覧性に優れているし、ソートやフィルター機能を利用し、用語集を推敲編集したり、情報を共有/分割することも簡単である。(3.3 参照)
- ひとつの用語集を多目的に活用

通常の記事チェックや日/英機械翻訳用の用語集は単一目的なので、他に流用できない。HT³は用語集を双方向の翻訳支援で利用できるのに加え、翻訳結果の訳語使用のチェック(日/英)にも

使用できる。また、「飛び込んでくる辞書」としての辞書引きツールにもなる。なお、HT³の「日/英」は日/英文中にある標準表記と異表記(同義語を含む)の文字列を色分け表示し、異表記を標準表記に置換する機能である。「日英/英日」は原稿中の標準/異表記を問わず、相手言語の標準表記の訳語を付与し、翻訳を支援する機能である。

● 日/英で異表記を標準表記に置換

日/英の色分け結果で、用語集の標準表記に一致した文字列は水色/緑になる。異表記の文字列は明るい紫/赤になり、これらは標準表記の文字列に置換することができる。HT³で用語集の標準表記用語に統一したり、統一されていることを確認したりできる。

● 日英/英日で用語集の訳語を使用した翻訳支援

翻訳原稿中の用語が用語集の標準/異表記の如何を問わず、用語集にある相手言語の標準表記が訳語として付与される。付与された訳語を訳文中に取り込み翻訳する。もはや用語集を参照することなしに用語集の訳語を完全に翻訳に活かすことができる。付与された訳語を使わなくても良いが、故意に使わないのと知らずに使わないのでは雲泥の差だ。

● 複数の用語集を優先順位をつけて統合使用

HT³は複数の用語集を統合して使用できる。複数の用語集に用語が重複しているとき、優先順位の高い用語集の方を活かし、低いものを無視する。誤解や曖昧さの原因である重複を許さないソフトだ。

● 何回でも楽に検収が可能

一般の文章チェックや翻訳ソフトでは、仮にそのソフトと用語集を使用して作業をしたとしても、後の検収は手作業である。検収ができないソフトは中途半端だ。HT³なら同一の用語集を利用して日/英の色分けで標準表記の使用箇所を視覚的に確認可能だ。検収も、検収の検収も簡単にできる。また、統計表示機能により文中の標準/異表記の個数を把握できる。

● 良いところを養める

通常の記事チェックツールは、スペルミスなど問題点を減点主義で指摘する。どれだけ正しいかを積極的に確認できない。一方、HT³は日/

英英の色分けで、標準表記の用語は水色になる。つまり、用語集の用語が正しく使われているので、「良い」と誉めるのである。それ以外のところ、特に黒のまま残ったところに重要単語があれば要注意だ。一般的に、チェックの対象が限定されると、チェッカーには心の余裕ができる。

3. HT³対応用語集の作成

HT³用の用語集作成について、「フロッピーディスク」という用語を例に以下に述べよう。

3.1 かにの原理

筆者の思い付いた「フロッピーディスク」の日本語/英語の語彙を[図1]に示す。同義語、省略形、異表記(言葉遊び)を含めてかなりの数だ。

日本語は「フロッピーディスク」、英語は「floppy disk」を標準表記とすると、それ以外のものは

すべて両者の異表記(HT³では標準表記以外は同義語も含めて異表記として扱う)ということになる。この概念をかにの形に例え「かにの原理」[図2]と呼ぶ。かにの甲羅は日/英の標準表記、左右の手足はそれぞれの同義語や異表記である。標準/異表記の入れ替えは自由だ。異表記に「芯出し」のような間違い表記を入れておくのも親切であろう。

3.2 HT³対応用語集の概念

[図2]の「かにの原理」を表にしたものが[表1]で、HT³の標準/異表記の概念である。

1対1の関係にある日/英標準表記をA/C列に入れる。n個ある日本語異表記はB列に入れる。同様にn個の英語異表記をD列に入れる。A/B列の日本語標準/異表記を日本語文書と前方最長一致方式で検索し、一致(ヒット)したものがあれ

日本語
フロッピーディスク、フレキシブルディスク、フロッピディスク、フロッピイディスク、フロッピー、フロッピイ、フロッピ、アーカイブディスク、アーカイヴディスク、アーカイブディスクット、アーカイヴディスクット、フロッピーディスクット、フロッピディスクット、フロッピイディスクット、FD、F/D
英語
floppy disk、floppy、FD、F/D、flexible-disk、floppy-disk、flexible disc、floppy disc、flexible-disc、floppy-disc、archive diskette、archive-diskette、archive disk、archive-disk、archive disc、archive-disc

図1 フロッピーディスクのさまざまな表現

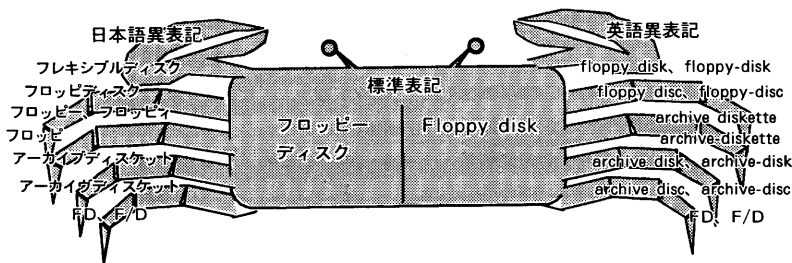


図2 標準表記と異表記に関するHT³の基本概念「かにの原理」

表1 標準/異表記に関するHT³の基本概念(表形式)

A	B	C	D
日本語標準表記	日本語異表記	英語標準表記	英語異表記
フロッピーディスク	フレキシブルディスク、フロッピディスク、フロッピイディスク、フロッピー、フロッピイ、フロッピ、アーカイブディスク、アーカイヴディスク、アーカイブディスクット、アーカイヴディスクット、フロッピーディスクット、フロッピディスクット、フロッピイディスクット、FD、F/D	floppy disk	floppy、FD、F/D、flexible-disk、floppy-disk、flexible disc、floppy disc、flexible-disc、floppy-disc、archive diskette、archive-diskette、archive disk、archive-disk、archive disc、archive-disc

なる。これは辞書の先引きを利用したものだ。色分け後に標準書体のまま残った部分に引きたい用語がある場合は、すぐに他の用語集を参照すべきである。

5. 辞書引きの手間を排除（引かなくてよい辞書）

1.4②の詳細を述べる。一般の国語辞典や英/和辞書などは、勉強もかねて引くことが多い。特に多義性のある語については、説明文からヒントを得たり、適切な語を選択したりするので、無駄足もやむを得ないという気持ちがある。しかし、専門用語集を業務で使用すると、速度と正確さの両立というジレンマに悩まされる。機械的な正確さで用語集の用語/訳語を執筆/翻訳に使用しなくてはならない。検収もすべきである。しかし現実には、時間的制約、忍耐力の限界、自信などで結果は妥協の産物となり不統一なものになる。

これらの根元には「辞書や用語集は引くもの」という諦めにも似た感情がある。良い辞書は語彙が多く必然的に引くのも大変になる。そこで登場したのが「用語集は引かなくて良い。用語集のほうから文書中を勝手に走り回りマーキング/訳語付与すべきだ」というHT³の考え方である。

HT³は、用語集の全用語と、文書中の文字列を「前方最長一致」方式で照合し、用語集を日日/日英/英日/英英/辞書引きに総合利用する。機械翻訳や文書チェック、辞書引きソフトなど文書系のソフトは数多くあるが、HT³のようなものは例を見ない。この点について、前出の澤田は次のように触れている。『...CD-ROM化にあたりまず問題となったのは、いかなる検索エンジンを搭載するか、という点である。我々は、当初辞書検索を念頭に置きJISにもなってるEPWING規約を検討した。しかし、我々は、単なる辞書引き機能のみでの開発に漠然と不満足なものを覚えていた。いわゆるテクニカルライターと呼ばれる人たちが対象の多くを占めているならば、この方々は本大辞典にどのような機能を望まれるのであろうか。辞書引きに加え何か便利な機能を付加できないか。概して、情報技術の進展動向は、音や画像といった先端部分を向いてきた。そうした中で発見でき

たのが、(株)日立国際ビジネスの開発による翻訳支援並びに用語統一ソフトであるHT³である。(中略)用語統一のしやすさということでありユーザフレンドリーな開発思想である。このHT³にJISの用語データベースを具備しさえすれば、テクニカルライターや管理者が欲するJISに基づく用語統一と言う機能が実現できる。もちろん、自らの用語も自由に登録しJIS用語を核としたデータベースを構築することもできる。こうして本CD-ROM版の開発は、HT³とのドッキングを前提として進めることとなり、(株)日立国際ビジネスとの共同開発として1997年2月より着手した。...』

6. まとめ

用語集を作成される方々に是非提案したいのが、「用語集を作るときは利用者が活用できる形で提供する」ということである。書籍やEPWING形式の用語集にするのではなく、真の意味での活用される用語集を目指し、HT³形式にしていだきたい。HT³形式にすることで用語集作りの手間が特に増えることはない。かえってHT³の辞書作成機能を利用することにより、①ソーティングが楽、②不本意な重複の発見、③追加修正の日付の記録、④解説文に見出し語が正しく使用されているかの検収、等々のメリットをご享受いただける。用語集の目的達成のために、HT³のご利用を関係諸兄にご案内する次第である。

商標表示

「Word」と「Excel」は米国Microsoft Corporationの、「HT³」と「エイチティースリー」は(株)日立国際ビジネスの登録商標です。

参考文献

- 1) 貝島良太 (株)日立国際ビジネス 1997 「用語と訳語の統一用ソフト HT³の開発」テクニカルコミュニケーションシンポジウム '97 論文集 153-157
- 2) 澤田 位 (財)日本規格協会 1997 「JIS工業用語大辞典CD-ROM版の開発について」テクニカルコミュニケーションシンポジウム '97 論文集 146-149

- 3) 富井玲子 (株)日立国際ビジネス 1998 「引く辞書から飛び込んでくる辞書へ 一用語集超活用ソフト HT³対応の辞書づくり」 テクニカルコミュニケーションシンポジウム '98 論文集 83-87

■国際会議ご案内■

European Association for Terminology
Conference on Co-operation in the field of Terminology in Europe

期 日 1999年 5月17日(月)-19日(水)
場 所 パリ
主 催 European Association for Terminology(EAFT)
テーマ Section A The Discipline of Terminology
1. Training
2. Information and Documentation
3. Resources(Terminological, Human, Financial)
4. Software
5. Harmonisation and Standardisation
" B The Status of Terminology
" C Co-operation in Terminology(National, Regional, European and International)
" D Special issues

5th International Congress on Terminology and Knowledge Engineering
(T K E '99)

期 日 8月23日(月)-27日(金)
場 所 オーストリア・インスブルック
主 催 Infotermほか

■詳細は事務局までお問い合わせください

外字の作成事例：『広辞苑』を中心に

上野 真志* UENO Masashi

本稿の主題となるのは、JIS第2水準までに含まれない文字を、『広辞苑』（の電子媒体版）ではどのように扱っているか、ということである。わざわざ、このように述べるのには、次のような理由がある。

コンピュータと漢字

コンピュータ上で漢字を使用可能にする試みは早くからあったが、実用化をめざしてあちこちで行われるようになったのは1970年代のことである。昨年文庫化された、杉山隆男『メディアの興亡』（文藝春秋）には、日経新聞がIBMと共同で、独自の漢字組版システムを開発する経緯が詳しく紹介されており、今でも興味深いルポルタージュである。この時期には、各企業が業務内容に応じた独自の漢字コードの制定に着手していた。その流れを受けて、JIS「情報交換用漢字符号系」として制定されたのが1978年のことであった。

このとき採用されたコード化の手法が7bitコードとの共存を考えた2バイトの組み合わせであり、（詳しいことは省くが）94区×94点＝8836文字が規定できるような体系になっている。実装にあたっては、数字やアルファベット、仮名、キリル文字などの非漢字を割り当てるエリアを15区×94点＝1410文字分作ったため、漢字が割り当てられるエリアはもう少し減ることになる。

このエリアは、常用漢字1945文字（制定当時は当用漢字1850文字）に比べると広大だが、約5万字（「諸橋大漢和」の収録数）ともいわれる漢字に比べるとあまりにも小さい。われわれが要求する文字数がどのくらいになるのかは、後に考察したいが、そもそも、第2水準という集合を決めることは、そこに含まれないものも生み出すこと

であり、外字の存在は当然なのである。

もう一つの「外字」

さて、現在の技術では、漢字の収録字数に関していえば、ほとんどあらゆる要求に応えることができる（もちろん、コストと時間を無視すれば、の話であるが）。しかし、その要求に対して、複数の解決方法があり、それが混乱をもたらすもう一つの要因になっているのではないか。根本解決に対する、対処療法の存在である。

JISのコード割り当て方式では、例えば02区や04区の後ろに文字の割り当てられていない領域が存在する。これを将来の拡張のための「保留領域」といい、改正ごとに増える文字は保留領域に割り当てられた。

一方、非漢字と漢字の境界にあたる09区から15区の空白域、そして漢字の終わる84区以降（85区から94区）は「自由領域」と呼ばれてきた。

JIS X 0208:1983の解説には「情報交換当事者の協定によって、一時的・局所的に文字を割り当てて利用しても構わない」とあり、これが対処療法の役割を果たしてきた。この自由領域を使って、メーカーは自分たちが必要だと思う文字を提供し（「システム外字」という）、またユーザーが必要とする文字（「ユーザー外字」という）を作成できる仕組みを提供してきた。

単に「外字」というときに、「ある集合に含まれない文字」の意味なのか、それとも「システム外字やユーザー外字」を指しているのかが明確でないために、混乱を来すことがあるので注意が必要である。

なお、1997年の改正で「自由領域」は「廃止」された。その経緯と影響についてもさまざまな議論があるが、その紹介はこの稿の任ではないので

* 岩波書店電子出版部

割愛する。

どれくらい不足しているのか

さて、それではJIS第2水準までに含まれない「外字」は、いったいどれくらいわれわれの目に触れるのだろうか。

一般にわれわれが使用することの多い小辞典タイプの漢和辞典には、だいたい8000~10000字が収録されている。それらの辞典の凡例によれば、文字の選定基準は、

- (1) 常用漢字
- (2) 人名用漢字
- (3) 上記(1)(2)の旧字体、異体字
- (4) その他、日常生活に必要と思われる文字。

また、日本や中国の主要な文献の読解に必要な漢字。

というところになる。少なくとも漢和辞典の世界では、この8000~10000字が日常生活に必要というのが、ある種の「定説」になっている感がある。

一方、最近は「JIS第2水準までを全収録」を謳っている漢和辞典も多い。第2水準までに含まれる漢字は6355字であるから、単純に計算すればこれらの辞典には、いわゆる「JIS外字」が1500~3500字程度収録されていることになる。先の「定説」と考えあわせると、現行の第2水準までではだいたい数千字が不足しているとも言える。この数千字が、「ある集合（この場合はJIS第2水準まで）に含まれない（けれども必要な）文字」という意味の「外字」である。

ただし、ここには注意しなければならない点がある。漢和辞典には、先の(3)のように旧字体・異体字が収録されるが、これにJISに含まれるものと含まれないものが存在するのである。例えば、先の「眞」や「學」は含まれるが、「黒」や「青」の旧字体は含まれない。収録字数にはこれらの旧字体・異体字を含めるので、JISに含まれない場合には、その分だけ単純に数が増えることになる。

『広辞苑』の場合

国語辞典の場合は、若干事情が異なる。漢和辞典が文字を解説するものであるのに対して、国語

辞典はことばを解説するものであるからである。

『広辞苑』を例にみてみよう。

『広辞苑』に限らず、一般的な中規模の国語辞典の凡例には、次のようなことが書いてある。

漢字の字体は、常用漢字ならびに人名用漢字はいわゆる新字体を、他は広く通用している字体を採用した。

(『広辞苑第五版』凡例)

原則として、現代仮名遣いにしたがい、常用漢字を用い、現在通用している文章体で記述しました。

(『カラー版日本語大辞典』凡例)

これでわかる通り、漢和辞典の場合と異なり、字体については一つの統制が働いている。

例えば、古典の用例中の文字は新字体に改めてある。項目「あだ - ばな【徒花】」の用例「ならぬ一真白に見えて、憂き中垣の夕顔や」(閑吟集)は、おそらく「眞白に…」と書かれていたであろうし、「出藍の誉れ」の出典となった『荀子』の巻名「勸學」は、『広辞苑』では「勸学」と書かれることになる。この規則は、人名や地名にも適用されるので、「はしご高」は「高」となるし、「邊」「邊」なども「辺」となる。

『広辞苑第五版 CD-ROM 版』では、この考えをさらに広げている。JIS第2水準までの文字についてはその字体を「通行の字体」と認めることにより、作成する外字の数を減らしている。作家「もりおうがい」は「森鷗外」として、カモメの左半分を「區」とした外字を作成してはいない。

このような手段をとった結果、『広辞苑』に使われている外字は、だいたい2000字となっている。このうち、非漢字が500字、そして漢字が1500字である。

解決方法その1

『広辞苑』のCD-ROM版は、EPWINGという規約に則って作成されている。CD-ROM上に日本語の電子出版物を載せるべく作成された規約であるから、その中には当然、JIS外字を扱う手段も規定されている。まず、その方法をみてみよう。

先に述べたように、JISのコード化の方法は、

94区×94点という2次元の表を用いたものである。これを2バイト平面といい、縦軸に最初の番号である区(第1バイト)、横軸に後ろの番号である点(第2バイト)をふる。また、実装においては、1~94という数字ではなく、16進コードで、21hから7Ehまでの94個の番号を用いる。(図1)

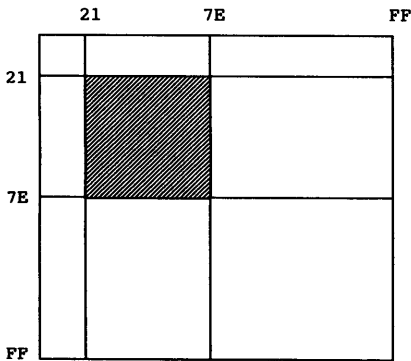


図1 JISの2バイト平面

図1の斜線部に、JISの第2水準までの文字が割り当てられているのだが、EPWING規約での外字はこの斜線部以外のエリアに割り当ててることとしている。外字の場合は第1バイトの最上位ビットをたてて

21h(00100001)→A1h(10100001)

22h(00100010)→A2h(10100010)

...

7Eh(01111110)→FEh(11111110)

とした上で、JISと同じ方法で2バイト平面が作成されているため、最大8836文字が定義可能である。図2の網点部のように位置する形になっている。

さらに、この拡張された2バイト平面は、タイトル(書籍)ごとに定義可能である。これは、EPWING規約の電子出版物が、タイトル(コンテンツ)とビューアー(検索ソフト)というセットでの動作を前提としており、その意味で閉じた世界を形成していることによる利点である。すなわち、EPWING規約の電子出版物は、JISの文字や「システム外字」「ユーザー外字」とらわれないことなく、それぞれが必要な文字を用いるこ

とができるようになっている。

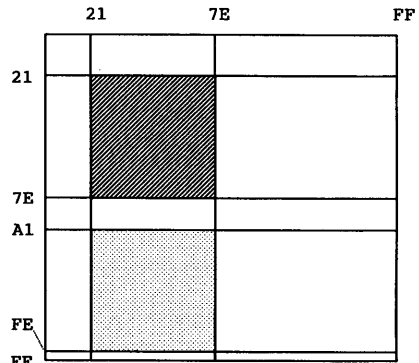


図2 EPWINGの2バイト平面

解決方法その2

もちろん、外字がタイトルごとに定義可能である、ということには欠点もある。タイトルAとタイトルBで、同一の外字コードに違う文字が割り当てられる(可能性がある)ためである。実際には、主としてコスト面から同一出版社のタイトルであれば同じコード割り当て(後から出版されたものは、それ以前のものに対して上位互換である)ということが多いが、それでも、出版社間での統一はなされていない。

この欠点に対する方法として二つの動きがある。一つは、いわば「EPWING 共通外字」とでもいうものを規定しようというものである。例えば、多くの辞典で語義の区分に使っている丸数字や、英和辞典などでは必須の発音記号、さらにはローマ数字など、各社がそれぞれ作成してきた外字には、共通のものが意外と多い。これらの文字の割り当て先を共通にすることで、出版社間の統一を一部でも図ろう、というものである。

もう一つ、興味深い動きとして、ビューアー(検索ソフト)側で、出版社間の差異を吸収しようというものがある。例えば、岩波書店が大日本印刷などと共同で開発した『こととい』というEPWINGのビューアーでは、EPWINGの外字割り当て方式はそのままに、『こととい』の内部コードを4バイトとすることで、タイトル間の差異を吸収し、さらにUnicode(これについては、ほと

んど触れてこなかったが)への対応まで可能とする仕組みを持っている。

検索の話

これまで触れてこなかったが、電子辞書においては、外字に関係してもう一つ問題が生じてくる。それは、電子辞書の根幹ともいうべき、検索に関することである。

例えば、「トウ小平」をその漢字表記から検索することは、トウが外字であるために、現在のところ不可能である(対処療法的に、タイトルを作成する側では、外字を含むキーワードは仮名にひらき、かつユーザーにその方法で入力してもらうという方法が行われている)。

異体字(先述した「渡辺」「渡邊」「渡邊」など)についても、一度で正しい結果を得ることが困難であることは容易に想像できよう。こちらの場合は、現在のところ、すべての異体字、すなわち、「渡辺」と入れると「渡邊」「渡邊」も検索してくる(他の場合も同様)ようにキーワードを作成する、ということしか解決方法がなく、しかも現実にはインデックスの増大を招くことになり、障害も大きい。

これらの点の根本解決をもたらすのは、大規模な漢字データベース以外にはない。実際にデータベースとフォントの作成を行っている文字鏡研究

会と、その漢字検索システム「今昔文字鏡」の取り組みに期待するところ大である。

終わりに

第1次規格が定められてすでに20年経っている。それにもかかわらず、実際のところJISの文字コード(と外字)の問題について、関心が高まってきたのは、たかだかここ数年のことにすぎない。例えば、1990年に刊行された横井俊夫『日本語の情報化』(共立出版)は、日本語の情報化とその技術について簡潔にまとめた好著であるが、残念ながら文字コードの問題には触れていない。

このような実態を、「日本語に対する無関心」や「戦後の国語改革の誤り」に結びつけて論ずるつもりはない。しかし、コンピュータがこれだけ普及した今、いわばインフラの位置にある文字コードについて、もはや無関心では済まされないのではないか。

しかも、そのインフラは、われわれ個々の利用者が操ることのできるものである。多様なニーズをもった利用者が直接システムに向き合うとき、要求を明確にしなければならなくなる。日本語に何を求めるのか、コンピュータで日本語をどう扱いたいのか、それらについての要求と判断が必要とされている。このことに意識的でありたい。

国内刊行用語辞典リスト(1998 後半)

凡 例

収録範囲

前号(第16号)収録以降(1998年8月から1998年12月)に国内で発行された専門用語辞典(集)を収録した。新語辞典、国語辞典、古語辞典、方言辞典などは収録対象としていない。

2. 情報源

Kinokuniya和書データベースBook Webほか

3. 書誌記述

記述項目は、書名、著者(编者)名、出版社名、出版年月、ページ数、価格、国際標準図書番号(ISBN)とした。

4. 配列

日本十進分類法(NDC)にしたがって、分類順に配列した。

5. 例示

分類見出し	件名								
↑	↑	書名							
007	情報処理	↓							
	コンピュータ技術基礎用語事典								
	新電気編集部.	オーム社.	97-12.	248p.	2,800円.	ISBN:4274945952			
	↓	↓	↓	↓	↓	↓			
	著者(编者)名	出版社名	出版年月	ページ数	価格	ISBN			

6. その他

このリストについてのご意見やお問い合わせは、編集委員会までお寄せください。

007 情報処理

第一種・高度情報処理用語辞典

金子崇. 経林書房. 98-10. 280p. 2,800円. ISBN:47673070233

初歩パソコンbooks 初歩パソコン用語集

月刊初歩のパソコン編集部. エグゼット・プラス(ビ・エヌ・エヌ発売). 98-10. 317p. 933円.
ISBN:4893696599

情報処理技術者試験 初級・上級システムアドミニストレータ図説用語集

松木則夫. 実教出版. 98-10. 169p. 1,700円. ISBN:4407051507

Visual Basic&VBA 用語辞典

杉松秀利. ナツメ社. 98-08. 399p. 2,000円. ISBN:4816324674

わかりやすいコンピュータ用語辞典

高橋三雄【監修】. ナツメ社. 98-08. 704p. 1,380円. ISBN:4816324402

誰でもわかるパソコン用語辞典

高作義明・川嶋優子. 新星出版社. 98-10. 430p. 1,600円. ISBN:4405056439

第二種・シスアド情報処理用語辞典

藤本喜弘. 経林書房. 98-12. 358p. 3,400円. ISBN:476730704X3

第二種・シスアド試験対応 情報処理用語辞典—新カリキュラム対応

福嶋宏訓. 新星出版社. 98-08. 453p. 1,600円. ISBN:4405025568

初級シスアドコンパクト用語辞典—新・標準カリキュラム完全準拠

中家裕之; 高度情報化利用技術研究会【編】. プレンティスホール出版. 98-08. 214p.
1,200円. ISBN:4894710552

最新パソコン基本用語辞典

ノマド・ワークス. 新星出版社. 98-12. 798p. 1,600円. ISBN:440505651X

210 考古学

日本考古学用語辞典 軽装版

斎藤忠. 学生社. 98-10. 568p. 3,800円. ISBN:4311750250

320 法律

会社法律基本用語集商法2第2版

早稻田司法試験セミナー. 早稻田経営出版. 98-08. 128p. 1,400円. ISBN:4847100557

330 経済

英和経済学用語辞典

多賀出版. 1998-11. 632p. 3,000円. ISBN:48115513111

国際協力用語集第2版

国際開発ジャーナル社. 98-08. 263p. 3,000円. ISBN:48753904838

最新英和ビジネス基本用語辞典

菊地義明. 洋販出版. 98-08. 495p. 2,800円. ISBN:4896846567

338 金融

会計学用語辞典

森川八洲男. 税務経理協会. 98-10. 385p. 3,800円. ISBN:44190279679

税務用語辞典六訂(銀行実務用語シリーズ)

経済法令研究会. 98-11. 430p. 3,500円. ISBN:4766814371

クレジット・金融用語辞典

西ヶ谷葉子. 金融財政事情研究会(きんざい発売). 98-09. 203p. 2,800円.
ISBN:4322160417

339 保険

保健医療用語事典

社会保険研究所. 98-11. 775p. 6,500円. ISBN:4789416089

340 財政

税法用語事典 3訂版

金子宏. 税務経理協会. 98-09. 548p. 4,400円. ISBN:44190309179

370 教育

和英教育用語集 日本の教育について英語で話すための

飯沢省三. 稲元印刷. 98-07. 235p. 1,400円. ISBN:4901136011

すぐに役立つ教職教養用語事典〈2000年度版〉(教員採用試験シリーズ)

下村哲夫. 一ツ橋書店. 98-11. 251p. 940円. ISBN:456500173X

430 化学

英日中化学用語辞典

工業調査会. 98-10. 1438p. 18,000円. ISBN:47693706606

マグローヒル化学用語辞典

日刊工業新聞社. 98-11. 631p. 5,000円. ISBN:4526042692

460 生物学

生物教育用語集

日本動物学会・日本植物学会【編】. 東京大学出版会. 98-09. 191p. 2,400円.
ISBN:4130622102

マグローヒルバイオサイエンス用語辞典

日刊工業新聞社. 98-10. 975p. 5,600円. ISBN:45260424636

491 医学

眼科手術用語集

日本眼科手術学会. 文光堂. 98-11. 148p. 1,800円. ISBN:4830655-707

生理学用語集

日本生理学会【編】. 南江堂. 98-08. 451p. 4,700円. ISBN:452420945X

500 工学・技術

応用物理用語大事典

応用物理学会. オーム社. 98-04. 1177p. 25,000円. ISBN:4274023648

マグローヒルエンジニアリング用語辞典

日刊工業新聞社. 98-12. 984p. 5,600円. ISBN:4526042854

507 特許

英和特許用語辞典 最新増補版

飯田幸郷. 発明協会. 98-10. 563p. 4,600円. ISBN:4827100330

509 工業経済・経営

絵で見てわかる工場管理・現場用語事典

日刊工業新聞社. 98-10. 183p. 1,800円. ISBN:4526042455

510 土木工学

景観用語事典

篠原修【編】；景観デザイン研究会【著】. 彰国社. 98-11. 307p. 3,600円.

ISBN:439510018X

全訂 都市計画用語事典

ぎょうせい. 98-10. 507p. 3,800円. ISBN:4324054266

520 建築学

イラストでわかる二級建築士用語集

中井多喜雄. 学芸出版社. 98-11. 333p. 3,000円. ISBN:476153074X

540 電気工学

電気工事絵とき基礎用語事典

オーム社. 98-11. 380p. 3,000円. ISBN:4274941965

最新日中英エレクトロニクス用語辞典

熊本工業大学・中国高等教育出版社. 熊本工業大学(発売:星雲社). 98-09. 1295p.
9,524円. ISBN:4795234205

550 海事工学

和英・英和船舶用語辞典 改訂第19版

東京商船大学船舶用語辞典編集委員会. 成山堂書店. 98-09. 591p. 5,000円.

ISBN:4425110145

海洋建築用語事典

日本建築学会. 彰国社. 98-12. 297p. 3,600円. ISBN:439510021X

560 採鉱冶金学

塑性加工用語辞典

日本塑性加工学会. コロナ社. 98-12. 450p. 6,000円. ISBN:4339043397

570 化学工業

日英中化学プラント用語辞典

重化学工業通信社. 98-10. 966p. 14,286円. ISBN:4880530565

日英中粉体技術用語辞典 vol.1

美洋(発売:重化学工業通信社). 98-11. 488p. 9,500円. ISBN:4944192002

580 製造工業

ワイン用語辞典(平凡社ライブラリー)

菅間誠之助. 平凡社. 98-10. 395p. 1,200円. ISBN:4582762646

590 家事

調理用語辞典 改訂

全国調理師養成施設協会（発売：調理栄養教育公社）. 98-12. 1468p. 6,571円.
ISBN:4924737356

610 農業

図解農業土木用語集 改訂版

東洋書店. 98-08. 560p. 3,500円. ISBN:4885952190

農業施設用語事典

農業施設学会. 筑波書房. 98-10. 391p. 4,000円. ISBN:481190172X

640 畜産業

ペット用語事典1999—2000年版 犬・猫編

どうぶつ出版. 98-10. 475p. 1,800円. ISBN:4924603481

670 商業

新広告用語事典

電通広告用語事典プロジェクト. 電通. 98-12. 268p. 2,700円. ISBN:4885531314

マーケティング・リサーチ用語辞典

日本マーケティング・リサーチ協会【編】. 同友館. 98-09. 219p. 2,000円.
ISBN:449602666X

700 芸術・美術

日本美術用語辞典 和英対照普及版

東京美術. 98-12. 793p. 8,000円. ISBN:480870658X

740 写真

楽しく読める写真用語事典 明快

玄光社. 98-10. 162p. 1,450円. ISBN:4768300928

780 体育・スポーツ

山歩きのための山名・用語事典（アルペンガイド別冊）

山と溪谷社. 98-08. 264p. 1,400円. ISBN:4635004643

800 語学

マチガイ言葉・あいまい用語辞典“日本語力”が身につく

河口章. 同文書院. 98-11. 294p. 951円. ISBN:4810375455

時事用語事典2000年版 就職試験対策必須本

池田書店. 98-10. 383p. 1,400円. ISBN:4262151697

新聞によく出る用語事典 最新版（知的生きかた文庫）

花田久徳. 三笠書房. 98-10. 389p. 752円. ISBN:4837909892

新聞を読むためのインドネシア語時事用語辞典

服部英樹・粕谷俊樹【編著】. 大学書林. 98-08. 368p. 4,200円. ISBN:4475000920

900 文学

最新 文学批評用語辞典

川口喬一・岡本靖正編. 研究社出版. 98-07. 344p. 2,800円. ISBN:4327461342

1999年研究報告会 ご案内

例年秋に開催しておりますシンポジウムを、本年は、情報知識学会との共催で、きたる5月に開催します。現在、下記テーマで発表テーマを募集しておりますので（締切り3月5日）、プログラムができましたら、改めてご案内いたします。会員各位のご出席をお待ちしております。

日時 5月22日（土）

会場 東京・凸版印刷会議室

予定の発表テーマ内容

- 1) 情報知識の構造解析、モデル化、意味理解、自己組織化、可視化
- 2) 情報・知識の表現、生産、組織化、検索、提供
- 3) 電子出版、電子図書館、フルテキストデータベース
- 4) マルチメディア、電子ミュージアム
- 5) 用語、シソーラス、電子化辞書、機械翻訳
- 6) デジタル・コンテンツの流通技術・体制と著作権処理
- 7) 専門分野における情報管理、レコードマネジメント
- 8) その他情報知識学、専門用語、図書館・情報学に関する諸研究・開発

J I S用語規格制定の動き

情報源：“標準化ジャーナル”，1998年1月号-1998年12月号

制定・改正されるJ I S

- 歯車-歯面の摩耗及び損傷-用語 (制定)
Gears - Wear and damage to gear teeth - Vocabulary
- 球面滑り軸受-用語 (制定)
spherical plain bearings - Vocabulary
- 機械振動と衝撃-人体暴露-用語 (制定)
Mechanical vibration and shock - Human exposure - Vocabulary
- 情報処理用語-人工知能-基本概念及びエキスパートシステム (制定)
Information technology - Vocabulary - Artificial intelligence - Basic concepts and expert systems
- A0002 建築モジュール用語 (改正)
Glossary of terms used in building module
- A0203 コンクリート用語 (改正)
Concrete terminology
- B0102 歯車用語 (改正)
Glossary of gear terms
- B0148 巻上機 (改正)
Lifting equipments - Terminology
- B0132 送風機・圧縮機用語 (改正)
Glossary of terms for fans, blowers and compressors
- B8032-1 内燃機関-小径ピストンリング-第1部:用語 (制定)
Internal combustion engines - Small diameter piston rings - Part 1: Vocabulary
- B8037-1 内燃機関-大径ピストンリング-第1部:用語 (制定)
Internal combustion engines - Large diameter piston rings - Part 1: Vocabulary
- C0364-2-21 建築電気設備-第2部:用語定義-第21章:一般用語の指針 (制定)
Electrical installations of buildings - Part 2: Definitions - Chapter 21: Guide to general terms
- D6201 フォークリフトトラック用語 (改正)
Glossary of terms relating to fork lift trucks
- H0201 アルミニウム表面処理用語 (改正)
Glossary of terms used in the surface treatment of aluminium
- H0400 電気めっき及び関連処理用語 (改正)
Glossary of terms used in electroplating and related processes
- R6004 研磨材, 研削といし及び研磨布紙の用語及び記号 (改正)
Glossary of terms and marks used in abrasives, grinding wheels and coated abrasives
- R9200 せっこう, 石灰及びマグネシアセメント用語 (改正)

- Z3500 Glossary of terms related to gypsum, lime and magnesia cement
粉末や(冶)金用語(改正)
Powder metallurgy terms
- Z4001 原子力用語(改正)
Glossary of terms used in nuclear energy
- Z8106 音響用語(一般)(改正)
Glossary of acoustical terms (general)
- Z8107 音響用語(機器)(改正)
Glossary of acoustical terms (transducers and instruments)
- Z8113 照明用語(改正)
Lighting vocabulary
- Z8114 製図用語(改正)
Glossary of technical drawing terms
- Z8126 真空用語(改正)
Vacuum technology - Vocabulary

官報告示・発行されたJ I S

- A8308:98 土工機械-基本的機種-用語(改正)
Earth-moving machines - Basic types - Vocabulary
- A8420-1:98 土工機械-トラクター第1部:用語及び仕様項目(制定)
Earth-moving machines - Tractors - Part 1: Terminology and commercial specifications
- A8421-1:98 土工機械-ローダー第1部:用語及び仕様項目(制定)
Earth-moving machines-Loaders - Part 1: Terminology and commercial specifications
- A8423-1:98 土工機械-グレーダー第1部:用語及び仕様項目(制定)
Earth-moving machines-Graders - Part 1: Terminology and commercial specifications
- B0111:97 プレス機械-用語(改正)
Presswork machinery - Vocabulary
- B0134:98 産業用マニピュレーティングロボット-用語(改正)
Manipulating industrial robot - Vocabulary
- B0144:97 電子部品実装ロボット-用語(制定)
PCB (printed circuit board) assembly robots - Vocabulary
- B0155:97 工業プロセス計測制御用語及び定義(改正)
Glossary of terms used in industrial process measurement and control
- B0181:98 産業オートメーションシステム-機械の数値制御-用語(改正)
Industrial automation system - Numerical control of machines - Vocabulary
- B0660:98 表面粗さ-用語-第1部:表面及び表面粗さパラメータ(制定)
Surface roughness - Terminology - Part 1: Surface and its parameters
- B3000:97 FA-用語(改正)
Factory automation - Vocabulary
- B3551:98 プログラマブル表示器-用語(制定)

- Programable display - Vocabulary
 B8650:97 プラスチック加工機械－用語(改正)
 Plastics moulding machinery - Vocabulary
 C0161:97 EMCに関するIEV用語(制定)
 International electrotechnical vocabulary - Electromagnetic compatibility
 D0004-1:98 土工機械－スクレーパー第1部：用語及び仕様項目(制定)
 Earth-moving machines - Scrapers - Part 1: Terminology and commercial specifications
 E4019:98 特殊鉄道車両用語(制定)
 Specified railway (guideway system) car - Vocabulary
 F0013:98 造船用語－船体－外ぎ装(改正)
 Shipbuilding - Vocabulary - Hull part - Outfittings
 F0015:98 造船用語－船体－内ぎ装(改正)
 Shipbuilding - Vocabulary - Hull part - Accomodation and hold
 F0031:98 造船用語－電気(改正)
 Shipbuilding - Glossary of terms - Electric
 F0041:98 造船用語－特殊船－種類(改正)
 Shipbuilding - Vocabulary - Special vessel - Classification
 F0042:98 造船用語－特殊船－機器(改正)
 Shipbuilding - Vocabulary - Special vessel - Equipment
 H0500:98 伸銅品用語(制定)
 Glossary of terms used in wrought coper and copper alloys
 H8200:98 溶射用語(改正)
 Thermal spraying terms
 K6200:98 ゴム用語(改正)
 Rubber - Vocabulary
 L0204-1:98 繊維用語(原料部門)－第1部：天然繊維(制定)
 Glossary of terms used in fiber - Part 1: Natural fibers
 L0204-2:98 繊維用語(原料部門)－第2部：化学繊維(制定)
 Glossary of terms used in fiber - Part 2: Man-made fibers
 L0204-3:98 繊維用語(原料部門)－第3部：天然繊維及び化学繊維を除く原料部門(制定)
 Glossary of terms used in fiber- Part 3: Materials except natural and chemical fibers
 L0306:98 製織機械用語(改正)
 Glossary of terms used in weaving machinery
 P0001:98 紙・板紙及びパルプ用語(改正)
 Paper, board and pulp - Vocabulary
 Q14050:98 環境マネジメント－用語と定義(制定)
 Environmental management - Vocabulary
 R1600:98 ファインセラミックス関連用語(改正)
 Glossary of terms relating to fine ceramics
 T0102:98 福祉関連機器用語(制定)
 Glossary of terms used in rehabilitation equipments for the

- physically disabled or handicapped persons
- X0013:98 情報処理用語 (図形処理) (改正)
Glossary of terms used in information processing (computer graphics)
- X0024:98 情報処理用語 (計算機統合生産) (制定)
Glossary of terms used in information processing (computer- integrated manufacturing)
- Z0106:97 パレット用語 (改正)
Glossary of terms relating pallets
- Z0110:98 産業用ラック用語 (改正)
Glossary of terms for industrial racks

編集後記

◆「広辞苑」の外字

「文字シンポジウム」には、正直に申せば気が進まぬ思いで参加した。私が常用漢字の枠に規制された「学術用語集」の世界に常住していたからであろう。しかし、スケールの大きな報告が次々に続いて、おおいに蒙を啓かれ非常に楽しかった。

諸報告のおかげで頭がかなり柔らかくなり、総会後の懇親会で余韻を楽しんでいるうちに、ふと「広辞苑」のことを思い浮かべた。普通の国語辞典と違って、「広辞苑」は使用漢字に枠をはめていないと記憶していた。第三版から電子編集になったが、漢字使用に制限が行われるようになったとは聞いていない。新しく刊行された第五版ではいったいどれだけの字種を使用したのだろうか？ 何人かの人にもこれを知りたいと同調したので、岩波の辞典部に問い合わせしてみた。

同じような質問がすでに多く寄せられたようで、JIS 第二水準以外に約 1,700字とか責任者が発表しているとのことだった。企業秘密事項ではないようなので、事情に詳しい人に寄稿を依頼した。

上野さんには、シンポジウムの内容をろくにお伝えもせず報告を強要して申しわけなかったが、シンポジウムで触れられなかった領域の説明もして下さって良い報告をいただき、喜んでいいる。新版刊行後も関連したことで寸暇もない様子なのに無理をしていただいた。

歴史的な事項が相当豊富な「広辞苑」で、外字漢字が 1,500種とは意外であった。第二水準までの文字が古い文化事象を記載する上ですでに相当抱擁力があつたということなのだろうか？ 五版で取り入れた外字はどんな字種なのかさらに具体的に知りたいと思う。

旧字体、異体字には私は二次的な関心しかない。文字種がおおよそ 6,355+1,500 で用が足りたということだが、ではボディになっている6,355種の第二水準までの文字でほとんど要らなかつた文字がありましたか？ そうした統計まで作れるのならとてもうれしいことだと思う。(牧野)

◆やや予定よりは遅れたけれども、電子情報通信学会編の用語辞典の原稿のチェックが終わり、コロナ社の手に渡った。99年夏には世に出ると思うとうれしい。テレビジョン学会から用語辞典を出して大分年月が過ぎた。10年ひと昔など言っておられない速度で変化し、増加する言葉の処理は大変である。昨年、健康番組に、聴覚障害者のために字幕を書くにあたって、南山堂の医学辞典を買った。18版約30,000語あるから、かなり大きな本である。こうなると多くの人が筆を執ることになり、かつ、前の版から、そのまま流用する言葉の説明文と、新しく採用した語のそれを含め、文の構成、表現に、やはり気になるところが出てくる。編集部の方の苦勞がわかる気がする。この点においては、新規に版を起こす方が楽である。

前述の電子情報通信用語では、文の構成について、良い例、悪い例を示して執筆を依頼したが、中には編集幹事の権限で、文の構成、表現を変えざるをえなかつた説明文が存在した。やはり辞典は、一つの商品であるので、読みやすく、わかりやすいこと、説明文の形式が揃っている方が、価値が上がるはずである。

いつのまにか、時間がたってしまったが、1993の「シンポジウム 専門用語の生態学」(会報No.5)での京都大学・長尾 真教授の講演は、岩波情報科学辞典を含め、何回読んでも教えられるところが多い。勉強に限りはないと、自らを励ます今日この頃である。(中山)

◆「ニュースレター」(No.23, 98-12)でお知らせしましたように、新会長、理事の分担が決まり、編集委員会に仲本氏が新たに入られました。本号発行後には、新委員会のもと99年度の企画検討に入ります。会員皆様からのご提案・ご意見をいただければ幸いです。

◆例年秋に開催のシンポジウムは、今年は、情報知識学会との共催で、きたる5月22日(土)に開催されます。近く詳細なご案内を差し上げますので、会員皆様のご出席をお待ちしております。

(戸塚)

編集委員会委員

太田 泰弘 文教大学
四ノ宮明夫 大正製薬
戸塚 隆哉 KMK デジテックス (委員長)
仲本秀四郎 I R I S
中山 亮一 リョウ・プロダクション
牧野 正久 東京理科大学
山下 泰弘 学術情報センター
山本 昭 国学院大学

専門用語研究 第17号

(1999年2月28日発行)

発行所 専門用語研究会

〒102-0082東京都千代田区一番町4-6

一番町中央ビル2F

日本総合技術研究所(JIST)内

Tel. 03-3262-8956

Fax. 03-3262-8960