

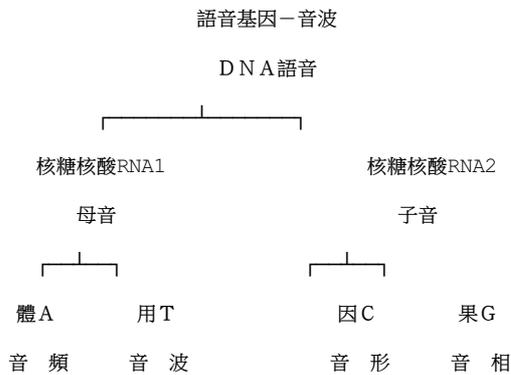
開放文學 – 漢文樂園 – 漢字基因

字音 字義 字辨 字用

五、字音：即語音。 六書中之「形聲」，適用於八成以上之漢字，可供語音辨識及合成用。

於語音辨識中，本系統採用「波形追蹤法」，可望做到不受環境干擾，能過濾噪音。尤其是此法於單音辨識之正確率極高，幾達百分之百；個人特徵亦可由波沿特徵，辨率八成；情緒緒亦同。程式空間約佔64KB，資料約佔20KB；辨識速度則受採音限制，若以技術論，以3GHz CPU每秒鐘可每秒可辨音一億個以上。

字音乃溝通基因，有母音及子音，母音由頻率決定性別、年齡，波形決定聲音感受；子音則由發聲部位因爆發、摩擦、送風、擠壓等產生之音形，及高低起伏長短頓挫之音相決定之。其基因結構如下表：



「字音」可供語音辨識及語音合成之用，任何一種漢字輸入法均須經過學習，推廣不易。若用語音辨識輸入，效果最佳。同時，當理解系統完成後，再配以自然語言，語音的辨識與合成更是不可或缺的工具了。

語音辨識早年皆採數學模式，即以行列式之統計分辨之。近來多採音訊壓縮比對式，效果較佳。本法直接取聲音之基本因素，於結構取音波之振盪為母音，取聲音形相子音。

漢語母音有「阿A伊E哎I歐O悠U」五種，子音則有齒、唇、喉、舌、風聲等。各音皆可化為點陣視訊，再分析其波形即可（詳情見拙著《模擬真實》）。

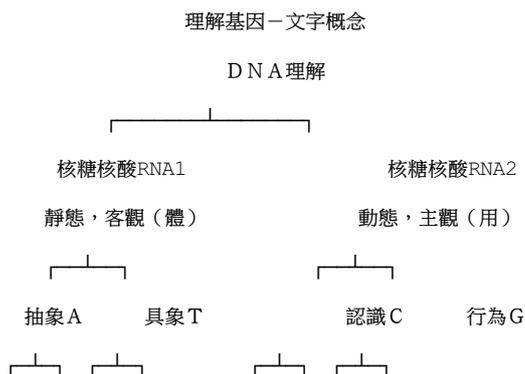
六、字義：

漢字傳衍千年，約有八萬餘留存，其中，經考究過的錯別字幾佔兩成，不再使用之閑字約五成，餘字約近三萬。於此三萬字中，絕大多數為形聲字，形聲字緣自部首與字身之結合，部首定義、字身發音。凡歷代人仕、山川河流、地理方位、動物植物等皆須有名，其數無窮無盡，且累出常變，將形聲字視若符號，實用正宜。

象形字時代久遠，難以稽考，所幸常有甲骨出土，學者已整理出三千定義明確之字，與當今概念若合符節。此外，尚有五千非符號之形聲字，其部首、字身皆係象形，見象知源、溯源得義。作者歷數年時間，翻閱古籍，整理為字義基因。

漢字基因各大功能中最重要的是「字義」基因，由於漢字的「象徵」特性，概念在大眾約定俗成的運作下，對一應客觀現象、事物以及主觀感受、認識等，都形成了合適的「字義基因」作為理解的介面。

於靜態，大自然中客觀之事體，相當於單核之一，其下有抽象之鹼基A，具象之鹼基T。而於動態，人類主觀認識之鹼基C，與行為之鹼基G是為用。從理論上說，但凡某一機體，能於大自然中將各種事物正確無誤地應用於其主觀動作、行為，即稱理解。



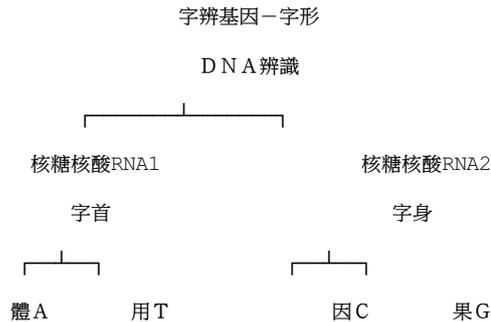
同時，漢字來自象形，故而概念中亦具有圖形基因，抽象之鹼基，可先建立象徵資料庫，具象客觀體則直接建立圖形庫；至於認識與行為鹼基，則直接接以概念基因、再經理解系統轉化為動態圖形即可。

當然，科學重視實證，上述基因必須經過驗證，始能奠定其價值。為此，文化傳信公司特別設計了「全自動圖文系統」，以輸入漢字（或由記憶調出）的方式，即時在電腦、手機的屏幕上，直接顯示與輸入一致的動畫圖形。

表面看來，此事不可思議，其實，在字義基因與圖形基因同步下，文字、劇本原本就是客觀圖形的表徵。關鍵在於字義及圖形的基因，人有意識，意識能理解概念，概念能代表圖形，電腦系統能及時、正確地將各種基因聯繫為一，就是圖文一體的全自動系統。

七、字辨：

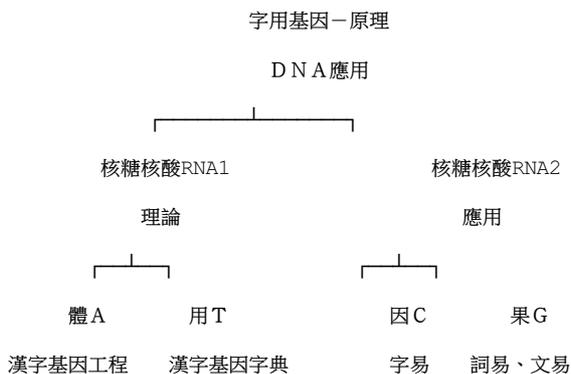
以掃描點陣資料輸入，利用「線性追蹤法」，比對前述七十三類字形基因特徵，轉換之字碼即為文字辨識。由於字形與字辨基因相同，完全符合人類認字（辨識）及寫字（組字）的根本原理。



「字辨」有多種功能，除了文字辨識外，也有影像辨識的基因在內。只是限於中文平台的機能，「字辨」尚須配合其他硬體設施。

八、字用：概念應用

有《漢字基因工程、漢字基因字典》《字易、詞易、文易》供各界使用。



由於紙本污染環境，以上五書僅在網上公佈，歡迎取用（網址：CNFLABS.COM 及 開放文學.COM）