

## 【翻 訳】

## 銭乃栄主編《現代漢語》(重訂本) 第七章 漢字と漢字文化

## 第四節 漢字の簡化、標準化と漢字のコード化

## 甲 斐 勝 二

## 翻訳にあたって

以下に訳出するのは、銭乃栄主編《現代漢語》(重訂本・2008年6月第1版・鳳凰出版伝媒集団江蘇教育出版社)の中、第七章《漢字と漢字文化》第四節《漢字の簡化、標準化と漢字のコード化》の部分である。執筆の担当は石汝傑氏(執筆当時は蘇州大学、現在は熊本学園大学教授)で、中国語学の研究者として日本でも著名である。

中国では《現代漢語》という名で各種の書籍が様々な出版社から出版されているが、漢字の直面する現在の状況についての記述は、訳者の手元にある数種のものの中では、この書籍が一番詳しく新しい。しかも漢字問題について非常に幅広い領域から触れている。おそらく、漢字の問題に興味を持つ方なら興味もあろうと、訳出を試みた。

漢字は、篆書から隸書が変わるときに形態上大きな変化を経た。隸書ができて以後やがて楷書が生まれても、隸書の形体から基本的に外れるものではなかったから、その後は二千年以上もそこで生まれた形をほぼ変えずに使われて来たことになる。日本に漢字が輸入されたのもこの楷書主流の時代である。ところが、漢字の老家中国では、20世紀の半ば過ぎ人民共和国の成立後、文字改革を行い、筆画を減少させた簡体字を新たな文字として採用し、従来の漢字は繁体字として必要な部分で限定的に使用されるようになっていく<sup>\*1</sup>。これは、篆書から隸書になったほどの変化だとみる事ができる。漢字史の上から見れば、大変な変化の時代に我々はいるわけだ。

篆書から隸書が変わったとき、そこにはそうなるべき必然的な理由があったように、繁体字から簡体字になったのもちゃんと理由があった。文字学習の容易化による識字率の向上そして人々の知識水準の向上のためである。この簡体化の作業によって、日本の漢字との形態上の差異が大きくなった結果、中国の簡体字に不満を持つ

人々も中にはいるが、同様の視点は中国にもある。簡化が一定程度まで来てその後ほとんど進んでいないのもまた理由があることだ。この点については訳出した文章でも触れているので参考にされたい。

また、現代の情報処理社会における漢字処理の問題もここでは触れられている。訳者は曾て学生から中国での漢字のコンピュータ入力の方法を問われたことがある。これは、漢字の検索方やコード化の問題にかかわってくる。訳者は中国語教科書などで現在発音表示の標準となっている拼音<sup>\*2</sup>ローマ字で打ち込む方法に慣れているので、それ以外には詳しくは知らなかったが、中国ではその他漢字の特徴に注目した各種のコード化による各種の入力方法が考案されていることが、訳文を読めばわかる。漢字のコード化の問題は、漢字検索法とも関係して漢字のコンピュータ処理の上で現在解決を待つ大きな問題である。この問題は漢字の成り立ちに基づく昔ながらの字形・字義・字音研究と漢字を切り離し、記号としての入力の快速性とか識別性などをめざす処理方法の研究を導く。これは、漢字学の中に、従来型の漢字の原型やそのその字義を探る領域に加えて、将来の漢字使用に有用な文字使用・文字検索方の領域がくわえられたものといつてよい。ここではその字義に意味のある形体の組み合わせである表意記号の合体として漢字を扱うのではなく、形体に多く注目される事が起こっている事はやはり注意しておきたい。これは、漢字が既に単に文字として記号化されていることを物語るからだ。そこには、日本の表記法のように発音を示す仮名と、意味の保守性を強く持つ漢字を共に使用し、仮名を「表音文字」、漢字を「表意文字」として認識しがちな視点とは別の視点が漢字に向けられていることが示唆されよう。検字法、つまり漢字の探し方についても、日本では漢和辞典の引き方として、画数引き・音訓引き・部首引き程度で固定しているが、中国では漢字を効果的に探すために、各種の検字法が生み出されていることも、訳文の中に示される。

\*1 これに対して台湾では、現在でも繁体字が継続使用されている。

\*2 拼音：《漢語拼音方案》に基づく表音文字体系の略称。1958年に公布、漢字の発音を、子音・介音・主母音・声調の5種の部分に分け、それぞれローマ字と記号を使って示す方式。現在では、漢字文の導入教材や、発音表記、検字法など、各所で用いられると共に、ローマ字型キーボードの入力も容易になった。

日本では、このような文字問題と言えば、漢字使用制限がすぐ思い浮かぶ。学校教育を除けば、実際には漢字の使用はかなり自由で、その読み方などは、かなり大胆に行われているのが実際の状況だ。現代語による漢字の訓読を漫画などではよく見かけるのは、「文選読み」の伝統にあると言うべきなのだろう。個人の名前も読みづらい難しい名前が増えてきた。中国でも、小学生の作文などで街の看板の誤用字について触れているから、事情は似たものなのかもしれない。しかしながら、中国の言語政策は日本のそれに比べて昔からかなり熱心に見える。たとえば、普通話（共通語）の普及推進のために国家通用文字法が制定されたり、普通話水準を計る試験が行われて、朗読にも一定の成績が求められる点などは、日本とは違うところだ。多くの民族、多種の言語、更に言えば各種の方言を域内に持ち続けてきた中国ならではのものであろう。

訳出においては、石先生に訳出の許可を得たが、漢字のコード化やコンピュータへの入力法など訳者にとっては初めてのものも多く、勝手な理解による誤訳を恐れる。御指正をお待ちします。

なお、固有名詞や組織の名前などに用いられる漢字は、日本人が読みやすいように、その必要性がなければ、中国の簡体字は日本語の対応する漢字に改めている。字形が問題になるような場合は、簡体字はそのまま使った。《》の記号は独立した書名や篇名を示す。日本の示しかたより便利なので、これを使う。脚注は、訳者による。

銭乃榮主編《現代漢語》（重訂本）第七章漢字と漢字文化  
第四節

《漢字の簡化、標準化と漢字のコード化》

一 漢字の簡化について

明清時代より俗体字、略字は木版印刷や文人・庶民の手によって絶えず現れ伝わってきた\*3。五四時期\*4以後、多くの有識者が主張したのは、世間で通用している略字を収集整理し、それを押し広めることであった。

1952年、中国文字改革委員会\*5が成立し、漢字の簡化及び整理の仕事を責任をもって進めることになった。

簡化字を試用した公に意見を徴集してのち、1956年1月、国務院は正式に《漢字簡化方案》を公布する\*6、そこには515の簡化字と54の類推可能な簡化偏旁が含まれていた。1964年にはそれを編成して《簡化字総表》\*7としている。1986年に再度公布したとき、今度は個別の字に調整を加えていた。第一表には簡化偏旁に用いない350の簡化字を収め、第二表には偏旁に用いる132の簡化字と14の簡化偏旁収め、第三表には第二表に並べられた簡化字と簡化偏旁によって作られた1753の簡化字が収められた。加えて第三表に未収録の文字で、第二表の簡化字或いは簡化偏旁を偏旁とするものはすべて、通常は同様に簡化せねばならないと規定したのだった。

《簡化字総表》の三種の表によって統計を取れば、2235の簡化字で2261の繁体字を取り替えたことになる。

《漢字簡化方案》の前の二表（544の繁体字が簡化の後515の字になった）の中の漢字に対し、その筆画数の統計を取った人がいた。

	筆画 総数	各文字の 平均画数	十画以下	十一画	十二画 以上
簡化前	8745画	16.8画	34字	33字	457字
簡化後	4602画	8.16画	409字	33字	71字

統計の結果では、簡化を経て、漢字の筆画数は半分に減っている。

漢字の簡化の方針は“世間の利用習慣に従い、ゆっくりと進む”というものだが、それは絶対多数の簡化漢字が人々の表記の実践のなかで生まれ伝えられてきたからだ。これを帰納すれば、概ね以下の幾つかの簡化方法にまとめられる。

- (1) 偏旁を変えてしまうもの：  
禮-礼 聯-联 難-难 劉-刘
- (2) 部分を取り除くもの：  
務-务 號-号 奪-夺 醫-医
- (3) 同音による取り替え：  
鬱-郁 穀-谷 醜-丑 後-后
- (4) 草書体の楷書化によるもの：  
樂-乐 為-为 書-书 舉-举
- (5) 古体を採用するもの：  
雲-云 鬚-胡 網-网 從-从

\*3 こういった簡略字・俗字は、所謂異体字（指し示す意味や音は同じだが形体が異なる）として扱われてきた。口語表記の文献に多く見られる。歴代の俗字を集めたものとしては、劉復・李家瑞編《宋元以来俗字譜》が、敦煌文書の俗字を集めたものとして、潘重規編《敦煌俗字譜》が著名。

\*4 五四時期：時代としては1915～23年ごろ。銭元同は、簡化漢字の使用を主張してを得ている。この結果当時の国民党政府も第一批簡体字表を公布（324字）することになった。

\*5 中国文字改革委員会：1949年に中国文字改革協会が成立（常務理事は呉玉章）し、その後1952年に文化教育委員会の下に中国文字改革研究委員会が作成される（主任・馬叙倫、副・呉玉章）、1954年に国務院の直屬機関として中国文字改革委員会に改編（主任呉玉章）される。

\*6 《国務院關於公布漢字簡化方案的決議》（1956年1月28日国務院全体会議第23次會議通過）参考資料1として、1956年《中国語文》2月号に載せられた簡化字方案の一部を付録に掲載しておく。漢字簡化第一表、漢字簡化第二表、漢字偏旁簡化表の三部分からなる。

\*7 《簡化字総表》：原著の《現代漢語》では付録として第一表、第二表がつけられているが、初めの部分だけを参考資料2として後ろに附録。

(6) 別に新しい文字を作ったもの:

- a 会意法: 塵-尘 筆-笔 竈-灶  
衆-众  
b 形声法: 護-护 驚-惊 懼-惧  
欄-栏

(7) 適切な類推:

- 東-东: 练 炼 拣 將-杲: 浆 浆 浆

(8) 偏旁の簡化:

- 貝-贝: 贲 贼 贯 愛-爱: 暖 媛 媛

ここで、一般漢字の簡化の偏旁として用いるものは、《簡化字総表》の第二表に並んだその字と偏旁だ、ということには注意したい。この第二表にない要素は勝手に類推して使ってはならないのである。例えば“东”という部分は、“練・煉・拣”という繁体字に用いて、“练・炼・拣”と簡化することができる、しかし、これら幾つかの文字は皆あらかじめ準備された第一表内の簡化字だから許されるので、“东”の簡化字から類推して“瀾(瀾)”“瀾(瀾)”を“瀾”“瀾”と書いてしまうことは許されない\*8。また、“盧”“瀘”“鱸”“鱸”等は簡化して“卢”“泸”“炉”“驴”となり、“盧”“蘆”“爐”“驢”の簡化字は“卢”“芦”“炉”“驴”となる\*9。“償”は簡化して“偿”となるが、“賞”は“尝”に簡化できない。というのは“尝”は“嗜”の簡化字だからだ\*10。これらの混乱は許されない。

《簡化字総表》第一表・第二表の中の簡化字の絶対多数は常用字であり、第三表中に類推されてできる漢字もまた常用字であるので、漢字が簡化されて後は使いやすくなり、漢字の書写速度を向上させることになった。漢字がこの簡化の段階を経たことで、多くの学者は、形が複雑すぎるという問題は基本的に解決されたと考えている。筆画数がかかり多いが常用されないような漢字群を簡化するかどうか、これはもはや緊急な問題ではないし、筆画を少なくしすぎると形体の差異度が下がって読む効率が悪くなくなる。現代社会で主要な問題は、漢字

の標準化を重めることだ。マスメディアの中、人々の書記活動の中、共に規範化された漢字が使われねばならず、自分で勝手に簡略字を作るという非社会的な現象はしっかり克服すべきである。同時に、文化遺産や国際交流を更に勉強し理解しようとするれば、繁体字を知ることも必要なことである\*11。

## 二 漢字の標準化について

標準化は現代社会が協力的にまとまるために必要な手段である。漢語は国連の公用語の一つであって、中国と世界各国との関係強化及び中国の国際的地位の向上に従い、漢語漢字を学ぶ人は日増しに増え、今や漢字に対する標準化の要求は切迫さをまし続けている。中国の漢字の標準化も漢字文化圏の強化と発展を推し進めるだろう。国内的にいうなら、漢字の標準化は文明教育の普及、普通話の推進に対しても重大な作用を持つ。更に重要なのは、現代社会が既に情報化時代に入っており、漢字の標準化は、コンピュータの中国語化を助け、中国語の情報処理問題を解決し、中国の現代化の足取りを速めることだ。よって、標準化とは、現代化された生活が文字に求めた要求なのだ。

漢字の標準化は、つまりは現代漢字に対して、全面的で体系的また科学的な整理を行った基礎の上に立ち、現代に通用する漢字の分量を定め、形体を定め、音韻を定め、順序を決めることなのである。

### 1 分量を定める

分量を定めるとは、つまり現代漢語で使う文字の数を決めてしまうことだ。

甲骨文字\*12より今まで、漢字は六万字以上の字形を積み重ねてきた。しかし、現代漢語を記すには、大体六七千字もあればよい\*13。コンピュータによる現代文献資料の統計を利用すれば、“現代漢語常用字表”と“現代漢語用字全表”を作ることができる。

1981年には中華人民共和国標準《情報交換用漢字コー

\*8 つまり、东の字は第二表には含まれていないので使えない。第一表に出ればかまわないが、それ以外の文字には応用できないということ。ある漢字の要素が簡化されたとき、その漢字の要素を含む別の漢字の扱いをどうするかということへの指示。簡化後の漢字をすべて表に掲げるわけには行かないので、このような形で原則を示している。

\*9 “盧”の簡化字は“卢”だが、これは第二表に含まれているので、第一表にない盧を含む漢字でも、この簡化字を使って書き直すことができる。つまり、“盧”“蘆”“爐”“驢”の簡化字は既に第一表に“卢”“芦”“炉”“驴”とあるが、その第一表にない“瀘”“鱸”“鱸”等も簡化して“泸”“炉”“驴”と書き直せるので、簡化字体の応用として第三表に挙げられることになった。

\*10:《簡化字総表》には、こういった簡体化以後の注意の他に、固有名詞の場合や、使用法について違いがある場合などそれぞれ注がつけられている。付録を参照。

\*11: 現在、簡体字は中国では漢語教育の普及でかなり広まっているが、香港や台湾、韓国などでは漢字と言えは通常繁体字である。また、繁体字は過去二千年に渡って使われて来ており、その間膨大な文献が繁体字で遺されている。よって、本格的に中国学を試みるとすれば、繁体字についての理解も必要であること論を待たない。

\*12 甲骨文字: 19世紀の末に発見された殷周時代の文字。亀の甲羅や牛の骨などを使った占いの表記に使われた文字で、現在中国で確かめられている漢字の最古の形体。

\*13 六七千字もあればよい: 2005年度版《中国言語生活状況報告》(下) 国家言語資源監測与研究中心編(2006.6商務印書館)の《新聞雑誌・ラジオテレビ・インターネット用字総表(報刊、広播電視、網絡用字総表)》によれば、1000字で91.16%、3000字で99.65%、4000字で99.23%、5000字で99.99%(小数点三桁四捨五入)の出現頻率であることが示される。

ド字符集基本集》(GB2312-80) が公布され、全部で現代漢字6763字が収められた\*14。二つのクラスに分けられ、第一級3755字が常用字とされ、現代の出版物の漢字中の出現率は99.9%である。第二級3008字が準常用字となり、現代の出版物の漢字中の出現率は約0.1%である。この字符集はコンピュータの中国語情報交換のために準備されたものだが、概ね現代漢字の通用字の数を代表するものである。

1988年1月国家語言文字工作委员会と国家教育委员会は共同で《現代漢語常用字表》を公布した\*15。その字表の第一表には常用字2500が収められ、第二表には準常用字が1000収められた。この字表は語文教学・辞書編纂・漢字機械処理及び情報処理などの面で常用字の根拠及び標準を打ち立てた。1988年3月、国家語言文字工作委员会と新聞出版署は共同で《現代漢語通用字表》を公布し、通用漢字7000字を収録した\*16。この表は主に出版印刷・情報処理等の面での用字の需要に対応するために作られたものだった。

“通常文字表”・“常用字表”を作り上げるほかに、更に“現代漢語用字全表”も作らねばならない。全表は常用字を収めるほかに、あまり使われない文字も収録し、使用瀬率によって文字を収録するのではなく、実際に使用する必要によって収録するのである。全表を作り上げるときには幾つかの問題を処理する必要があるが、この解決が難しい。例えば、現代漢語の用字を古漢語の用字から区分すること、例えば“踟躕不前(躊躇して前進しない)”の“踟躕”、“愆前毖后(過去の失敗を教訓として今後を慎重にする)”の“愆后”などを表の中に含めるかどうかなどは、原則を打ち立てて、慎重に処理しなければならない。このほかに更に特殊な用字をうまく処理すべきだ。例えば科学技術用文字、人名と地名の専用文字、少数民族の名称専用文字、企業特殊文字、音訳専用文字、方言文字、外国産文字(日本の国字など)だ。

現代漢語の用字総量が確定すれば、これらの漢字の使用頻度に基づき、コンピュータを利用して瀬率調査作業を行い、現代漢字のクラス別数量表を作成しなければならない。関係部門が、文芸作品、ニュース報道及び雑誌、口語材料、自然科学と工学、社会科学と哲学という五つの分野の材料に基づき、抽出統計を行ったところ、現代漢語を使って書かれる文章用字はかなり決ったものであ

ることが、以下の表の用に示されている。

頻度順 番号	10	40	160	950	2400	3800	5200	6600
累計 頻度	11%	25%	50%	90%	99%	99.9%	99.99%	99.999%

国家計測委员会の“漢字情報処理工程”事務室が作った《漢字頻度表》統計によれば、一級文字の500字で、すべての出版物及び日常用字の77.419%を占める。これに二級字500字を加えれば、つまり前部の1000字で、90.819%を占め、もしさらに三級字の500字を加えると、前部の1500字で95.898%を占める。四級字は1500字でも、全部の出版及び日常用字の3.699%しか占めないし、五級字は2991字でも、0.4%を占めるだけだ。この統計の数字から、漢字の使用頻度に大きな差があることが分かる。

別の測定の結果によれば、2011076字の言語資料の中、《現代漢語常用字表》中の3500字による占有率は99.4%で、その内、常用字2500字の占有率は97.97%で、準常用字1000字の占有率は1.51%である。

漢語の文字利用が集中するのは、現代漢語の大多数の言葉が常用字によって構成されているためである。中国人民大学の言語文学研究所の統計資料では、語彙形成能力の最も強い十字で、1338の単語を作ることができ、50の常用字で5577の単語を、100字で9250の単語を、3736字で45000余りの単語を構成でき、これで現代漢語の総単語数の90%前後を占める。4990の文字があれば《現代漢語詞典》\*17の中のほとんどの単語が構成されてしまうのだ。

《現代漢語常用字表》中の2500字は、小学校国語課の識字教育の“小学生用漢字”の標準とすることができよう。《漢字頻度表》の前部1000字は、“識字教育用漢字”とすることができよう。こうすれば、中国の識字教育\*18は科学的基礎に基づいて構成されることになる。

## 2 形体を定める

形体を定めるとは、漢字使用の統一体を決めることである。定形の仕事は二つの方面の内容を持っている。異体字の整理と別字の整理である。

異体字というのは、漢字の形態変化の過程で生まれた同音同義ながら形の異なる漢字だ。あるものは四つも五

\*14 国家標準総局発布1981年5月1日実施 16区から87区の区画に1区94字枠でまとめられ、区画の二桁の数字と配列される二桁の数字計四桁の数字でコード化する。内55区までがA～Zまでの拼音の発音順に並べられた第1級文字となる。参考資料7参照。

\*15 通知には、“応用語文教学及びその他の方面の需要のために、《現代漢語常用字表》を作って発布する”とあり、説明の中では基づいた統計資料や計算方法などが記されて、常用字として2500字が、次常用字として1000字が示される。参考資料3に一部を示す参照。

\*16 1965年に発布された《印刷用漢字字形表》(6196字)を実際の需要に合わせるため、そこから50字を除き、新たに854字を追加してできたもの(通知前書きによる)。参考資料4に一部を示す。

\*17 《現代漢語詞典》: 人民共和國成立後、商務印書館より1973年に初版が出版され以後五版を重ねる、中国を代表する現代中国語辞典。

\*18 中国の識字教育: 漢字を言語表記に利用している中国では、現代化を妨げるものとして以前より識字率の低さが問題になっていた。そこには、漢族以外に少数民族も域内に含む中国の事情もある。文字改革の目的がそもそもこの識字率の向上にあること論を待たない。この識字運動については、大原信一《中国の識字運動》(東方書店1997年)に詳しい。

つもある。例えば：“劫・劫・劫・劫”、“窗・窓・窓・窓・窓・窓”である。

異体字の生まれた主な理由は以下の通り：

- a 古今字\*<sup>19</sup> 例えば— 礼—禮 岄—时 艸—草
- b 繁簡字\*<sup>20</sup> 膽—胆 髣—彷彿 會—会
- c 声旁・形旁\*<sup>21</sup>の取り替え可能による、異体字の発生  
咏—詠 擘—譚 猪—猪 姪—侄 跡—蹟  
稷—棕 稭—楷
- d 造字法\*<sup>22</sup>の違いによるもの  
磊(会意)—礫(形声) 野(形声)—埜(会意)
- e 偏旁が位置を変えたもの  
峯—峰 碁—棋 碁—期 够—夠 群—羣
- f 仮借字\*<sup>23</sup>に形旁がついたもの  
念—唸 贊—讚
- g 部分・筆画の変形  
吳—吳 賴—賴

一字多形は、いたずらに漢字利用の困難を増すので、必ずや整理を加え、規範の字形を確定し、その他の書き方をやめねばならない。1955年、中国文字改革委員会と文化省は、《第一次異体字整理表》を発表し、異体字810組、全1865字を配列し、1055の異体字を取り除いた。1986年10月新たに発表された《簡化字総表》と1988年3月に発表された《現代漢語通用字表》に基づくと、“沂 瀟 晔 蒼 河 鰭 紬 划 鯨 匡 讎 翦 邱 於 澹 幣 衍 菰 溷 徽 薰 黏 桉 愣 暄 凋”の26文字が規範文字として認められて、異体字としての淘汰を受けなくなった。かくして、《第一次異体字整理表》は810組の異体字から796組に減り、淘汰された異体字は1027に減ることになる。これ以前に、“阪 控”の2字も規範字に復帰している。1956年から1964年までの間、國務院の批准により、常用字によって地名の中の非常用文字が取り替えられて、30以上の地名の用字が改められ、対象となる非常用字が取り除かれた。このほかに、《簡化漢字表》中では同音取り替えの方法で漢字を簡化\*<sup>24</sup>し、幾つかの繁体字を淘汰した。すべて、1185の文字が整理されたのである。

異体字を整理する原則は、一般での使用法や簡便な方

法に基づくというものだ。例えば以下に挙げた( )の中  
の文字が淘汰された異体字である。

一般の使用法に従った場合：

冰(氷) 吊(弔) 泛(汎・汜)  
暖(煖・暎・煥)

簡便な形に従った場合：

厨(廚) 岳(嶽) 決(決)

この二つの場合が重なったときには、原則上一般の使用法を優先する。例えば“輝(輝)”。部分の構成では、書写においての書きやすさによって、左右配列の文字を選び、上下構成の漢字は除く。例えば、“略(畧)”“協(脅)”“群(羣)”。

同じ文字のなかでも筆画によっては、印刷と手書きの間に、また印刷の各種字体の間に違いがあり、異体字となるものがある。1965年1月30日文化省と中国文字改革委員会が共同で公布した《印刷通用漢字字形表》では、6196の漢字の標準印刷体を決めている。筆画の数、筆画の形、筆画の順序及び構成方法に対して説明を行っており、漢字印刷の規範を建て、また印刷体と手書き体との基本的な一致もみたのだ。例えば：

令(令) 直(直) 者(者) 普(普) 平(平)  
吳(吳) 研(研) 呂(呂) 着(着) 修(修)  
这(这) 没(沒) 益(益) 既(既) 牡(牡)  
耕(耕) 教(教) 林(林)

以上の作業によって、通用漢字の規範字体は既に基本的には確定している。現在行われている字形は長期にわたって変わることはないだろう。私たちは標準字体をつかって書くというよい習慣を養い、廃された異体や異形の文字を使わないようにすべきで、ましてや嘘字や、別字などは用ないようにしなければならない。当然ながら、漢字の形体の決定が、書法芸術と問題を起すこともない。書法にはそれ自身の芸術形式があり、標準字形の書法のほか、さらに各種の字形及び独自の書写スタイルがあってよいからだ。

このほかに、漢字の筆画の名称、筆画順、要素の数量と字形のモデルなど各方面で、十分な研究に基づいて統一された基礎の上に国家標準が打ち立てられることにな

\*<sup>19</sup> 古今字：同じ文字と認められながら、旧来の形と現在の形が異なるもの。

\*<sup>20</sup> 繁簡字：同じ文字と認められながら、形体や筆画数に差があるもの。

\*<sup>21</sup> 声旁・形旁：声旁とは、漢字を構成する各要素の内発音を示す役割の部分。たとえば「エイ」と読む「詠」は「永」が声旁となる。これに対し形旁は、意味に関わる部分。たとえば「擘」の場合は「言」となる、いわゆる偏にあたる。

\*<sup>22</sup> 造字法：漢字の造字方は通常四種類。事物を形でかたどる「象形」、上や下など抽象概念を形で示す「指事」、指事や象形でできた要素を組み合わせて作る「会意」ここでは石を積み上げて重ねる意味を示す、その中でも意味とその単語の音を組み合わせたのが「形声」と呼ばれる。たとえば「磊」は、石を組み合わせて石の多さを示す会意文字(発音は示されない)、一方「礫」は石で関係する意味を示し「累」でその語の発音を示す形声文字となる。

\*<sup>23</sup> 仮借字：ここでは、同音文字による当て字用法のこと。「念」は本来「覚えている」の意だが、同音で「朗誦する」意味の語の当て字として使われた。その後「口」偏がつき「唸(朗誦する)」ができそれぞれの意味に対する漢字となる。

\*<sup>24</sup> 同音取り替えの方法で漢字を簡化：たとえば、1956年の簡化字法案の漢字簡化表では、「復」という簡化字の中に同音の「復・複・覆」を統合してしまった(その後「覆」は分離独立)。その結果、現在ではかかる文字類ではその一つの字の意味解説に複数の柱が立つことになっている。現在簡化字を利用した中国古典の原文表記の問題点としてしばしば指摘される原因ともなっている。

る。以下は統一された楷書の印刷体とその文字サイズポイントである\*25。

	一号	二号	三号	四号	小四号	五号	小五号	六号
黒体	学	学	学	学	学	学	学	学
老宋	学	学	学	学	学	学	学	学
正楷	学	学	学	学	学	学	学	学
宋体	学	学	学	学	学	学	学	学

1965年1月公布された《印刷通用漢字表》では“横（よこぼう）、竖（たてぼう）、撇（はらいぼう）、点（てん）、折（おれ）”の5種の基本筆画を決めている。以下に挙げたのが筆画表で、表の各画筆類の後ろがその名称と文字例である。

漢字に中国で国家標準の基礎ができあがれば、つぎには漢字文化圏の他の各地区と協調して、漢字情報の交流の面で互いに調和のとれた標準を打ち立てる必要が出てくる。如何にして繁体字と簡体字、また日本の簡体字と中国の簡体字の字形の問題を処理するか、これらは議論・検討すべき課題であって、漢字のコンピュータ処理の国際化に役に立つだろう\*26。

横 类	一	横	三	彳	提(挑)	汉			
竖 类	丨	竖	卅	丿	竖 钩	了			
撇 类	丿	撇	老	一	平 撇	手	丿	竖 撇	用
点类	丶	点	衣	丶	左 点	办	丶	长 点	示
	㇇	平 捺	之	㇇	捺	义			
折 类	一	横 钩	买	㇇	横折折折	凸	㇇	斜 钩	伐
	㇇	横 撇	水	㇇	横折折折钩	乃	㇇	卧 钩	心
	㇇	横 折	录	了	横撇弯钩	队	㇇	竖折撇	专
	㇇	横折堤	诅	㇇	竖 提	民	㇇	竖折折	鼎
	㇇	横折钩	刀	㇇	竖 折	匡	㇇	竖折折钩	与
	乙	横折弯钩	几	㇇	竖 弯	西	㇇	撇 折	么
	㇇	横折折	凹	㇇	竖弯钩	已	㇇	撇 点	女
	㇇	横折折撇	及	㇇	弯 钩	家			

\*25「文字サイズポイント」：原文は「字号」。中国の印刷で用いられる文字サイズとその番号のこと。以下の表では、1号～6号まで示されるが、ここでは印刷の都合上字体は正確な大きさではない。

\*26：以下の訳文にも登場するように、現在しばしば利用されている国際的な漢字コードにアメリカの企業の共同提案により作成され、国際標準化機構のISO10646の中に組み入れられたunicodという文字体系がある。この体系は中国や日本・韓国などの漢字の事情には詳しくなく、英語中心の視点であることが指摘されている。

### 3 音を定める

音を定めるとは、それぞれの漢字の規範化された標準音を決めることである。

普通話在北京語音を標準音として決めている。つとに20年代には国語音を定める仕事が既に始まっており、現在通行している漢字音のほとんどがその時既に審議されて決まっていた。しかし、実際の使用では、依然として一文字に複数の発音があるという状況が存在していた。例えば、“波”の字は“pō”と読むときもあれば、“bō”と読むときもあった。“帆”の字は“fán”と読めば“fān”と読むときもあった。“指”の字は“zhǐ·zhí·zhǐ”等の幾種類かの発音があったのである。複数の発音を持つ漢字は、一字一字検討して決めていかねばならない。その場合、各種の状況に基づき、漢語語音の発展法則を原則とするのだ。多くの人々に普通話が学びやすい、という視点から、一般の使用状況に従い、現実状況を受け入れるという態度に基づくので、概ね以下の幾種かの方法で整理を加える事になる。

#### (1) 必要な異音は保存する。

例えば“畜”の字を動詞(蓄積する)の意味に用いるときは“xù”、例えば“畜産、畜牧”。名詞の意味(禽獸・家畜)に使うときはchù、例えば“家畜”“畜力”。また、例えば“血”を文語音として読むときは“xuè”、例えば“貧血”“嘔心沥血”。口語音で読むときはxiě、たとえば“鸡血”“流血”。さらに、語素を異にする場合、発音は異にする、“学校”の“校”は“xiào”で、“校正”の“校”は、“jiào”と発音するのだ。地名や姓名、民族名などにも異音を残しておく。

#### (2) 文字によっては複数の発音方法<sup>\*27</sup>をやめる。

例えば“教室”の“室”には“shì”と“shǐ”の二通りの発音があったが、今では“shì”一つを使う事に決まっている。“比較”の“较”には“jiào”と“jiǎo”の二通りの発音があったが、今では“jiào”一つを使うことに決まっている。

#### (3) 使用の現状に従うという原則で発音を改める。

例えば、ある古音を現状の発音に変えてしまうも

の、“叶公好龙”の“叶”は本来“shè”と発音したが“树叶”の“叶”(yè)にしたがい同様に“yè”と発音する。また、“往”の字はももとは、“往前看”のように前置詞として使うときは“wàng”と発音し、その他の場合、“来往”“往事”などは“wǎng”と発音したが、現在では“wǎng”の一つにまとめられている。

#### (4) 意味が同じ或いは近似する字の複数発音を減らす。

例えば本来は“寻思”の“寻”は“xín”と発音し、“寻找”の“寻”は“xún”と発音したが、現在は“xún”の一つで発音する。また、そもそも“指甲”の“指”は“zhǐ”と発音し、“指头”の“指”、“食指”の“指点”の“指”は“zhǐ”と発音したが、今では“zhǐ”の発音に統一されている。

漢字は一つでも意味がその発音によって変わる場合、いかに発音をしわけるか、それは漢語学習上とても難しい所である。《新華字典》<sup>\*28</sup>に収録する8000余りの漢字をみると、複数の発音を持つ多音字が800程度を占め、約一割の比率である。

普通話発音審議委員会では1957年から1962年の間に3回に分けて《普通話異読詞審音表初稿》を発表し、更に1963年にはそれらをまとめて《普通話異読詞三次審音総表初稿》とした。1985年12月、修訂を経た《普通話異読詞審音表》を国家言語文字工作委员会、国家教育委員会とラジオ・テレビ省が共同で、発表し、現代漢字の発音の規範が一層はっきり示されるようになってきている<sup>\*29</sup>。

### 4 順序を定める

順序を定めるとは、現代漢字の配列の順序を決め、標準の漢字検索方法(以後検字法と訳す)を定めることである。

漢字の配列順を決めて、人々が使いやすくすることは、大きな価値がある。それぞれの字典・辞典の編輯や、目録・索引の編輯、書籍資料の整理検索では、統一された規範順序によることが望ましい。現在のところ、常用字の調べ方には、部首法、発音法、筆画法、番号法の4種類がある。

\*27 複数の発音方法、原文は「異読」、別の発音でその漢字を読むこと。

\*28 《新華字典》: 1958年に初めて出版された小型字典。現在では字典として代表的な位置づけを得ている。現在は中国社会科学院言語研究所による編集で10版まで進む。日本では第10版の日本語版が出ている(宮田一郎編 光生館 2005・6)。

\*29 国家言語文字工作委员会/国家教育委員会/広播電視部關於《普通話異読詞審音表》的通知(1985/12/27)のこと。この通知の中では、1963年に出された《普通話異読詞三次審音総表初稿》の修訂について述べられ、修訂の原則として、“取約定俗成、承認现实的态度(既に通用しているものを取り、現実状況を承認する態度)”が示されている。これによって、たとえば従来“往”にあった2種の発音wǎng・wàng は wǎng に統一され、迹の発音が“jī”から“jì”に変わるようになった。参考資料5参照。

(1) 部首法

部首法とは、中国では最も基本的な漢字の検索法である。この方法は漢字の形体の作られ方に基づいており、簡単に習熟して使う事ができる。《辞源》\*30《辞海》\*31は部首法で項目を並べており、《新華字典》《現代漢語詞典》等の発音配列の詞典・字典でもだいたいものが部首検字法を附録している。

現在の各種字書では、部首の建て方が一致していない。例えば《新華字典》には189の部首があり、新《辞海》には250の部首がある。部首の建て方の原則も十分な統一がなされていないので、部首の分け方や部首の建て方を統一し、同じ部内で漢字の順番を一致させるべきである。その他、探しづらい漢字の統一された検字表の作成も必要である。

中国文字改革委員会、国家出版局は、1983年に《漢字統一部首表（草案）》を制定し、部首201種を建てた。以下に紹介しておこう（筆画数と第一筆の形で配列し、括弧内は繁体字或いは変形部首である）\*32。

(2) 音序法

音序法とは主に漢字音を順序とする検字法である。昔の《広韻》・《集韻》\*33は四声及び韻部に基づいて配列した字書であった。《現代漢語詞典》・《新華字典》等は共に漢語拼音方案字母表の字母順序を規則として配列したものだ。字音を知っているという前提であれば、探し方は便利で、字形の影響を受けない。同音の字が多いので、同音の漢字を配列する時は、やはり配列法を考える必要がある。

- 一画: 一、丨、丿、丶、乙(冫、フ、乚、丿)
- 二画: 十、厂(厶)、匸、卜(㇇)、冂(冂)、八(㇇)、人(人、亻)、勹、匕、儿、几(凵)、冫、彡、冫、凵、冂(冂)、刀(刂、刂)、力、勹、又、廾
- 三画: 干、工、土(土)、井、大、尢(兀、允)、寸、弋、口、冂、巾、山、彳、彡、夕、女、彡(彡)、广、冂、冫、彡(彡、彡)、尸、己、弓、中(中)、女、飞(飛)、小(㇇)、子、马(馬)、彡、彡
- 四画: 王(玉)、无(无)、韦(韋)、木、支、犬(犴)、歹(歹)、车(車)、戈、比、牙、瓦、止、支、(女)、日(日、日)、贝(貝)、见(見)、牛、手(手、扌)、毛、气、长(長、長)、片、斤、爪(爪)、父、月(月)、氏、欠、风(風)、爻、文、方、火(火)、斗、户、心(心、心)、母(母)、水(水、水)
- 五画: 示(示)、甘、石、龙(龍)、业、目、田、四、皿、生、矢、禾、白、瓜、鸟(鳥)、疒、立、穴、疋(疋)、皮、彡、彡
- 六画: 耒、耳、老(老)、臣、冫(冫、西)、而、页(頁)、至、虍(虎)、虫、肉、缶、舌、竹(竹)、白、自、血、舟、色、齐(齊)、衣(衣)、羊(羊、羊)、米、聿(聿、聿)、艮、艸(艸)、羽、彡(彡)
- 七画: 麦(麥)、走、赤、豆、酉、辰、豕、鹵(鹵)、里、足(足)、邑(邑右)、身、彡(彡)、采、谷、豸、龟(龜)、角、言(言)、辛
- 八画: 青、卓、雨、齿(齒)、非、黽(黽)、隹、阜(阜左)、金(金)、鱼(魚)、隶
- 九画: 革、面、韭、骨、香、鬼、食(食)、音、首
- 十画: 鬲、髟、門、高
- 十一画: 黄、麻、鹿
- 十二画: 鼎、黑、黍
- 十三画: 鼓、鼠
- 十四画: 鼻
- 十五画: 龠

\*30《辞源》：商務院書館の編纂。1908年（光緒34年）に編纂開始、1915年に出版、1931年に続編が出版、1939年に合訂版が出版。この版は現在でも利用されている。後1958年以降古典語彙の収録に特化して修訂が行われ、1979年より4冊本として出版される。古典籍用のため繁体字を使用、評価が高い。部首は康熙字典の214部首順による。

\*31《辞海》：中華書局の編纂、1937年の出版。1958年より修訂を加え、現在では簡体字を使用し、総合的な辞典となっている。

\*32 これに対して日本の現在の漢和辞典は概ね康熙字典の214部首に従う。

\*33《広韻》：隋代にできた《切韻》が北宋の初めに増訂されたもの。詩の押韻部を206部分け、平・上・去・入の四声調に従って区分し、それぞれ同音の音をまとめて配列している。隋唐の中古音を反映しており中古音研究のための貴重な資料。

《集韻》：宋代の編纂。形式は《広韻》を踏まえるが、音注には修訂があり、宋代の語音を知る資料とされる。



注意すべきは、現行の音配列法は通常は“字”単位であって、単語を単位として音配列を行ってはいないことであり、また声調の順番も考慮する必要があることだ。例えば、“fāngdǐng 房頂”は“方便fāngbiàn”と“方法fāngfǎ”の間に置かれるのではなく“方法fāngfǎ”の後ろに並べられる事になる\*34。

### (3) 筆画法

筆画法とは画数と書き順を結びつけた検字法である。まず、画数で区分し、それから書き順によってすべての漢字を配列する。現行の筆画法は一般に“札”字の五筆準配列法を採用している。つまり“一 | ノ、し”の順序である。筆画法と音序法の同じ所は、筆画法も“字”を単位として言葉を並べ、初めの文字が同じ言葉は、次の文字の筆画数や筆順で順番を決めてゆく、こうやって順次言葉を配列する所である。筆画法で漢字を配列すると、同じ分類に入る漢字がはなはだ多くなる\*35。このため、項目の少ない索引や、検字の補助として使われるくらいになっている。

### (4) 番号法\*36

番号法とは、数字によって漢字を表示する検字法である。それぞれの文字に一つの数字があり、数字の大きさに基づき、小から大へと順番に漢字を配列するものだ。常用字の番号検字法には“四角号碼(四隅符号検字法\*37)”がある。これは漢字の筆画の形体を“頭”“横”“垂”“点”“叉”“挿”“方”“角”“八”“小”の10種類に分け、それぞれが、0から9までの数字で代表される。それぞれの漢字はその四隅の形で、四つの数字に変換されるのだ。例えば“花”は、4421とされる。この検字法の欠点は、人工的に分類されてしまった感じが強い筆画も結構あり、そのため多くの漢字のデジタル化が難しく、重複してしまうものもかなり多い所であった。

このほかに、一般の辞典は二つの異なる検索法を採用し、互いの補充を計っている。

これを要するに、以上のそれぞれの漢字配列方法はそれぞれで完全というわけではなく、更に研究を進めて、標準を確定する必要があるものばかりだ。筆画数と筆順に違いがある漢字はその画数と筆順をはっきりさせなくてはならない。標準の漢字順と検字法があれば、それぞれの方面の需要を満たすことができ、現代化された用字という要求に応えることができるだろう。

### 参考

漢字を覚えることは外国人が漢語を学ぶ時の難しい所であるが、字典を引くととなると外国人には一層難しいこととなる。

漢字の古い文字検索法は、漢字に関する知識をかなり深く理解してからでないと使えないものが多かった。もし、ある字を捜すときに必要とされる知識が、その文字を見つけてからでないと分からないような知識によるとなると、これはもう自己撞着だ。外国人に漢語を教える時に必要な漢字の探し方には、以下の基本的特徴が必要なのである。

#### 1、異なる文字社会の人々でも、その探し方が簡単に理解できなければならないこと。

形・音・義の三項目の中では、字音の差が言語間では最も大きいものだし、字の意味もまた簡単に理解できるものではないのに対し、字形は文字の根本であり、どんな種類の文字でも字形は二次元でできているので、図形の一つだといってよい。文字を捜すときに利用する項目を、漢字というレベルで考えるのではなく、図形というレベルで考えてしまうと、線分の違いという方向に向かう。漢字の筆画中の“横・縦・払い・点(おさえ)(一・|・ノ・し)”はちょうど4種類の異なる線分であり、他の文字の基本的線分と一致するため、留学生も容易に区別がつく。これは、文字を捜す項目として取り上げるのにふさわしい。一方、25種類に及ぶ漢字の筆遣いの形体は、“横・縦・払い・点(おさえ)(一・|・ノ・し)”の筆遣いの形体の組み合わせに適切に分解することができる。

#### 2、漢字の探し方は漢字の字形独特の弁別的特徴を反映できるものであること。

字形の面から漢字の探し方を言うならば、それは漢字の持つ弁別的特徴をしっかりと把握したものでなければならない。例えば、“四角形”の図形：“目 唱”等、“四角い枠”が主体となっているもの：“电 旧”等、四角い枠が重要な構成要素になっているもの：“扣 照 疇”等、さらに“枠・角”の組み合わせや、縦横の線分、交点、等は、すべて漢字字形の重要な図形的特徴である。例えば“龔”は全部で36画、“鞞”の字には13の交点があるというものがそうだ。

\*34 この点日本の中国語辞典、たとえば岩波中国語辞典などがローマ字表記によって配列したものは単語重視の視点で注目される。しかし、日本で出版される中日辞典の多くは、現在拼音ローマ字順による漢字を親字とした配列が主となってしまっている。また、声調による順番とは、同音の場合声調の種類も配列順を決める要素となること。

\*35 例を挙げれば、“石”の文字は、五画の中で、最初の筆画“一”の区分の下に属することになる。この分類に属す漢字は、卡北占凸……など数十種類になる。資料6参照。

\*36 番号法：原文は「号碼法」、直訳して「番号法」、主に数字を使う。

\*37 四隅符号検字法：原文は、四角号碼検字法。1970年代中期～80年代前期に流行する。日本でも大漢和辞典が利用し、《漢語林》(大修館書店、1987)のような小型辞典でも一時期利用したことがある。“四角号碼”の名でよく知られるので、訳文では以後“四角号碼”を利用。

3、漢字の探し方は、すべての漢字形体を表現できるものであり、かつ10の要素を上限とすること。

要素の数が10に上限とされるのは、総数を制限して字を捜す難度を下げる他に、素早く打ち込める小型キーボードのキーが0から9までだからでもある。

4、その探し方は簡単で、すぐ使えるようになること。

以下に外国人に漢語を教える時便利な漢字の配列方法を紹介しよう。簡易五形検字法<sup>\*38</sup>である。

簡易五形検字法とは、全部で5つの数字を使い、その中の前の3数字が外側の形に対応する数字で、後ろの2数字が内側に対応する数字である。

外形の数字は、四角形の漢字で、左上隅、右上隅、右下隅の順に対応する。この3種の隅の外形の特徴は全部で10類にわけられ、それぞれ0～9の数字によって示される。1は、平ら形(対)、2は豎形(上)、3は点・おさえ形(站)、4は支え形(让)、5は交差(犯)、6は枠(黙)、7は折れ曲がり(斤)、8は人形(短)、9は个形(怕)、0は特殊形(草・街)である。(括弧の中の例字はともに左上の隅の型を例として挙げたもの)

内側対応の数字は、漢字の内側の特徴に注意して、それぞれ平ら形(横線に似たものを指す)と交差の数をとり、それを数字に変えるのだが、数字が10を超えると1をその数とする。例えば、“承”の文字の外側の数は173で、“轟”の外側の数字は501であるが、その内側には“承”は5の横線があり、3つの交点があるので、内部数字は53である。“轟”には18の横線があり(これを1とする)、3つの交点があるので、内側の数字は13である。よって、“承”の数値は17353、“轟”の数値は50113である。字典の本文で数字の17353,50113を探し当てれば、すぐさまそれらを探し当てられる。

簡易五数字法は、四角号碼の検字法の長所を取り入れながら、それ自身にも以下のような優れたところを持っている。

(1)、漢字の発音、部首、筆画、筆順や筆遣いを前提としないので、字形だけで探せる。例えば、

俊	89521	暨	71191	武	13441
丿	70110	儂	81617	罐	85115
轟	50113	驪	46318	靄	11614
灑	31410				

(2)、規則は簡単で、捜す時の方法も明快で簡単、速度は速く、番号の重複も少ない。例えば四角号碼で数字の重なる<sup>\*39</sup>ものも、以下のように分けられる。

而10720	丙10721	雨10721	兩10721
兩10722	爾10725	麗10730	兩10731
霽10731	霽10741	霽10741	霽10742
霽10742	霽10750	霽10763	霽10763
霽10771	霽10711		

(3)、字典は数字の小から大に向かって配列すればよく、二次検索は不要であって、これだけですぐ見つけられる。

(4)、わずか10種の筆形を覚えればよく、筆形の分析には3箇所が必要なばかりである。

この方法で編纂された辞典として既に、新世界出版社の《漢字音形速查》、北京大学出版社《漢英快捷小辞典》、台湾新文語出版社の《繁簡通》・《漢印通》(漢語、ヒンディー語)等の字典がある。

三、漢字コードについて

1. 電子計算機中国語化の要害 — 漢字入力コード

1946年世界に初めての電子計算機ENIACが出現して以来、電子計算機は思いがけないスピードで発展しつつある。80年代以来、マイクロコンピュータが速やかに社会と家庭、個人の生活に浸透し、今や各種の情報を伝える最もよい道具となっている。とりわけインターネットの普及により、そのかけがえのない作用を一層目立たせることになった。

電算機の発明が始まったときは西洋の言語、中でも英語と一緒に結びついていたので、文字入力是一般に西洋のタイプライターのキーボードを用いていた。よって、中国における電算機の普及は漢字の抵抗を受けたのである。西洋のタイプライターが発明されてこのかた、中国語のタイプライターの小型化の問題はずっと解決されていない。中国語のタイプライターの文字のタイプ速度は英語の十分の1に及ばない。しかも、タイプライターの重量は西洋のタイプライターをはるかに超えて重いから、タイプライターが各家に普及するはずがなかった。今やこの問題が電算機にも出てきた。中国語情報を電算機に入力する問題を解決するために、各種各様の漢字コード方案が生みだされたのである。

現在の電算機技術では漢字のコードに、以下の幾つかの要件が求められている。

\*38 簡易五形検字法:原文は「易五碼查字法」。易は簡易の「易」とこの方法を創出したと聞く易洪川氏(現広州大学教授)の「易」をかけるか。  
\*39 数字の重なる:以下に引用された漢字は、四角号碼ではすべて1022と記される。このようは重複を避けるために、四角号碼には補助符号が加えられるが、それでもこれらの文字の場合1022-7となって同じなため、複数の漢字が同じコードに多数並ぶことになる。

### (1) 文字コードの単義性。

拼音の検字、四角號碼の検字のやり方では、異なる漢字に同じコードが多くなる。しかし、電算機では、同じコードの文字は少ないほどよく、できれば一コード一字だ。常用の3500漢字の中で、同じコードの存在率は1%を超えてはならず、通用7000字で云えば、3%~5%を超えるべきではない。さもなければ、電算機の漢字入力には各種の混乱が起こるだろう。

### (2) 規則の一貫性と簡易性

文字を見ればコードがすぐ分かる必要がある。電報番号は単義性の点では完全に合格である。しかし、あのような羅列形式の番号振り付け方式では、字形や字音と無関係で、“無規則のコード割り当て”といえよう。漢字を見てそのコードを推定することもできないし、コードを見て漢字を知ることもしない。漢字のコード割りは文字をみればその番号がわからねばならない、つまり、拼音の検字法と同じように、コードをあれこれ考えて文字入力時に影響が及ぶものであってはならない、ということだ。

### (3) 使いやすさ。

電算機の入力はキーボードを通してなされる。現在既に文字や発音の識別技術による入力手段があるとはいえ、しかし、実用的には幾つかの問題がある。漢字の入力は一字ずつの入力方法を採用することも可能だ、つまり大きなキーボードを用意し、一つのキーに一つの文字を対応させるのだ。しかしこれをするとは巨大になりすぎて、必要な文字のキーを捜すのにとっても不便だし、入力スピードも極めて遅くなる。

曾て日本でキーボードのキーの数が異なる三種のタイプライターで入力速度の計測が行われた。計測の対象はともに日本の熟練したキーパンチャーだ。結果は以下の通り：

アルファベットタイプライター (26鍵)	450打/分
50音仮名タイプライター (50鍵)	250打/分
漢字仮名混合タイプライター (2000鍵)	50打/分

漢字のコードは、それぞれの漢字を幾つかの対応コードに分析し、しかる後に並べるものなので、それらのコードの総数がキーボードのキーの数と対応することになる。

### (4) 各対応コードの快速性

英語の「play」の語を入力するには、キーを四回叩かなければならない。しかるに漢字の“玩”の文字では、すべての漢字が配列された巨大キーボードを使えば、一度打てばよい。しかし、そのキーを捜すのには手間がかかり、費やす時間は実際のところ英文の入力時間より長くかかるだろう。よって、対応コードの総量を減らすという前提により、やはりそれぞれの対応コー

ドで打つキーの数を極力減らすべきである。しかし、その一方で異なる漢字が同じコードを持つ可能性を上げるわけには行かない。例えば、筆画法での入力は、キーを八回打たねば、一つの“玩”の字を入力できない。同様に採用しかねるのは、筆画入力法はキーの数を大いに削減することができるものの、しかしこの削減はコードの分量の長大化をその代価としているわけで、加えてコードの重複となるとなおさら問題になってくる。

### (5) 入力の確実性

英文の文字は少ないので、熟練した人なら目をつぶっても入力できるし、打ち間違い率も低く確実性も高い。一方すべての漢字の並ぶキーボードは打ちながら探し、その上打ち出したものが正しいかどうか確認もせねばならず、目をつぶって打つこともかなわず、間違いも起きやすい。

このほかに、幾つかの必要事項がある。例えば、設備が少なくすむように、構造が簡単であるように、技術的に実現しやすいように、外部コードから内部コードへの変換が簡単であるように、対応するソフトウェアの内部デスクに占める割合が減るようになっていくなどである。これらは共に漢字をコード化するときに考えておかねばならないことだ。

以上の原則の求めに基づき、長い間、各種の漢字の入力コード化方案が絶えることなく出現してきた。80年代の末になると、急激な増加を迎えその数は700種あまりにまで達していたが、しかし実際に用いられるものとなると、十余りの方法に過ぎない。しかも、最も優れた方法だと公認されたものはまだできあがっていない。漢字入力のコード化は確かに電算機の中国語化中の一大難事なのである。

## 2 漢字コードの種類

情報の特徴によって分類すれば、現在漢字のコード入力の方法は三類型に大きく分けられる。形型コード、形と音型コード及び音型コードである。

### (1) 形型コード

漢字の字形によるコード化は二種類に分けられる。筆画番号と文根(文字要素)番号である。

#### a. 字画コード

字画は漢字を作る一種の“文字”と見なすことができ、字画に対して対応する番号を決めることが可能だ。その後その字画がその漢字の中でどの場所に出るかに従い、字画の登場順が決まってくる。こうすると、それぞれの漢字のコードは、その字画がその漢字中で現れる順番に従って並べられた数字コードとなる。

参考

字画コード型の中で、比較的代表的なものには李金鎧<sup>\*40</sup>等の人を作った「筆形コード化方案<sup>\*41</sup>」である。「筆形コード化方案」は、現代漢字の字画に元来の五筆の基礎と共に、「又(十)」と「方(口)」を加え、さらに「折」を二つに分けて、左に曲がるのを「折」とし右に曲がるのを「弯」と呼ぶことで、合計八種類とし、コード化をしやすいとした。このコード化は八種類の字画を1・2・3・4・5・6・7・0の8つの数字に分類して、それぞれ、横・豎・撇・点・折・弯・又・方に対応させるものだ。

同じコードの文字がないように、字画の位置への注目は書き順にはよらない。なぜならば、書き順によると、1344は「犬」でもあるし「太」でもあり、1132は「开」でもあり「井」でもありうる。よって、字形番号では、厳格な位置順を決めている、つまり「先高後低」(或いは「先上後下」)、「先左後右」である。例えば、「犬」は、先ず「ノ・丶」に注目し、その後「一・丶」に注目して3414となる。番号が長くなるように、字形コード法はたくさんの要素を持つ漢字の場合、多くとも三つの部分に注目し、三つを超える漢字については、第一と第二及び最後の一つの部分の字画コードに注目する。例えば、「麤」は、4322536で、「43」は、「尸」の「丶・ノ」に、「225」は「皿」の「丨丨→」に、「36」は「几」は「ノ・乙」に対応するコードである。このほかに、最もよく使う常用字を簡略番号とする。例えば、「的」の簡略番号は3(「丶」に注目する)で、「是」の簡略番号は2(「丨」に注目する)、「主」の簡略番号は4(「丶」に注目する)である、さらに二桁或いは三桁の簡略番号もある。

例

工121 五1251 人34 之454 弓515

女613 丰711 品000 汇4416 帅23225  
 千37 入43 子515 民51616 艾727  
 目011 打51115 堅225471 小534 犬3414  
 专5114 讯4667 田07 团0513

この方法は、確かに漢字の字画用に用いるキーが少ない、しかし、以下の問題点を伴ってもある。それはつまり字画のコード配列が書き順の習慣に背いているということだ。しかも、コードの長さが時には8桁になるとすると、一文字の入力速度がかなり落ちる。現実的に言えば、人間の指が処理できるキーは、三四十個にはなるから<sup>\*42</sup>、字画番号の入力キーの利用率は低く、人間の持つ能力の十分な利用によって入力速度を向上させるものではない。

b 字根コード (構成要素コード)

字根 (構成要素) 番号とは、漢字の構成要素を単位としてコードを決める方法である。

字根コードの中で、現在最も用いられているのが、「漢字五筆字形コード」で、考案者は王永民<sup>\*43</sup>だ。

「五筆字型」の入力法は「横・豎・払い(撇)・抑え(捺)・折れ」の5種類の字画を順番に「1・2・3・4・5」の数字に割りあて、書き順の最初の字画に当たる数字に従い入力キーを五つに大分けし、それぞれの区分の中で再び五つの位置づけに分け、区号と位号とを名付けて、二桁の数字11～55の25個の番号で表示する。以下の表がその例だ。

それから125種の主要字根を決めてコード表に加え、形が似ているもの(例えば：王・主・五)、淵源が同じもの(例えば：耳・卩)の字根を25個のキーに振り当てる。こうやってできあがった「五筆字型」の字根表

区位 位号	1	2	3	4	5
横 1	王一	土二	大三	木	工
豎 2	目丨	日𠃉	口川	田	山
撇(はらいぼう) 3	禾丿	白リ	月彡	人	金
捺(右はらい) 4	言讠	立彳	水彡	火	之
折(おれ) 5	巳乙	子㇇	女㇇	又	纟

\*40 李金鎧：未詳

\*41 字形コード化方案：原文は「筆形編碼方案」

\*42 ここでは、キーボードに利用されるキーの数を言う。数字1～0、及びアルファベット、及び補助記号を合わせれば40程度のキーを通常は使えることになる。

\*43 王永民：1943年生、1978-83年にかけて五筆入力法を考案し、普及に努め、現在中国内では90%以上の使用率になるという。現在北京王碼公司總裁(百度百科)。

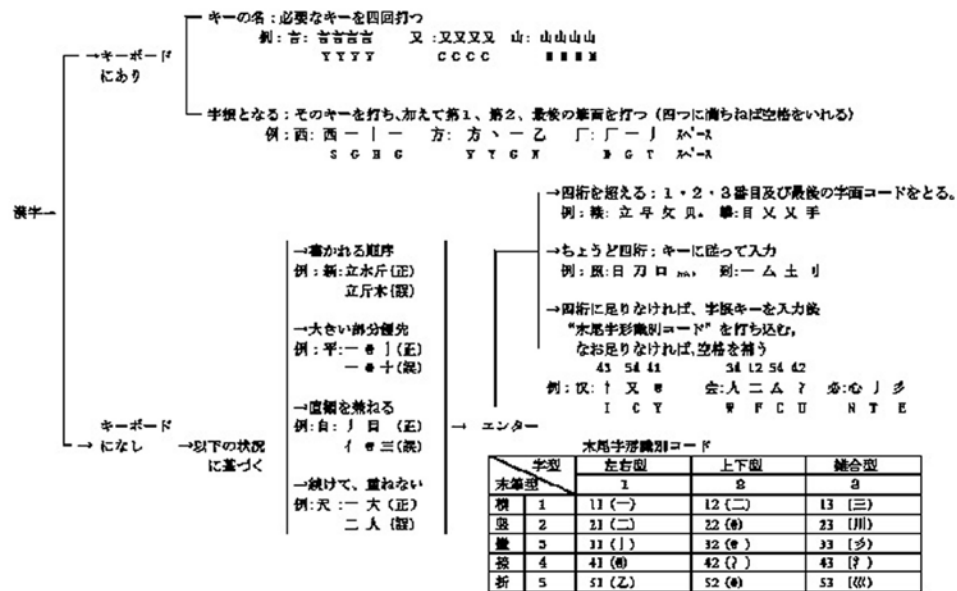
のキーボードは以下のようにになる。

金 斤 夕 35Q	人 八 34W	月 夕 33E	白 斤 32R	禾 斤 31T	言 斤 41Y	立 斤 42U	水 斤 43I	火 斤 44O	之 斤 45P
工 夕 15A	木 丁 14S	大 夕 13D	土 十 12F	王 五 11C	目 止 21H	日 早 22J	口 川 23K	田 力 24L	:
Z	至 夕 55X	又 夕 54C	女 夕 53V	子 夕 52B	己 夕 51N	山 夕 25M	<	>	?

これに“書き順・筆画多数を優先・直感を兼ねる・繋げるが重ねない<sup>\*44</sup>”の四原則に従って、漢字を順番に四つの字根に分け、対応する字根のキーナンバー(英文ローマ字)を打ち込んでいくと、通常は四つのキーでそれぞれの漢字が打ち込める。字根が少ない一部の文字には、番号の情報量を増やすために、必要な識別記号を加えてやる。この識別番号は、字型のコード(左右構造を1、上下構造を2、雑合構造を3とする)と書き順の最後の字型を組み合わせて作る。以下の図が五筆字型コード化の進行図である。

参考

漢字要素コード型の中で、もう一つ影響が比較的大きいものは“天竜蒼頡番号”であり、朱邦復<sup>\*45</sup>の考案による。このコード化方案では、それぞれの漢字を字首と字身の二つに分けており、すべての漢字は基本的に300の字首と2000の字身に分けられる。蒼頡番号は24個の漢字を選んで中文字母とし、この24の中文字母ですべての漢字の基本単位としていて、それぞれの字母は5個を超えない。例えば、“鋸”の文字のコードは“金 尸 + 口”でキー上にある“金”、“尸”、“+”、“口”の四つのキーを推せば、“鋸”の字が入力できる。蒼頡番



\*44: その漢字の分析において、以下の順で各キーボードに割り振られた要素を探すというもの  
書き順: “拆”は“扌 斤 \”の順で打ち込み、“斥 扌”の順ではない。  
筆画多数を優先: キーボードにある要素のなかでは画数の多いものを選ぶ。“豪”は“宀 冫 豕”で打ち込み、“宀 口 冫 豕”の順では打ち込まない。  
直観を兼ねる: 見た感じのまともまり大切にすること。“魚”は“魚 一”で打ち込み“夕 田一”で打ち込まない。  
繋げるが重ねない: “天”は“- 大”で打ち込み“二 人”では打ち込まない。“一”と“大”はつながっているが、“二”と“人”は重なっている。  
\*45 朱邦復: (1937 ~) 現在香港上市公司文化伝信集団の副主席。蒼頡入力法を發明し、中国語コンピュータに大きな貢献をしたところから、香港台湾では「中国語コンピュータの父」と称される (百度百科)。

号の中文文字分類は4類であり、60の補助字型が配分される。それは右の通り。

漢字を入力するときは、まず漢字（文字キーに既にある漢字は除く。例えば“日・A”“月・B”“金・C”“木・D”“水・E”“火・F”“土・G”“竹・H”“戈・I”“十・J”“大・K”“中・L”“一・M”“弓・N”等々）を

哲理类	笔划类	人体类	字形类
日 A 四	竹 H (竹) 卩	人 O 人 亻	尸 S (尸) 冫
月 B 冂 冃	戈 I (点) 彡	心 P 忄 小	女 T (女) 艹
金 C 钅 釒	十 J (十) 山	手 Q 扌 扌	山 U (山) 卩
木 D 扌	大 K (大) 彡	口 R	女 V (女) 艹
水 E 氵 冫	中 L (中) 冫	X 彡 彡	田 W (方) 口
火 F 灬 火	一 M (横) 冫	Z 彡 彡	ト Y 彡
土 G 土	弓 N (弓) 冫		

字首と字身に分解し、外側から内側へ、上から下へ、左から右の順番にキーを打つ。例えば“甘”のコードは“甘、一”となる。また“商”の番号は“ト・金・月・口”が選ばれる。字首の番号は二桁に限られ、字身の番号は3桁に限られる。よってそれぞれの字のコードの長さは5桁を超えない。例えば、“輯”の字首“车”は、規則によれば“十・一・十”で、二桁を超えるので、最初と最後の番号をとって、“十・十”とする。字身の番号も同様だ。例えば、字身自体がまた更に次のレベルの字首と字身に分けられるときは、下位のレベルの字首の始めと終わりの番号を取り、下位のレベルの字身の始めと終わりの番号をとる。例えば“溯”の字は、“水・甘・山・月”をコードとして“鋸”は、“金・尸・口”をコードとするが、その中でも“尸”は下位のレベルの字首となる。

(2) 形音コード

形音コードは漢字の発音と漢字の部首・字形・字義を組み合わせて作るコード化方式である。例えば、文字の声母、韻母<sup>\*46</sup>の基礎の上に部首の番号或いは字画の番号を組み合わせ、一つの文字の代用番号とする、或いは形についた字元（漢字の構成要素）の音番号に字画の音番号と漢字本体の発音の声母とを組み合わせ

て一字の記号とする。

この種類のコード化で、現在比較的多く使われているのが、“見字識碼” 方案である。“見字識碼” コード化方案は支秉彝<sup>\*47</sup>の発明による。

このコード化の基本原則は、字音・字形の双重関係に基づきコードを作り上げるというものだ。先ず漢字を構成部分（字元）に解析し、極力四つの字元で作られる字の並びに分ける。例えば“路”なら 口・止・文・口に。この四つの字元の拼音字母のイニシャルに基づき、その字に対応するコードとする。つまり“路”ならば、K Z W K<sup>\*48</sup>となる。もし、三つの字元に分けられるだけだったら、第四番目の字母は、その字の最後の筆画の漢語音のイニシャルをとる、例えば、“余”は人、于、八、丶で、R Y B D<sup>\*49</sup>となるが、最後の“D”は漢語の“点”の発音diǎn イニシャルである。もし、二つの字元にしか解析できないときは、第三番目の字母に最後の一筆の拼音のイニシャルを番号化する他に、第四番目の字母としてその字の発音の拼音イニシャルを使う。例えば“吳”は、口・天・丿・吳で、K T N Wとなる。Nは“捺(nà)”、“W”は“吳”の“wú”のイニシャルである。このコード化は、字元の分析を字型の基礎に置いているのだが、字元が文字化されると、今度はその字音が基礎となってくる。よって音と形の結合となるものなのだ。実際の使用において、漢字を見ればそのコードが分かるし、重複率も低く、拼音もそれほどしっかり知らなくてもよい。ラテン文字を採用しているので、欧文キーボードを使って打ち込むこともでき、ブラインドタッチも可能だし、入力速度も速い。

とはいえ、字元の解析を極力四つにしようとするために、まったく分析のできない文字、例えば“一”“乙”等は、それぞれI I I I、H H H Hで示されることになる。しかしこれはたいした問題ではない。五つや六・七或いはそれ以上に分けられる文字となると、字元の確定に根拠が乏しくなってしまう問題がおきる。例えば、“彰”の場合は“立・日・十・彡”に分けられるのに、“章”だと“一・立、日、十”となってしまう、漢字の要素のもつルールに合致しない。こうなると学習や記憶の難度が上がり、“非理性的”な感じがしてくる。実際、“見字識碼”の評価はもはや高くない。

\*46 声母、韻母：中国語の発音は、音節が単位となる。音節構造は語頭子音・介音・主母音・尾音とならび、これに声調が被さることになる。語頭子音を声母と呼び、介音・主母音・尾音をまとめて韻母と呼ぶ。  
 \*47 支秉彝（1911-1993）1976年より漢字のコード化の研究を始め見字識碼を制作する。1980年中国社会科学院院士。中国で漢字コンピュータ処理に取り組んだ先駆者として知られる（百度百科）。  
 \*48 K Z W K：各漢字の拼音ローマ字を（ ）中に記せば、口(kǒu) 止(zhǐ) 文(wén) 口(kǒu) となる。K Z W Kはそれぞれのイニシャルをとって組み合わせたもの。  
 \*49 人、于、八、丶で、R Y B D：人(rén)、于(yú)、八(bā)、丶(diǎn)。以下同様

### (3) 発音コード

発音コードは、漢字の拼音<sup>\*50</sup> (ピンイン) を使って入力する方法だ。台湾では“注音字母”<sup>\*51</sup>による入力方法となる。

実際のところ、漢語拼音による入力は“漢字コード化”の一種とは見なせない。なぜならば、その表記は別にコード化する必要はないからだ。それは、普通話 は語音と完全に一致しているから、漢語の拼音さえ分かれば、直接キーボードを操作できる。それを一種のコード化と見るのは、それには“縮ピン”、“双ピン”等の学びうる形式があり、更に重複番号の選定や、単語区分の処理などの面で、その他のコード化の方法と似た所があるからだ。

多くの発音コードは声調を区別せず、漢語拼音が用いるローマ字に基づき、そのまま声母と韻母を入力する。このような入力方法を“全ピン式”と呼ぶ。“全ピン式”の有利なのは、漢語拼音と同じローマ字そのままにキーを打てばよいところだ。これは英文タイプライターに習熟した使用者にとっては、非常に楽に入力できる。しかし、漢字の中には何度もキーを打たねばならない漢字があるのが欠点だ。例えば“状”“zhuang”は、キーボードで六つのキーを打つ必要がある。

その後、二つ三つの文字で示される子音や母音を一セットにして、使用に混乱が起きない文字に割り当てる方法も出た。例えば：

拼音	ang	eng	ong	ing	an	ao	en	ai	zh	ch	sh	ü
キー	h	g	s	y	j	k	f	l	a	i	u	v

このようにすればそれぞれの漢字の入力コードは三つの字母以内に圧縮することができる。例えば“状 zhuang”は“AUH”<sup>\*52</sup>である。

“全ピン式”以外に、“双ピン式”の使用もずいぶん

Q	W	ia	E	R	uan	T	Y	U	I	O	P	un
iou		ua			üan	üe	ü	sh	ch	uo		ün
A	S	ong	D	iang	F	G	H	J	K	L		;
		iong		uang	en	eng	ong	an	an	ai		ing
Z	X	C	V	B	N	M						
	ei	ie	iao	zh uei	ou	in	ian					

広がっている。“双ピン”とは、即ち各漢字の声母・韻母でおのおの一つのキーを押すだけのものだ。現在比較的利用されている省略方“双ピン式”キーボードの設計は下図の通り。

“双ピン式”は初学者には記憶への負担が些か大きいけれども、キーを打つ数が少ないので、入力の時間を短くすることができる。

漢語拼音を使って漢字を入力すると、それぞれの音節にある大量の同音漢字のため、番号の重複が多くなるという問題が出てくる。同音字を選択するには再度キーを叩いて必要な文字を入力することになるが、選択の方法は番号付きの同音字を提示領域に示し、示された漢字字形にしたがって該当の数字のキーを押しその字を確定するというものだ。

一方“単語入力”は“文字入力”よりも大いに重複を減らすことができる。なぜならば、一部の単音節単語以外に、二音節単語、多音節単語を一つの単位として入力すれば、同音現象を大きく減らすことができるからだ。よって、ソフトの中には単語処理能力をいれておくべきで、その中の単語量は多ければ多いほどよい。或いは各種の学問領域で使う専用述語ごとに分類した単語カードを別に挿入利用する形式にするのもよいだろう。現在多くの表音コードによる入力ソフトは使用者が自分で単語を作ればすぐさま自動記憶を行うようにできる。

如何に容易に“単語入力”を可能にするか、これが表音コード入力型ソフトの水準を向上させるカギである。例えば現在ある方式は、一度使った言葉を自動的に一番前に持ってくるようにしているし、あるものは、声母だけを入力すればよく、また二つの文字の声母と一つの文字の韻母を入力することで二音節単語を打ち出し、三四音節の単語は声母の入力だけでよいという

\*50 拼音：《漢語拼音方案》に基づく表音文字体系の略称。1958年に公布、漢字の発音を、子音・介音・主母音・声調の5種の部分に分け、それぞれをローマ字と記号を使って示す方式。幅広く受容され、現在では、漢語ローマ字表記の国際標準となっている。

\*51 注音字母：1918年国家教育部から公式に公布された中国では初めての法定表音文字方式。1920年以降小学校で使われ始める。大陸では1958年まで使用され、その後拼音ローマ字方式に変わるが、台湾では使用を続けている。子音・介音・主母音+韻尾・声調の4区分で発音を示し、主母音と韻尾を一つで示すので、要素が拼音ローマ字方式より一つ少なくなっている。

\*52 “AUH”：A (zh) + u + H (ang) = zh u ang .

ものもある。またあるものは拼音に変換された言葉を直接使用者の文中に示し、キーを押して必要な言葉に変えるというものもある。こうすれば、漢字の提示枠で漢字を選ぶ必要は減るか不要になる。

このほかに、さらに音形コード入力というものもある。漢字の音と形の両方の情報特徴を利用して漢字を入力するもので、発音を主とし、字形を補助とする。発音コード入力と比較すると、重複率が低く同音異義の言葉の区別が比較的簡単で、発音できない文字も入力できる所で優れるが、その習得が中々難しいところが欠点だ。

これを要するに、発音表記型の種類はあまり多くはなく、しかも漢語音節の性質や原理は既に明快になっているので、この発音表記は共通の基礎をもつことになる。差異も少なく統一も簡単だ。しかし、漢語の拼音を知らない人には拼音表記を利用するものは使えない。同音意義字が多いから、文字の提示枠の中から必要な文字を選び出すにはかなり面倒で、ブラインドタッチも結構難しいものだ。

現在にいたるまでに、漢字のキーボード入力方法は三段階を経て発展してきた。つまり文字処理・単語処理・文処理である。文字処理のレベルでは主に漢字の集積、漢字変換コード、漢字コード・キーボード設定などの面での研究が進んでいる。単語処理のレベルでは主に単語のコード化、単語瀬率統計・単語品詞の自動区分、常用字の集積などの面で研究が進んだ。文の処理のレベルで進んでいるのは、単語の組み合わせ、言葉の意味、言葉の使い方の面でのルール及び句型の統計などの面での研究である。単語処理能力があるかどうか、及び単語処理能力の高低が、そのコード化の能力を測る重要な標識であるが、文処理能力をもった入力法となると極めて高い人工知能を必要とする。コンピュータに言葉と文中での同音異義語を識別させ、自動的に文字の組み合わせを調整するソフトが既に運用が始まっている。このようなソフトは高性能の知能が必要だ。使用者は一文をまるまる漢語拼音で入力すると、ソフトのほうで自動調整をして正確な漢字に変換する、そして使用者はわずかの単語等を調整すればよいようになる。

もちろんいかなるコードでも、使う者の需要にもとづいて選択される。しかし、情報を交換する時には、交換用のコードは入力時に用いたコードではなく、国家標準と国際標準に従わねばならない。つまり、“自然言語”と同様に、各種の入力コードは“方言”であり、

交流の時には“共通語”－国際語を用いる必要があるわけだ。

1980年、中国では《情報交換に用いる漢字コード集・基本》<sup>\*53</sup>番号GB2312-80を示した。その中の“GB”とは“国标（国際標準）”の二字の拼音の<sup>\*54</sup>イニシャルであり、2312-80は、国家標準の順番と年代である。この基本集積集には、一般番号が202字、序数記号60字、数字22字、ローマ字52字、日本仮名169字、ギリシア文字48字、ロシア文字66字、漢語拼音番号26字、漢語注音字母37字、漢字6763字が含まれている。その中で漢字は二つのレベルに分けられ、第一級の漢字は3755字、第二級の漢字は3008字。合わせて7445の番号で、習慣上簡単に“国際コード<sup>\*55</sup>”と呼んでいる。

“国際コード”は区分けの番号で、漢字と番号を94行94列の四角い枠の中に並べたものだ。こうやってそれぞれの漢字はその位置の行と列の記号でただ一つが表示される。“区”とは“行”にあたり、“位”は“列”に相当する。それぞれの区の中には94字があり、発音順或いは部首順に番号が決まる。例えば“工”の字は、25区04位で、その文字の番号は“2504”となる。このような標準番号は、それを直接利用して漢字入力をしようとするもので、これを必要とする者は電報を打つような専門の人材に違いない。よって、一般にはこれは内部番号となり、入力の際のコードはそれぞれの簡略なコード化の方法を使い、その後内部でこれに変換するのである。

基本集積集の補充として、1987年に国家標準局が公布した《情報交換に用いる漢字コード化文字番号集》第二補助集（GB7589-1987）と第四補助集（GB7590-1987）には、よく使われる規範簡化文字を収録、それぞれ7237字、7039字を収録し、部首によって配列している。1990年にはまた第一補助集（GB12345-1990）が公布された。ここには基本集積集に対応する繁体字を集めたもので、図形や文字番号を合わせて7583字が収められている。その中にはGB2312の中に収められたすべての文字記号の他に、6866の繁体字が追加収録されている。このほかに第三第五補助集があるが、これはそれぞれ第二第四補助集に一つ一つ対応する繁体字の文字記号集である。

1993年国際標準化組織（ISO）が《通用多八位字符集》（UCS; Universal Multiple-Octed Coded Character Set）を発表した。この国際標準第一部分のコードがISO10646.1である<sup>\*56</sup>。ISO10646.1では各種の文字のすべての字符部分を128組に分ける。各組では更に続けて

\*53 原文は：《信息交換用汉字编码字符集》参考資料6 参照

\*54 イニシャル：“国标”を拼音ローマ字で書けば、guóbiāo、この中のgとbをとって大文字にしたものがGB。

\*55 国際符号：原文は国際碼。

\*56 現在いわゆるunicodeのこと



256の表にわけ、その表の横と縦を行と位に分ける。そこに区分けされた20902区画のコードをつかって、中国・日本・韓国の各種の漢字を統一して表示するのである。ISO10646.1に対応するために、中国では国家標準GB13000.1-1993を作った。更に多くの技術的問題を処理するため、その橋渡しとして、さらに“漢字拡大内部コード規範”(略称GBK)を作成している。

2000年、今度は《情報交換用漢字コード字符集基本集の拡充》(GB18030-2000)が発表された。これはGB2321-1980とGB13000-1993の後に続く最も重要な漢字コードの標準であり、中国の電算機系統が必ず遵守すべき基礎的性格をもつ標準の一つである。この標準はGB2312-1980、GBKと、内容が重なり、27484字の漢字を収録して、区画はすべて150万区画を超え、人名や地名の用字問題を解決するための方法を提供し、漢字研究や古籍研究にとっては統一された情報の基礎的環境を提供するものでもある。

忘れてはならないのは、ユニコード(unicode)標準である。これは、アップルとマイクロソフト等の企業が電算機の多種文字文処理の問題を最終的に解決しようとして、その制定を訴えた通用多種文字字符集である。その漢字集の処理方法などは、ISO10646.1に吸収されている。

### 3 漢字コードの前途

漢字コードの中では、漢語拼音の発音コードの利用が最も広く行われ、90年代の上海の統計では、コンピュータ資格受験者のなかで、漢語拼音を用いて入力した者は95%以上を占めた。“五筆字型”コードも既に大多数のコンピュータ操作系統で使用者のために使われる入力法の一つとなっている。この方法では一分間に224字の漢字を打ち込んだという記録があり、これは英文を打ち込む速度よりもっと早い。しかしながら、漢字の構成の法則が十分に理解されていないために、各種のコード化方案は依然として一定の随意性もち、自然さに欠けるところがある。これらのコードの統一のためには、漢字の標準化の過程における構成要素の数や構造などの確定を待ち、また他面では国家の各種方案への全面的な評価というものも待たれる。

漢字のコード化の発展のためには、あちらこちらで用いられる各種コードが統一に向かう事も必要だ。配列の順序及び字典検索使用の可能性も極力考えねばならない。

科学技術の発展にともない、文字入力の面では、新しい動きが出てきた。文字の識別と発音の識別により、コンピュータに文字・発音を識別させ、キーボードは

使わず、直接言語情報を入力するというものだ。今日、漢字を書いて入力するコンピュータソフトと、音声で直接入力するコンピュータソフトはもうすでに作られている。

### 参考

#### 漢字改革に関する異なる意見と各種の考え<sup>\*57</sup>

1. 漢字の簡化の問題に於いて、ある意見は、現行の漢字にはまだ多くの文字に、例えば、“繁”“餐”“醒”等のように筆画が繁雑なものがあるので、世間でよく使われている簡体字の利用を継続推進するべきだという意見がある。その意見の中には、“同じ発音の文字に置き換える”という方法は字数を絞るというやり方で漢字の簡化を進めるもので、これは拼音化の方向にも符合し、かなりのところで採用可能だろうと考えるものもある。別の意見として、今のところ人々は現行の漢字の字形が一定期間、相対的安定性を保つことを強く願っており、漢字の簡化の任務はもはや差し迫ったものではないという考えもある。さらに情報論の視点から見れば、以下のような意見もある。つまり、使用頻度が少ないコードほど、コードの長さは長くなるものだから、ほとんど出てこないような珍しい字は、その筆画が多少多くてもかまわない、逆に認識率が高くなるし、また現在よく利用されている漢字(とりわけ常用字)は筆画上もやはこれ以上絞る余地は少なく、現代社会の中で求められているのは漢字字形の安定化なのだ、という意見だ。更にある学者は、簡体字に些か手を加え、繁体字と簡体字が一つ一つ対応するようにすべき、つまり“钟(鐘・鍾)”“发(發・髮)”といった一対複数の簡化字は、分化すべきだと提案している。

2. 漢字の拼音化の問題において、ある意見は《漢語拼音方案》の効能は漢字に発音を振り付け、識字の効率を上げ、普通話の学習を助けるところにあると考える。日常生活でもその応用範囲を拡大させ続けることができる。例えば手旗信号・灯火信号・工業製品の番号、目が不自由な人の文字、耳の不自由な人の手話に用いることができ、中国の人名や地名を翻訳するとき用いることができ、発音順検字の制作と拼音式のコンピュータのコード化などがそうだ。しかし、これらは《漢語表音方案》を使って漢字の代わりにしてしまふのと等しいものではない。ある人は更に進めて、漢語の語素語としての性格が漢字を決めたのだから、漢字は純粋な表音文字となることはできず、表音化の路を進むことはできないと考える。別の意見では、漢字は世界の文字に共通する表音化の道を進まねばならないと考える。彼らは、漢字の学びに難さ、認識し難

\*57 参考：漢字を表記文字とする中国にとっては、漢字の規範化に伴い非常に大きな問題となる。以下に並べられる各種態度は、漢字改革への各種意識を示すもの。

さ、書き難さ、入力し難さが人々の文化水準の向上と現代化に影響を与えているので、必ずや“最も進歩した”音素文字によって取り替えねばならないと考えるのだ。曾て中国の書記言語は将来必ずや拼音文字になると考え、漢字の排除を主張した人は、現在では拼音文字セットを作り漢字と共に使用するのがよいという主張の側の方へ一層傾いている。中には2種類の文字の併用は不可能だから主と従の関係が必要だ、と考えるものもある。

3. 新文字をどうやって作るかという問題においては、各種の異なる意見がある。第一の意見(多数)は、将来漢字は表音化を果たすはずだから、国家が正式に公布して人々が熟知している《漢語拼音方案》によって、漢語の表音方法を研究し、作成し、試験する基礎とすべきで、これには修正を加えてもよいが、例えば声調を文字で示すなどで他に新しい方法を作るべきではない、と考えるものである。また、自分で新しい拼音方案を作ってしまった人もいる。例えば声韻双拼<sup>\*58</sup>のもの、注音字母式三拼<sup>\*59</sup>のものだ。ローマ字の使用法では国際的な発音の習慣などを考慮する必要はないという主張があり、ローマ字など使わずに民族独特の形式のコードを使えばよいと主張する人もいる。拼音方式の選択においては、単語を一つ続けて書く純粋な表音方式を用い、幾つかの同音異義語に対しては区別をするだけでよいという人もいるし、幾万かの漢字の中から幾千かの常用字を選択し、拼音文字の“字根”とすべしと主張する人もいる(傅斯年<sup>\*60</sup>《汉语改用拼音文字的初步談》)。あるものは、漢語の特徴に基づき、新しい漢字方案は必ずや純表音のコードにコード化さ

れた意味を示す義符を加えて構成されるべきだと主張する(例えば唐蘭<sup>\*61</sup>の主張する《拼音形成文字》)。また北京口語音の基礎の上に方言と古音を利用して更に同音異義の字を分別し異なった表記形式にせよと主張する者もある(例えば趙元任<sup>\*62</sup>先生は70年代に2085字の漢字とローマ字拼音を共に利用する《通字方案》を提案した)。また、まずは押し広め或いはずっと拼音と漢字の混用文字を使い続けよと主張する者もある。常用単語は漢字を使い、非常用語は拼音を用いる、或いはこれの逆にせよというものだ。あるものは一セットのローマ字番号の文字を新しく作り、漢字と一緒に使うことを主張する。その原則とはいえば：①文字であること。声母・介音・韻母・声調・義符にはそれぞれ一つの字母をあて、その字母は26の文字の中に限定する。②語素であること。漢語の単音節語の特徴で決まるものであるため、常用の字根は当然違った形にならねばならない。③補助であること。それぞれの漢字の番号となり、義符コードの確定においても、概ね漢字がそもそももつ形旁に基づく。④実用的なもの。注音のためではなく漢字同様に実際に使われるものであるために、理論上の整合性は犠牲にしてもかまわない。また実際に使われるときに認識しやすく、弁別しやすく、印刷しやすく、時間を省けるものであるために、たとえ学習に多少時間がかかっても仕方がない。さらに、この文字セットが漢字の字母文字と共に誕生したなら、漢字のコード化の問題も同時に解決するだろうというもの。

4. 漢字は改革の必要はない、或いはもうこれ以上改革は不要だという人もいる。(了)

\*58 声韻双拼：漢字の発音を声母と韻母の二つに分けて示す方法。伝統的な反切の方法を受け継ぐ。

\*59 注音字母式三拼：注音字母は、一字の発音を最大：声母＋介音＋韻母（除く介音）の3種に分けてその発音を示す。これを三拼とよぶ。双拼に比べて介音を独立させたため分析的といえる。

\*60 傅斯年：(1896-1950) 歴史学者。中央研究院歴史言語研究所所長をへて国立台湾大学学長。

\*61 唐蘭：(1901-1979) 現代文字学、金石学者。北京大学で古文字学を教えた。中国文字学理論・古文字学で著名な業績を上げる。

\*62 趙元任：(1892-1982) アメリカ籍の華人。天津に生まれる。アメリカ留学の後1925年精華大学教授。後1938年アメリカに定住、1945年アメリカ言語学会会長。中国現代言語学の開拓者の一人。

附錄

參考資料 1 漢字簡化方案 (一部) 1956 《中國語文》 2 月号より

# 漢字簡化方案

(中華人民共和國國務院公布)

## 國務院關於公布漢字簡化方案的決議

(1956年1月28日國務院全體會議第23次會議通過)

漢字簡化方案，1955年1月由中國文字改革委員會提出草案，經全國文字學家、各省市學校的語文教師以及部隊工會的文教工作者約20萬人參加討論，提供意見，再經1955年10月全國文字改革會議通過，並由國務院漢字簡化方案審訂委員會審訂完畢。

漢字簡化方案分三部分。第一部分即漢字簡化第一表所列簡化漢字共230個，已經由大部分報紙雜誌試用，應該從1956年2月1日起在全國印刷的和書寫的文件上一律通用；除翻印古籍和有其他特殊原因的以外，原來的繁體字應該在印刷物上停止使用。第二部分即漢字簡化第二表所列簡化漢字285個和第三部分即漢字偏旁簡化表所列簡化偏旁54個，也都已經經過有關各方詳細討論，認為適當。現在為慎重起見，特先行公布試用，並責成各省市人民委員會負責邀集本省市政治協商委員會委員徵求意見，在3月底以前報告國務院，以便根據多數意見再作某些必要的修正，然後正式分批推行。

### 漢字簡化第一表

這個表里有230個簡化漢字，按注音符音序排列。括弧里邊的字是原來的繁體字。

罢 (罷)	表 (錶)	蒙 (朦濛)	达 (達)	电 (電)	台 (臺檯颱)
卜 (蔔)	边 (邊)	弥 (彌彌)	斗 (鬥)	垫 (墊)	头 (頭)
备 (備)	宾 (賓)	蔑 (蔑)	担 (擔)	独 (獨)	体 (體)
宝 (寶)	补 (補)	庙 (廟)	胆 (膽)	夺 (奪)	铁 (鐵)
报 (報)	辟 (闢)	面 (麵)	当 (當噹)	对 (對)	条 (條)
办 (辦)	朴 (樸)	范 (範)	党 (黨)	断 (斷)	听 (聽)
板 (闆)	扑 (撲)	奋 (奮)	灯 (燈)	冬 (冬)	团 (團團)
帮 (幫)	么 (麼)	丰 (豐)	敌 (敵)	东 (東)	难 (難)
别 (別)	迈 (邁)	妇 (婦)	淀 (澱)	动 (動)	拟 (擬)
标 (標)	霉 (霉)	复 (復複覆)	点 (點)	态 (態)	乐 (樂)

参考資料 2 簡化字総表1986年新版（原書附録の中の一部）：以前（1956年版）に比べて配列が拼音に変わったり、簡化の文字数が増えている他、重なる部分を以前のものと比べるだけでも、簡化に修正が加えられているのがわかる。

## 第一表

### 不作简化偏旁用的简化字

本表共收简化字 350 个，按读音的拼音字母顺序排列。本表的简化字都不得作简化偏旁使用。

A	补〔補〕	迟〔遲〕	递〔遞〕	奋〔奮〕	顾〔顧〕	
	碍〔礙〕	C	冲〔衝〕	粪〔糞〕	刮〔颯〕	
			丑〔醜〕	凤〔鳳〕	关〔關〕	
肮〔骯〕	才〔纔〕	出〔齣〕	电〔電〕	肤〔膚〕	观〔觀〕	
袄〔襖〕	蚕〔蠶〕 <sup>①</sup>	础〔礎〕	冬〔冬〕	妇〔婦〕	柜〔櫃〕	
B	灿〔燦〕	处〔處〕	斗〔鬥〕	复〔復〕	H	
	层〔層〕	触〔觸〕	独〔獨〕	〔複〕		
	坝〔壩〕	辞〔辭〕	吨〔噸〕	G		汉〔漢〕
	板〔闆〕	聪〔聰〕	夺〔奪〕			号〔號〕
	办〔辦〕	丛〔叢〕	堕〔墮〕	盖〔蓋〕		合〔閤〕
	帮〔幫〕	D	E	干〔乾〕 <sup>③</sup>		轰〔轟〕
	宝〔寶〕	忏〔懺〕		儿〔兒〕		〔幹〕
	报〔報〕	偿〔償〕	F	赶〔趕〕		胡〔鬍〕
	币〔幣〕	厂〔廠〕		个〔個〕		壶〔壺〕
	毙〔斃〕	彻〔徹〕	导〔導〕	巩〔鞏〕		沪〔滬〕
	标〔標〕	尘〔塵〕	灯〔燈〕	沟〔溝〕		护〔護〕
	表〔錶〕	衬〔襯〕	邓〔鄧〕	构〔構〕		划〔劃〕
	别〔髣〕	称〔稱〕	敌〔敵〕	购〔購〕		怀〔懷〕
	卜〔蔔〕	惩〔懲〕	余〔糴〕	谷〔穀〕		坏〔壞〕 <sup>④</sup>

① 蚕：上从天，不从天。 ② 缠：右从厘，不从厘。 ③ 乾坤、乾隆的乾读 qián(前)，不简化。 ④ 不作坯。坯是砖坯的坯，读 pī(批)，坯坯二字不可互混。

参考資料3 現代漢語常用字表 (一部)《語言文字規範手冊-增訂本》(語文出版社1991)) より

### 常用字 (2500字) 笔画顺序表

一 画						无	元	专	云	扎	艺
一 乙						木	五	支	厅	不	太
二 画						犬	区	历	尤	友	匹
二 十	丁	厂	七	卜		车	巨	牙	屯	比	互
人	入	八	九	几	儿	切	瓦	止	少	日	中
了	力	乃	刀	又		冈	贝	内	水	见	午
三 画						牛	手	毛	气	升	长
三 于	干	亏	士	工		仁	什	片	仆	化	仇
土	才	寸	下	大	丈	币	仍	仅	斤	爪	反
与	万	上	小	口	巾	介	父	从	今	凶	分
山	千	乞	川	亿	个	介	乏	公	仓	月	勿
勺	久	凡	及	夕	丸	欠	风	丹	匀	乌	凤
么	广	亡	门	义	之	勾	文	六	方	火	为
尸	弓	己	已	子	卫	斗	忆	订	计	户	认
也	女	飞	刃	习	叉	心	尺	引	丑	巴	孔
马	乡					队	办	以	允	予	劝
四 画						双	书	幻			
丰	王	井	开	夫	天						

参考資料 4 現代漢語通用字表 (一部) 《語言文字規範手冊—增訂本》(語文出版社1991) より

1 画	乂 儿 九 匕 几	土 工 才 下 寸 丈 大 兀 与 万 弋	个 么 久 勺 丸 夕 凡 及	子 孑 也 女 飞 刃 习 又 马 乡 幺	韦 云 专 丐 扎 廿 艺 木 五 支 斤 卅 不 仄 太 犬 区 历 友 歹 尢 匹 厄 车 巨	牙 屯 戈 比 互 切 瓦	气 壬 升 天 长 仁 仃 什 片 仆 仇 化 仇 帀 仍 仅 斤 爪 反 兮 刈 介 父 爻
[一]	[一]	[一]	[一]	[一]	[一]	[一]	[一]
乙	刁 了 乃 刀 力 又 乚	上 小 口 山 巾	广 亡 冂 丫 义 之	丰 王 井 开 元 夫 天 元 无	止 少 日 日 中 贝 内 水 冈 见	手 午 牛 毛	
2 画	二十 丁 厂 七	千 乞 川 亿 彳	尸 巳 巳 弓 己 卫 子				
[一]	[一]	[一]	[一]				
卜	三 千 于 于 亏 士						
[一]							
八 人 入							
[一]							
[一]							

参考資料5 普通話異讀詞審音表(一部) 1985年 《語言文字規範手冊-增訂本》(語文出版社1991)より  
 発音を一つにしたものは(統読)とし、複数の発音があるものは(一)(二)で示し、口語音と文語音に違いがあるものには、(語)および(文)で示しめしている。

**A**

阿(一)ā  
 ~匍 ~罗汉 ~木  
 林 ~姨  
 (二)ē  
 ~洩 ~附 ~胶  
 ~弥陀佛  
 挨(一)āi  
 ~个 ~近  
 (二)ái  
 ~打 ~说  
 癌 ái (统読)  
 藪 ǎi (统読)  
 蕩 ǎi (统読)  
 隘 ài (统読)  
 谄 ān (统読)  
 掩 ǎn (统読)  
 昂 áng (统読)  
 凹 āo (统読)  
 拗(一)ào  
 ~口  
 (二)niù  
 执~ 脾气很~  
 坳 ào (统読)

**B**

拔 bá (统読)  
 把 bà  
 印~子  
 白 bái (统読)  
 膀 bǎng  
 翅~  
 蚌(一)bàng  
 蛤~  
 (二)bèng  
 ~埠  
 傍 bàng (统読)  
 磅 bàng  
 过~  
 鮑 bāo (统読)  
 胞 bāo (统読)  
 薄(一) báo (語)  
 常单用,如“纸很  
 ~”。  
 (二) bó (文)  
 多用于复音词。  
 ~弱 稀~  
 淡~ 尖嘴~舌  
 单~ 厚~

参考資料 6 筆画法を使った索引の実際の例 (《中国民間信仰風俗辞典》中国文聯出版公司1992) より

### 巫卜禁忌类

#### 1. 巫卜

##### 二画

〔一〕

七夕卜巧 …………… (756)

〔丨〕

卜 …………… (757)

卜宅 …………… (757)

卜居 …………… (758)

卜筮 …………… (759)

〔丿〕

八字 …………… (759)

八卦 …………… (760)

##### 四画

〔一〕

云占 …………… (761)

五行 …………… (761)

历谱 …………… (762)

〔丿〕

牛卜 …………… (762)

手相 …………… (763)

手卦 …………… (764)

风水 …………… (764)

〔丶〕

方术 …………… (766)

六壬 …………… (767)

六甲 …………… (767)

文王课 …………… (768)

##### 五画

〔一〕

术数 …………… (768)

石敢当 …………… (768)

〔丨〕

占 …………… (770)

占卜 …………… (771)

占候 …………… (772)

占梦 …………… (772)

田鸡卜 …………… (774)

〔丿〕

丛辰 …………… (774)

鸟占 …………… (774)

##### 六画

〔一〕

厌胜 …………… (775)

厌殃法 …………… (776)

〔丿〕

杂占 …………… (776)

〔丶〕

问仙 …………… (777)



参考資料 7 : 情報交換用漢字コード字符集基本集の例 (GB2312-80) 1981年5月1日  
 《語言文字規範手冊-增訂本》(語文出版社1991) より

**16区** 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19  
 01 } 啊阿埃挨哎唉哀皑癌藹矮艾碍爰隘鞍氨安俺 A  
 19 }  
 20 } 按暗岸胺案肮昂盎凹敖熬翱袄傲奥懊澳芭捌扒 B  
 39 }  
 40 } 叭吧芭八疤巴拔跋靶把耙坝霸罢爸白柏百摆佰  
 59 }  
 60 } 败拜裨斑班搬扳般颁板版扮拌伴瓣半办絆邦帮  
 79 }  
 80 } 梆榜膀绑棒磅蚌镑傍谤苞胞包褒剥  
 94 }

**17区** 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19  
 01 } 薄雹保堡饱宝抱报暴豹鲍爆杯碑悲卑北辈背  
 19 }  
 20 } 贝钡倍狈备惫焙被奔笨本笨崩绷甬泵蹦迸逼鼻  
 39 }  
 40 } 比鄙笔彼碧蓖蔽毕毙毖币庇痹闭敝弊必辟壁臂  
 59 }  
 60 } 避陛鞭边编贬扁便变卞辨辩辨遍标彪膘表鳖憋  
 79 }  
 80 } 别瘰彬斌濒滨宾摈兵冰柄丙秉饼炳  
 94 }

**18区** 01 02 03 04 05 06 07 08 09 10 11 12 13 14 15 16 17 18 19  
 01 } 病并玻菠播拨钵波博勃搏铂箔伯帛舶脖膊渤  
 19 }  
 20 } 泊驳捕卜哺补埠不布步簿部怖擦猜裁材才财睬 C  
 39 }  
 40 } 睬采彩菜蔡餐参蚕残惭惨灿苍舱仓沧藏糙槽  
 59 }  
 60 } 曹草厕策侧册测层蹭插叉茬茶查碴捺察岔差诧  
 79 }  
 80 } 拆柴豺搀掺蝉馋馋缠铲产阐颤昌猖  
 94 }

