

論漢字的規律

龔嘉鎮

四川達州職業技術學院中文系

所謂規律，是指事物在發展中的本質聯繫與必然趨勢。漢字的規律主要有歷史演變規律、共時用字規律和邏輯構字規律，第一條側重反映了漢字發展演變的基本趨勢，後兩條側重反映了漢字內在本質聯繫的普遍性。

一、漢字的歷史演變規律

漢字是一個不斷適應漢語發展需要而始終緩慢演變著的自組織系統。它的演變並不在於某一端，而是以系統的方式進行的。漢字系統內部存在著相互排斥又相互協調的三對矛盾：音化與意化、繁化與簡化、變異與規範。這三對矛盾的雙方既對立又統一，以此推動著漢字系統的不斷自我完善。記詞方式意音化、形體結構簡明化、社會用字標準化，是漢字系統的歷史演變規律¹。

一、記詞方式意音化

漢字的記詞方式是體現在字式結構上的。對漢字字式結構的研究，東漢許慎以「六書」說開其先，其後影響較大者當數南宋鄭樵的「母子相生」說與清初戴震的「四體二用」說。到了上個世紀，唐蘭先生首先提出「三書」說，其後的學者遞有發展，把漢字分為表意字、假借字和形聲字三種基本類型。事物在共時平面上的各種類型往往反映了歷史進程中的不同階段，而歷史演變中的各個階段總是沉澱為共時現象中的不同類型。劉又辛先生堅持從造字方式，尤其是假借字、形聲字在消長上的巨變化來研究漢字的發展歷史，於1957年提出了「漢字發展三階段」說，近二十多年又進行了更為深入的研究而作出這樣的表述：「我們把漢字發展的歷史劃分為三個階段：第一個階段是圖畫文字階段，商代以前的文字應該屬於這個階段。第二個階段是以表形文字（象形字、指事字、會意字）為基礎、以表音文字（假借字）為主體的表音文字階段，從甲骨文到秦代的文字都屬於這

1 龔嘉鎮：《現行漢字形音關係研究》（武漢：湖北人民出版社，1995年），頁14。

個階段。第三個階段是以形聲字為主體，還保留了一些表形字和表音字的形音文字階段，從秦漢到現在的漢字都屬於這個階段。」²

漢字起源於圖畫，是由記事的圖畫逐步發展而為記詞的圖像的，最初的漢字都具有很強的圖像性，是據義構形的表意字。在單音詞為主的情況下，這樣逐詞象形造字是難以適應漢語發展需要的。隨著有詞無字矛盾的日漸尖銳，已有的表意字被廣泛地當作音符使用，「依聲托事」的假借字逐步成為漢字的主體，原本以形表意的漢字一度出現了借形表音的趨勢。一字多用的假借手段雖然經濟，但兼職過多就喪失了差示原則，借用普遍又難以在方言區流通。面對書面語中的大量同形詞，漢字體系迫切需要重建形義之間的聯系，來區別同音語素以堅持表意方向。戴震說得好：「古字多假借，後人始增偏旁。」「諧聲以類附聲，而更成字。」³ 正是這種「以類附聲」手段的使用，使大批表音的假借字轉化為了形兼音義的形聲字。相承而生的形聲字保持了漢字體系的歷史傳承性，它「形聲相益」信息量大，以文組字能產性強，不僅分化了古字的兼職，而且促成了漢字的系統類化，因此很快發展成為漢字的主體。「以類附聲」相承造字一般增加的是形符，表示的是意化的趨勢。正是在意化對音化的衝擊下，漢字執兩用中地選擇了記詞方式意音化。

漢字系統中有一個最年輕的專業字群，那就是《元素周期表》中的 105 個元素字。這些字除了「金」字之外都是形聲字，而且大部分是科學家近代以來新造的字。先看聲符，104 字的聲符大多是常用的獨體字，據筆者統計，聲符與形聲字聲韻調全同者有 81 字，同音不同調者有 3 字，其餘 20 字有的可視為省聲字，有的聲符在標音上明顯可以類推。以此推測，歷史上的形聲字在產生之後一個相當長的時期裡，其表音功能也是很強的。再說形符，表金屬元素的字一律以「金」為形符（唯一例外的「汞」，其繁體「銻」原也從「金」）；表非金屬元素的字則按元素在通常條件下的狀態一分為三：如「碳」這樣的固態者從「石」，如「氦」這樣的氣態者從「气」，如「溴」這樣的液態者從「氵」，而常溫下為液態的金屬元素「汞」，其字也從「水」。元素在印歐語中一般是多音節詞，中國科學家把它漢化為單音詞，只取元素的第一個音節來為它造一個形兼音義的形聲字。先以這個音節的音譯字作聲符，再配一個表示元素屬性義的形符，如此既譯音又譯意的元素字，聲符標音以示源，形符舉類而表義，僅用四個類別性形符就系統地展示了各元素之間內在的層級類屬關係。我們不贊成再造新字，但體會科學家為元素造形聲字的良苦用心，可以使我們更深刻地認識到漢字的形聲化決不是偶然的。漢語語素單音節化的根本特點，決定了漢字要以單音節語素為單位來構形；漢民族政治統一而地域遼闊方言複雜的歷史環境，要求表音的假借字還要能夠形以示義。在單音成字據義構形上，信息量大能產性強的形

2 劉又辛、方育國：《漢字發展史綱要》（北京：中國大百科全書出版社，2000 年），頁 321。

3 清·戴震：《答江慎修先生論小學》，載趙玉新點校《戴震全集》（北京：中華書局，1980 年），頁 63。

聲字具有最優的整體功能，只有這種把形符聲符互補地統一於方塊形之中的造字方式，才能適應單音節語素既有音又有義的特點，才能滿足各方言區人們由形及義地認識漢字以理解書面共同語的要求。

二、形體結構簡明化

回顧漢字的發展史，不管是總覽源流看大勢，還是解剖麻雀看單字，漢字都存在著繁化與簡化的現象。星，甲骨文作、，像天上羅列的眾星之形，後又用它來表星光明亮的晶，為區別遂加聲符「生」而繁化為「𠂔」，使用中求省便再簡化為「星」。古文字時期尤其是形聲化過程中，由於相承造字增偏旁，漢字在整體上呈現出繁化的趨勢。裘錫圭先生指出，「形聲化就往往是跟字形的繁化聯繫在一起的。」「字形繁化的現象，大多數起了改善字形的表意、表音功能的作用。」⁴ 加旁字的產生不僅是漢字數量的增加，而且是漢字形體的繁化。因為這新增的大批加旁字從整體上增加了漢字的平均筆劃，在大勢上表現的是繁化。秦始皇實行「書同文字」，漢代在通行規範小篆的同時又擴大了隸書的應用，並且在東漢完成了隸變。這個時期的漢字在整體上出現了簡化的趨勢。一般來說，新造字因相承構形求區別而往往加旁趨繁，既成字為書寫方便圖省力而常減省趨簡，繁化與簡化都曾推動了漢字的發展。

漢字既是符號又是工具，符號信息要明確與工具使用要經濟的矛盾，始終存在於造字與用字的整個過程之中。反映在字形變化上，就是繁化與簡化這對矛盾的對立與統一。繁是求明確但過繁則使用不方便，簡是圖經濟但過簡則字形少區別。王鳳陽先生說得好：「文字演進，一言以蔽之，就是一個求區別、求簡易的過程。」「文字只有在區別與簡易相對統一時才能比較穩定。」⁵ 人們總是希望用簡省的字形傳達出明確的信息，努力尋找簡與明的最佳結合度。正是在簡化與繁化的對立中，漢字執兩用中地選擇了形體結構簡明化。

現代簡化漢字的首倡者當推 1909 年發表《普通教育應當採用俗體字》一文的陸費逵先生。1935 年 8 月 21 日，當時的國民政府教育部公佈了《簡體字表(第一批)》，此表計收 324 個簡體字，然而半年後即宣佈「暫緩推行」而收回。新中國成立後，國家積極開展漢字簡化工作，發佈實行的《簡化字總表》包括三個表，計 2235 個字，平均每字 10.3 劃，所替代的繁體字平均每字 16.1 劃，整整減省了三分之一的筆劃。由於第三表的 1753 字是根據第二表的簡化字和簡化偏旁作偏旁類推簡化出來的，所以簡化字的基本部分是第一、二表中的 482 個簡化字。許長安先生從簡化方法的角度考察了這 482 字，其中形聲簡化字 128 個(省聲者 124 字，形聲皆省者 3 字，省形者 1 字)，佔 27%；簡作會意的字 9 個，佔 2%；同音代替的字 99 個，佔 20%；符號代替的字 33 個，佔 7%；部件省略的

4 裘錫圭：〈談談漢字整理工作中可以參考的某些歷史經驗〉，《語文建設》第 2 期(1987 年)。

5 王鳳陽：〈漢字的演進與規範〉，《語文建設》第 4 期(1992 年)，頁 14-21。

字 74 個，佔 15%；草書楷化的字 139 個，佔 29%⁶。第一、二表的這 482 字加上《簡化字總表》附錄的習慣被看作簡化字的 39 個異體字，計 521 字。李樂毅先生對這 521 個簡化字作了探源查始出時代的工作，結果發現 80% 以上的現行簡化字都是在 1950 年代以前就已經流行或存在的，其中源自先秦兩漢的竟多達 165 字，佔 32%，而新中國成立後所造的簡化字不過 101 字，還不到 20%⁷。

《簡化字總表》以「約定俗成」為原則，在過去簡化的基礎上，總結和肯定了在群眾中長期流行的簡化字的成果。它用十年時間簡化了 2235 個現行通用的漢字，其減少的筆劃達到三分之一，動態地看其實際節省的功夫更多。客觀地說，這次簡化是成功的，影響是深遠的，它順應了漢字的簡化趨勢，適應了群眾的社會需求，其在提高民族文化素質上所產生的積極作用和實際效應更是巨大的。

對於漢字如何簡化，人們的認識是經歷了一個過程的。最初較多地著眼於單字筆劃的減省，甚至有人認為簡化就是減少筆劃；後來轉而著眼於字式結構的改變，有人主張採用每個音節只用一個聲符的方法來改造形聲字；近二十年來，沿著現行漢字也是一個以文組字的系統的思路，人們逐漸把目光集中到基本字符的規整上來。獨體字是漢字中單符成字的基本成分，它以字符的身份參與構字時，或用其音作聲符，或以其義作意符。因此，簡化一個獨體字充當的基本字符，不僅可以類推簡化一批字，而且還能保持原有的字式結構和構形系統，簡化字中最令人滿意的就是這樣的字。通過對金、言、糸、貝、魚、鳥、車、門、馬、頁、食 這 11 個現行常用部首也是現行常用形符的簡化，一下類推簡化了 1156 個合體字，竟然佔到 2235 個簡化字的一半多一點。漢字經過數千年的演變發展，雖然在整體上還是具有歷史傳承性的有理性符號系統，但不合規律不合理的現象，不僅隨處可見而且古已有之了。如何解決尊重約定俗成與遵循系統規律之間的矛盾，就自然成了漢字簡化工作中的難題。相信今後的漢字整理工作，能夠在宏觀把握現行漢字構形系統的基礎上，把約定俗成與簡化有理結合起來，更好地順應漢字形體結構的簡明化。

三、社會用字標準化

從甲金篆隸楷到今天的簡化字，漢字一直處於不斷的變易之中。這不僅因為漢字所記錄的漢語漢文化在不斷地發展著，使用漢語漢字的社會在不斷地演進著，而且在於千千萬萬的個人在書寫時的隨意性也使變易不斷地滋長著。但是，社會畢竟是一切文字的本質屬性，個人用字要受到約定俗成的制約，社會要求個人用字要遵循規範，否則就難以傳達信息而「沒有共同語言」了。從秦始皇的「書同文字」（小篆）、東漢的「熹平石經」（隸書）、唐代的「開成石經」（楷書），直到今天的《簡化字總表》，無不是在變易基礎上建立起來的新規範，無不是用法典化、標準化的方式來推行新規範。歷史地看，變易是絕對的

6 許長安：〈實事求是地評價簡化字〉，《語文建設》第 1 期（1991 年）。

7 李樂毅：〈簡化字的研究和簡化字的前景〉，香港《語文建設通訊》第 66 期（2001 年），頁 6-9。

漸變，規範是相對的穩定。變易積極的一面可以推動新規範的建立，其消極的一面則將衝擊現行規範的推行。面對變易與規範的對立，來的無論是積極的推動還是消極的衝擊，漢字都因勢利導地選擇了社會用字標準化。

先討論變易的消極衝擊與社會用字標準化。規範是順應客觀規律的人為干預過程，社會生活要求語言文字規範化，語言文字規範化反映了一個國家、一個民族、一個地區的文明程度和現代化水平。今天，全球化、信息化的浪潮正在席捲全世界，改革開放的中國已經加入世界貿易組織。市場經濟的空前繁榮、信息技術的高速發展、國際交往的不斷擴展，都對社會用字標準化提出了更高更緊迫的要求。國家在現行漢字的定量、定音、定形、定序上制訂了一系列規範，頒佈施行了《國家通用語言文字法》，採用法制和宣傳的手段，標本兼治地抓源頭、抓窗口、抓學校，大力推行社會用字標準化，初步遏制了社會用字混亂的勢頭。但是，出版物、影視屏幕、計算機的用字不規範現象仍然時有發生。有的電視台晚會節目演職員字幕中，常把「閭」字錯寫成廢止了的《第二次漢字簡化方案(草案)》中的「𨵿」；一些報刊上的手寫體標題，常把「美」字的下半部份寫作「火」；大型電視連續劇《水滸傳》把一個罵人的「鳥」，幾十次地錯讀成 niǎo。許嘉璐先生曾經強調：「當前要大力提高國民的語文素質，加強國民語言文字修養，提高國民語言文字規範意識和正確使用語言文字的能力。」⁸ 許先生的這番話並非只是針對普通百姓而言，從事新聞文教工作的人更應提高自己的規範意識和語文素質。

現在討論變易的積極推動與社會用字標準化。變易的積極一面具有合理性，有助於推動規範的建立與完善。這種積極的推動在當前主要表現在四個方面：現行多音字的進一步規範，現行異體字的進一步規範，人名用字的規範，政區名稱用字的規範。

《現行漢語通用字表》所收 7000 字中有 625 個現行多音字，其現行異讀音計 705 個⁹。多音字可大分為隨聲分義和同義異讀兩類，隨聲分義的多音字頗為經濟，同義異讀的多音字徒增麻煩。1985 年發佈的《普通話異讀詞審音表》，計涉及了 849 個多音字，審定統讀了其中的 587 字，對規範現代漢語語音和推廣普通話發揮了積極作用。但《審音表》還保留了 31 個文白異讀和若干習慣異讀，而這些字大多是常用的字。沒有別義功能的同義異讀是歷史音變過程中的過渡現象，是普通話中的消極因素，因此要求多一些統讀的呼聲一直是很高的。

國家 1955 年發佈的《第一批異體字整理表》，後又三次小作調整。據筆者統計，實際整理 795 組異體字，淘汰了 1025 個異體字。1988 年發佈的《現代漢語通用字表》是國家在漢字定量定形上的基本規範，其實也是在現行漢字範圍內對異體字的一次整理淘汰。

8 許嘉璐：〈在全國語言文字工作會議上的報告〉，《語文建設》第 2 期(1998 年)，頁 7。

9 龔嘉鎮：〈現行多音字分析〉，香港《語文建設通訊》第 45 期(1994 年)，頁 21-31。

2001年12月發佈的《第一批異形詞整理表》，對「普通話書面語中經常使用、公眾的取捨傾向比較明顯」的338組異形詞，給出了推薦使用詞形。進一步整理現行異體字的工作還需加強。

中國的人名用字具有深厚的人文性，一般多選用褒義字或中性字，以表達長輩的祝福與希望，或表白自己的心志與追求。國家語委先後兩次對第三次全國人口普查數據進行人名用字的抽樣統計，通過對75萬人名的統計共得人名用字4542個。其中現代漢語常用字佔63%，現代漢語通用字中的非常用字佔23%；另有629個非通用字，佔14%，其使用覆蓋率為0.2%。這個統計表明，大多數人取名是選用常用字的，但確有少數人（中國人的0.2%就是250萬）選用了異體字和古僻字¹⁰。這不僅衝擊了用字規範，而且影響了社會交際，看來《人名用字表》的制訂似應加快步伐。再說政區名稱用字。我國的省、市、縣、鄉四級政區單位計有5.3萬個，政區名稱用字中有近800個是非現代漢語常用字，其中的生僻字就有200多個¹¹；再加上形形色色的地方讀音（方言或古讀音），以及隨意性較大的少數民族語譯音用字。因此，全面規範我國政區名稱的用字及其讀音，不僅是地名國家標準化的基礎工作，也是語言文字規範化的重要內容。

要求進一步整理多音字、異體字的呼聲，要求人名用字、政區名稱用字有所規範的呼聲，其實就是變異對標準化的呼喚。借助變易積極方面的推動，建立規範完善規範就得依靠社會用字標準化；抵制變易消極方面的衝擊，堅持規範推行規範更要依靠社會用字標準化。我們應該辯證地認識規範與變易的關係，自覺地推動社會用字標準化。

二、漢字的共時用字規律

漢字歷來被人認為字數太多而難以學習掌握。甲骨文不過4500字種，東漢的《說文解字》就增至10516個，清代的《康熙字典》收字47043個，當代的《漢語大字典》收字54678個，即使普及到發行四億冊的《新華字典》也收字上萬。1980年代以來，隨著計算語言學的迅猛發展，人們終於發現漢字在總數很多這一表象後面，存在著一條極具應用價值的共時用字規律：漢字的使用具有時代特徵，一個時代的社會用字相對集中而常用字種遞相穩定。每個時代所用的字種不僅隨著社會文化的發展而遞有變化，而且一個時代流通使用的字種都大致保持在六七千左右，其中遞相穩定的高頻常用字就更少了。非常用字、非通用字雖然數量很大，但出現的概率極低。共時社會的用字集中，在很大程度上緩解了漢字繁多所帶來的困難。總字數為572984個的《史記》僅用字種4971個，總字數為731017個的《紅樓夢》僅用字種4462個。《現代漢語頻率詞典》對181萬字的各類現代

10 張書岩：〈人名用字調查和規範化設想〉，載陳原主編：《現代漢語用字信息分析》（上海：上海教育出版社，1993年），頁188-204。

11 商偉凡、王際桐：〈試論我國政區名稱生僻漢字的治理〉，《語文建設》第6期（1999年），頁13。

漢語抽樣語料進行動態統計，計得現代漢語詞種 31159 個，字種 4574 個，而構詞能力在 10 條以上的 1689 個最常用漢字，其出現的累計字次居然佔到全部語料的 91%¹²。這一數據有力地解釋了一個人們熟視無睹的現象：人民教育出版社 1990 年代編語文教材要求小學三年級掌握 1740 個生字，而三年級小學生一般都能借助字典閱讀通俗讀物了。

國家於 1988 年發佈了《現代漢語通用字表》和《現代漢語常用字表》，前表收字 7000 個，後表由前表中的 2500 常用字和 1000 次常用字組成。抽樣檢測表明：2500 常用字在現代漢語中的覆蓋率為 97.97%，1000 次常用字的覆蓋率為 1.51%，這 3500 字的累計覆蓋率為 99.48%¹³。國家把小學生、初中生的識字量分別定為 3000 個和遞增為 3500 個的教學要求，不僅切實可行，而且科學合理。上述兩個字表是國家系統整理現行漢字的重大成果，是國家在漢字定量定形上的最新規範。它著眼於現行通用而實行了分級定量，根據漢字的使用頻度與人們的文化層次，將現行通用漢字遞增定量為 2500 字→3500 字→7000 字三個級別。正是計算語言學的豐碩成果，使人們認識了漢字的共時用字規律，從而初步實現了對社會用字科學定量的世代夙願。

漢字是記錄漢語的書寫符號系統，漢字的用字規律其實就是漢語的構詞規律在社會用字上的反映，因此，討論漢字的用字規律不能不討論漢語的構詞規律。漢語構詞之要，無外乎構詞材料與構詞規則，而構詞研究的開展則應立足於對漢語根詞的深刻認識。

任何一種語言在詞的使用上都具有不平衡性，因而都歷史地形成了自己的基本詞匯。基本詞匯所代表的是與人類社會生活密切相關的基本概念，這些詞因其「基本」而全民通用而長期穩定而成為構造新詞的基礎。基本詞匯的核心部份是根詞，根詞最能反映出語言特點的本質。各種語言都有自己的根詞，漢語缺乏形態變化，其根詞因而尤為發達而頗具特色。第一，漢語的根詞一般都是單純孤立的單音詞。詞素單音節化是漢語的根本特點，上古漢語以單音詞為主，單音詞是漢語詞匯的源頭與基礎，這是認識漢語的基本出發點。第二，漢語的根詞一般都是內涵豐富的多義詞。漢語詞義演變的基本方式是義的引伸，漢語漢文化的悠久完整，為語詞特別是根詞的詞義引伸提供了跟隨社會發展的歷史大環境，從而使歷時的引伸列沉積而為共時的一詞多義。第三，漢語的根詞一般多是表示基本類屬概念的通用詞。「方以類聚，物以群分」¹⁴，漢民族的思維方式特別注重類的集合與區分。在為適應漢語發展需要而演變的漫長過程中，所有語詞都經受了優勝劣汰的歷史選擇，只有哪些最概括地表示基本類屬概念和基本語法關係的詞才能成為

12 北京語言學院語言教學研究所編：《現代漢語頻率詞典》（北京：北京語言學院出版社，1986 年），頁 II，頁 9-10。

13 國家語言文字工作委員會漢字處編：《現代漢語常用字表》（北京：語文出版社，1988 年），頁 7。

14 《周易·系辭上》，載《十三經注疏》（北京：中華書局，1980 年影印本），頁 76。

通用穩定的根詞。第四，漢語的根詞一般就是構詞能力最強的詞根。在缺乏形態變化而以句法構詞為基本構詞方式的特定條件下，漢語中現成的根詞天然地兼職甚至專職地成為詞根，生命力最強的基本詞通過職能的轉變而成為構詞力最強的基本語素。總而言之，漢語的根詞單音多義信息量大，通用穩定生命力長，高頻能產構詞力強。漢語根詞這些簡明經濟的特性，不僅決定了造句平面上基本詞的相對集中，而且決定了構詞平面上基本語素的更為集中，這就是漢語語詞雖多而基本語素較少的根本原因。

宏觀動態地觀察現代漢語詞匯，可以概括為四條：高頻的單音詞，大量的合成詞，能產的基本語素，意合的構詞模式。《現代漢語頻率詞典》的動態統計為我們提供了三組數據：佔總詞種數 12% 的單音詞的累計詞次竟佔總詞次數的 64.3%，而佔總詞種數 73.6% 的雙音詞的累計詞次則只佔總詞次數的 34.3%；出現 2000 次以上的 69 個高頻詞，內中 62 個單音詞的累計詞次居然佔總詞次數的 35.9%；構詞在 100 條以上的 70 個單音語素，所構詞種佔到總詞種數的 35.7%¹⁵。合成詞在詞種數量上早佔絕大多數，單音詞在使用頻度上仍居壓倒優勢，單音詞的這種優勢是由相對集中的單音基本詞的通用高頻來實現的，而大量的合成詞則是單音基本詞同時以基本語素的身份按照有限的句法構詞模式靈活高產有求必應地再生出來的。人們長期習慣於用舉例分析的方式來討論語素義與合成詞義的關係，苑春法、黃昌寧先生採用系統計量的方法，對 40958 個二字複合詞進行了分析統計：複合詞義是兩個語素義的組合者佔 89.8%，複合詞義發生轉化而不再是語素義的組合者佔 0.7%，介於二者之間即複合詞義與語素義有關但不全是語素義的組合者佔 8.8%。這項極有價值的統計有力地說明了：「語素在構詞時，一般總是保持原來的意義不變」，絕大部分二字「詞的意義是由兩個語素的語義共同決定」的¹⁶。複合詞是不同語素按一定語法關係結構而成的，其詞義當然不是語素義的簡單相加，但就宏觀而言，合成詞義一般總是根植於語素義及其構詞方式之中的。明乎此，將大有益於構詞研究與語文教學。漢語的基本語素是相對集中長期穩定的，構詞模式更是為數有限而有規律的，基本語素的科學篩選與構詞模式的系統歸納，因此而成為當前漢構詞研究的重點；而把語法與語義的分析結合起來，把定量與定性的分析結合起來，則是詞研究所應堅持的原則。運用現代信息技術來全面深入地研究現代漢語的基本語素及其構詞模式，是一項十分重要的基礎工作，清華大學的黃昌寧先生和北京大學的俞士汶先生在這方面進行了開創性的研究。

15 北京語言學院語言教學研究所編：《現代漢語頻率詞典》（北京：北京語言學院出版社，1986 年）。第一組數據引自頁 1489，第二組數據是根據該書〈表六 前 300 個高頻詞分佈情況分析〉的數據計算得到的，第三組數據是根據該書〈表八 漢字構詞能力分析〉的數據計算得到的。

16 苑春法、黃昌寧：〈基於語素數據庫的漢語語素及構詞研究〉，《世界漢語教學》第 2 期（1998 年），頁 7-13。此處所用數據是根據該文第四部分的統計數據計算得到的。

漢語是在音詞高度發達的基礎上出現構詞複合化的。漢語複合構詞有兩大特點：論規則，構詞同於造句；看材料，詞根就是根詞。原本高頻的根詞兼作高產的詞根，而詞根又不斷合成新詞，這種一體兩用即基本詞與基本語素的同一性，直接造成了基本語素的相對集中。與此同時，漢字由記詞轉而記錄語素，用語素合成的方法來構造新詞，不僅使構詞更為方便，詞義更為明確，而且使社會用字更為集中了。這就在客觀上大大提高了常用漢字的編碼能力與使用頻度，從而一以當十地用幾千漢字去記錄數萬語詞。這種「詞多字少」的現象，是在詞的複合化與字的多義性基礎上產生的。第一，構詞規律決定用字規律。漢語詞多而基本語素集中，漢字字多而常用漢字集中，記錄基本語素的字就是常用漢字，反之，常用漢字記錄的就是基本語素。2500現代漢語常用字在現代漢語中的覆蓋率之所以高達97.97%，就因為這些字大多記錄的是高頻高產的基本語素。第二，字的內涵大於詞的內涵。漢字在記錄漢語的歷史發展過程中，先後從增加字的義項(引伸、假借)和增加詞的長度(合成構詞)兩個方面來控制漢字的數量。「凡字，有本義，有引伸之義，有假借之義。」一字多義而異詞同形的現象相當普遍。於是，在構詞的基本語素相對集中的基礎上，記錄基本語素的常用漢字就更為集中了。綜合構詞與用字兩方面的規律，似乎可以得出這樣的認識：漢語詞多字少以字構詞而生詞熟字。這裏的「字」指的是記錄基本語素的常用漢字。

研究漢字的最終目的還是要透過字形去認識它所記錄的語詞。漢語的基本語素決不是散沙的堆積而是類聚的集合，這個集合不僅是一切合成詞賴以層層生成的物質基礎，而且是漢語語義層級網絡的核心體系。漢字的社會用字集中常用字種較少的共時用字規律，在現代信息社會中具有廣泛而重要的應用價值。它不僅可以使我們透過現行常用漢字來把握現代漢語的基本語素與構詞規律，從而科學地研究現代漢語的語義體系與有效處理漢語未登錄詞，而且可以使我們利用現行常用漢字來加強語素與構詞法的教學，從而大大提高中小學語文教學與對外漢語教學的效率。

三、漢字的邏輯構字規律

學習和教學漢字，首先要碰到兩個問題：漢字數量如此之多，先學哪些字(循序漸進)；漢字形體這麼複雜，應該怎麼教(執簡御繁)。這在小學語文教學和對外漢語教學中尤為突出。1970年代末，國家評選出的第一批特級教師是幾位小學語文老師。他們用創造性的勞動使小學生在兩年內認識掌握了2500-3000漢字，被人們稱為「手執金鑰匙的人們」。二十多年過去了，能在小學低年級完成小六年識字任務的教改班級、學校乃至片區，已經是各地都有了。他們的經驗各有特色，但根本的一條是都在努力按照漢語的構詞規律和漢字的構字規律去進行字詞教學。

漢字是一個先象形以畫文後孳乳而組字的有理性符號系統。漢字雖然形體複雜數量多，但它是由若干基本字構成的，獨體字就是漢字的原始字根和漢系統賴以形成和發展的基礎；漢字雖然形體複雜數量多，但它的構字模式是有限有理的，是由獨字通過「比類合誼」或「形聲相益」的構字方式繁衍出來的。數百基本字符在按形聲或會意的方式組合時，具有自己的邏輯構字規律。漢字的構字規律與漢語的構詞規律可謂基本一致：漢語詞多語素少，漢字字多字符少，其基本語素、基本字符就更為集中了；漢語用語素組成詞，漢字用字符組成合體字，其組合方式都簡便靈活而邏輯有理；漢語中的根詞又是構詞的詞根，漢字中的獨體字又是構字的字根，都是一體兩用而高頻高產。一句話，字多文少以文組字而生字熟旁就是漢字的構字規律。漢字構字之要，盡在字根與孳乳。研究漢字的孳乳，或可由字根而字族以把握漢字的結構體系，或可由漢語而漢字以認識漢字的發展規律；其樞機，前者在於獨體為本而以文組字，後者在於「以類附聲」而「造字相承」。

東漢的許慎首創對漢字的據形繫聯分部類從的工作，他指出：「倉頡之初作書，蓋依類象形，故謂之文；其後形聲相益，即謂之字。文者，物象之本；字者，言孳乳而浸多也。」¹⁷他強調先有文而後有字，字是由文孳生繁衍而逐漸增多的，並運用六書理論從造字原理上分析了九千多小篆字的字式結構。南宋的鄭樵在《通志·六書略》中進一步指出：「小學之義，第一當識子母之相生，第二當識文字之有間。」他強調文與字的區別，明確了「獨體為文，合體為字」；獨體之文又因在構字中的功能不同而分為主義的母與主聲的子，「二母為會意，一子一母為諧聲」，如此母子相生而成無窮的合體之字¹⁸。這就不僅強調了獨字的孳乳合體字的字符，而且指出字符有主義主聲之分，其以文組字的方式是會意與形聲。此後，研究字根如何孳乳為字族再集合為漢字系統的字原偏旁之學一直遞相漸進，並發展而為今天的漢字構形學。

漢字有兩個構形系統，一個是外部的字體結構，一個是內部的字式結構。分析字形結構的目的、方法也因此而各有側重，前者著眼於析字，逐層拆分而析形編碼；後者著眼於造字，一次切分以析形求理。漢字的字體結構經過甲金篆隸楷的歷史演變，雖一脈相承但變化很大。現行漢字的字體結構是由筆畫組合成部件，部件層層拼合為整字的。由於這是字的外部結構，字體的部件共不在乎是否成字而與整字的音義有甚麼聯繫，所以拆分帶有較大的主觀性，這就是漢字編碼之所以出現「萬碼奔騰」局面的原因。漢字的字式結構是立足於造字方式體現造字理據的結構，其字符與整字有著音或義的聯繫。漢字的字式結構古今基本一致，變化主要發生在各種字式的比重、以形表意的方式和部分

17 漢·許慎：《說文解字·敘》，引自清·段玉裁：《說文解字注》（上海：上海古籍出版社，1988年影印本），頁754。

18 宋·鄭樵：《通志略》（上海：上海古籍出版社，1990年影印本），頁112，頁3。

字符的性質上；後起的形聲字發展成為漢字的主體而多達80%以上，單字的象形表意逐步進化為分部的類符繫聯，一些字符演變成了既不表音也不表義的記號¹⁹。

讓我們還是從字體、字式的古今變化上，來考察一下現行漢字系統中的字根、字式與構字規律吧。

第一，漢字字體的一變再變與獨體字的字根性質。如同由一個語素構成的詞是單純詞，由一個字符充當的字就是獨體字，獨體字的性質就是整體象形而單符成字。甲骨文中的「人」字寫作，仿照手腳分工的直立人的形象描畫出來以記錄「人」這個詞的獨體字，後來它又帶著這個意義去充當字根參與構字，共在以文組字的過程中將原來的具義虛化擴大為一種概括的類意義，用來指人以及與人有關的各種事物、行為、性狀等，從而逐步形成了一個從「人」的類化性字族。在隸變為了更利構字的單人旁，其表意方式早已由字的象形表意發展為系統的義類聯繫了。說獨體字在字演變中喪失了表意性，是就它的圖像表意性而言的；但它的文字表意性卻隨著漢字系統的發展而得到加強，尤其是它在以字符的身份充當造字部首、高頻形符時所表現出來的那種表意性就更為高級更為概括了。《漢字信息字典》統計的7785個現行漢字中有323個獨體字，這些獨體字雖然只佔總字種數的4%，但使用字次卻佔到總字次數的26%²⁰。國家語委1997年發布的《漢字基礎部件表》，從對GB 13000.1字符集的20902個漢字的拆分中歸納出了560個基礎部件，其中有393個主形部件，另有與之同源或形近的附形部件167個。據筆者統計，393個主形部件中，由獨體字及其變體充當的有234個，佔60%；前100個高頻主形部件中，由獨體字及其變體充當的更多達73個，即佔73%。以上歷史分析與宏觀數據說明了，現行漢字仍然是字多字符少，兩三百獨體字仍然是現行漢字中使用頻度高構字能強的基本成分。

第二，部分字符變為記號與合字的字式結構。漢字的字符原本分為表義和表音兩類，兩類後來都有演變為記號的但以聲符為多。部分聲符不表音，根本原因有兩條：漢形音聯繫弱，聲符與形聲字在漸變過程中未能始終同步；漢字堅持表意性，在結構求平衡書寫圖簡便時一般總是簡化聲符。《說文解字》在分析正篆時，明確標出「省聲」的形聲字有284字，如「炊，從火，吹省聲」；明確標出「省形」的形聲有41字，如「亭，從高省，丁聲」；明確標出「省形」的會意字有83字，如「昏，從日，氏省」²¹。三項計408字，約佔總字數的4%（許說久妥處當另作別論）。小篆尚且如此，在經過隸變、楷化、簡化後的現行漢字之中，字符不標音或不表義的情況就更多了。部分字符變成了記號，但這並沒有從體系上改變漢字合文成字的字式結構，絕大部分漢字仍然是合體的形聲字或會意

19 參看裘錫圭：《文字學概要》（北京：商務印書館，1988年），頁32-36。

20 李公宜、劉如水主編：《漢字信息字典》（北京：科學出版社，1988年），頁1071。

21 潘天烈：《〈說文解字〉導讀辭典》（成都：四川辭書出版社，1995年），頁114-116。

字。不僅因為現行聲符系統的表音功能雖然不強但遇保持著表音的性質，而且更在於為數不多的核心義符的高頻高產從系統上加強了漢字的類聯繫。在 7000 現代漢語通用字中，構字在 20 個以上的現行形符只有 54 個，而它們所構成的形聲結構計 4898 字，居然佔到現代漢語通用中形聲字總數的 87%²²。這 54 個高產的現行形符又是高產的造字部首，參構了大量會意字；它們單符作字時還是使用頻度極高的常用字。其所以如此，就是因為它們所表示的是漢民族社會中最基本最核心的類概念。當然，其它形符所表示的也多是類意義，那已是下一個層次上的以形分類了。為數有限的義符通過形義統一的原則和以文組字的方式去進類聯繫，從而有力地維繫和加強了現行漢字表義的系統性和構字的邏輯性。漢字系統是在字體字式的演變中發展的，現行漢字仍然是有理性的符號系統，大多數現行漢字仍然可以進行析形求理的分析。

漢語的構詞規律和漢字的構字規律，決定了科學的字詞教學要從語素和字根抓起。利用詞多字少以字構詞而生詞熟字的構詞規律，把常用字當作語素教，可以使學生從掌握常用字入手來成批學習合成詞。用字多文少以文組字而生字熟旁的構字規律，把獨體當作字根教，可以使學生從掌握獨體字入手來成串學習合體字。面對大量的合體字，不加分地教整字，全盤拆解地教筆畫，都是不可取也辦不到的。面對大量的合成詞，忽視語素與構詞法而只顧逐詞講解，也是費力不討好的。如何自覺地從獨體字、常用字入手來有序列成體系地進字詞教學，已經為越來越多的人們所關注。因為只有在掌握了充當基本字符的獨字的基礎上，在掌握了記錄基本語素的常用字的基礎上，再來類推學習合體字合成詞，其效率才能因教與學的有理有序而大大提高。至於漢字教學所用的基礎部件，首先應該從常用字中歸納，最好從對現代漢語的覆蓋率達到 97.97% 的 2500 常用字中歸納；其次，部件的切分要有利於拼符構字，由獨字及其變體充當的高產表義偏旁要盡可留用；第三，運用基礎部件進行教學要水漲船高，隨著學生水平的提高，要及時使用成字部作乃至合體部件來開展字式結構分析，以加強漢字教學的有理性、人文性和科學性。

許慎在分析小篆中的省聲、省形現象時，是仍然把這些字當作形聲字或會意字的。今天，人們把聲符不標音或形符不表義但還能類推保持了形聲結構的字仍然視為形聲字，其原因就在於這些字還可以進行析形求理的字式結構分析，還可以從字符的類推與組合方式中找到與整字在音義上的某種聯繫，而這在字詞教中尤其有意義。總覽漢字，「獨體為文，合體為字」，「比類合道」以成會意字，「形聲相益」而為形聲字；執簡御繁學漢字，首先要掌握基本義符，功夫多下在字根字式。7000 現代漢語通用中的形聲字佔

22 康加深：〈現代漢語形聲字形符研究〉，載陳原主編，《現代漢語用字信息分析》（上海：上海教育出版社，1993 年），頁 76。

81%，分析現行漢字的字式結構，重點仍在形聲字。

綜上所述，漢字的規律就在於歷史的表意化單音化趨勢與內在的有理性系統性聯繫，前者反映了我們民族辯證用中的思維方式，後者體現了漢語字優化編碼的經濟原則。

著眼於歷史演變，漢字是一個記詞方式意音化、形體結構簡明化、社會用字標準化的自組織系統。漢字是脫胎於圖畫的自源文字，以單音詞為主的上古漢語要求它以單音詞為單位構形，尤其是民族的長期統一而方言複雜，使漢字歷史地選擇和堅持了據義構形單音成字的基本走向，就是漢字之所以能夠適應漢語漢文化的根本原因。漢字的發展以系統的方式進行的，我們民族辯證用中的思維方式對漢字的演變產生了深刻的影響。面對字式上的意化與音化、字體上的繁化與簡化、字用上的變異與規範的矛盾，我們的民族總是在承認矛盾的對立一面的同時，更多地注重矛盾的統一，往往利用矛盾雙方相互制約相互補充的關係來進行系統內外的協同整合，通過調整內部結構與改善外向功能，從而維護了漢字系統整體上的動態平衡與質的穩定。回顧在假借字泛濫的情況下「以類附聲」地創製形聲字以加強漢字的表意性，回顧在古文字形體繁難的情況下通過隸變來實現字體的全面簡化，回顧在「文字異形」的情況下用小篆作規範在全國推行「書同文字」，似乎無不是執兩用中辯證統一。結合民族思維方式和漢字所處的歷史環境來思考問題，可以使我們更好地把握漢字的歷史演變規律。

著眼於內在聯繫，漢字是一個碼元基因化、結構合成化、編碼經濟化的有理性符號系統。基因是生物體遺傳的基本單位。漢字據義構形形義統一的特點，決定了漢字的字根(獨體字)在合文成字時，漢語的詞根(常用字)在組字成詞時，就是以形表意遺傳語義的基本單位。獨體字以字根身份參與構字而「生字熟旁」，常用字記錄的語素參與構詞而「生詞熟字」，這熟旁、熟字就不僅生成而且表示了生字、生詞的語義。7000現代漢語通用字中的形聲字，其形符具有表義功能者仍然高達87%²³。上文所引苑春法、黃昌寧先生的統計表明，4萬二字複合詞中間詞義是兩個語素義的組合者居然高達90%。這兩個定量分析數據有力地說明了：高頻高產的獨體字、常用字的以形表義，在現代漢語的合體字、合成詞中仍然具有基因性質的遺傳作用。這個認識對於漢語字詞教學與中文信息處理具有極大的應用價值。徐通鏞先生指出，「語言是現實的編碼體系。一種語言如何將現實編成『碼』，使之成為語言的基本結構單位，這與該語言社團的思維方式有關。」²⁴ 漢字就是一種現實的編碼體系，其編碼機制不僅反映了漢民族的思維方式，而且體現了優

23 康加深：〈現代漢語形聲字符研究〉，載陳原主編：《現代漢語用字信息分析》（上海：上海教育出版社，1993年），頁81。

24 徐通鏞：〈「字」和漢語研究的方法論〉，《世界漢語教學》第3期（1994年），頁1。

化編碼的經濟原則。第一，漢語中的字、詞在使用頻度上，字符、語素在構字構詞的能力上，是很不平衡的；換言之，社會的用字用詞是相對集中的，造字的字符、構詞的語素是更為集中的。第二，現成的高頻獨體字往往就是高產的基本字符，幾百基本字符構成了幾萬個合體字；常用字記錄的高頻單音詞往往就是高產的基本語素，幾千基本語素構成了幾十萬條合成詞。這就是所謂的「字多文(獨體字)少」、「詞多字(常用字)少」。第三，漢語的構詞規則同於造句規則，而漢字合文成字的方式只有形聲、會意兩種，雖然缺乏形態標誌，但意合高產的組合中卻頑強而較好地保存著構形的系統性和語義的遺傳性。要言之，所選碼元的少而高產，編碼模式的精而有理，說明漢語的構詞構字是較為經濟的優化編碼。需要特別指出的是，漢字經過幾千年的演變，尤其是隸變、楷化以來的簡化、訛變，有一部分字符尤其是一些聲符已經省變了音義全無的記號，有一部分字尤其是一些常用字因結構破壞而確實喪失了構形的理據。現行漢字中缺乏規律性的部份，就主要集中在這些音義全無的記號與喪失構形理據的字上。總結表述如下：

漢字的規律

表意化單音化 的歷史趨勢	歷史演變規律	記詞方式	意化·音化	形兼音義意音化	辯證用中的 思維方式
		形體結構	繁化·簡化	單音成字簡明化	
		社會用字	變異·規範	相對穩定標準化	
有理性系統性 的普遍聯繫	共時用字規律 (構詞)	根詞作詞根 生詞熟字	構詞同造句 組字成詞	詞多語素少 用字集中	優化編碼的 經濟原則
	邏輯構字規律	獨體作字根 生字熟旁	形聲與會意 合文成字	字多字符少 造字有理	

【本文屬專著類】