

以聲符部件為主之漢字學習系統設計研究

The Design of Chinese Character Learning System

Based on Phonetic Components

張嘉惠 Chia-Hui Chang

國立中央大學資訊工程學系

Department of Computer Science and Information Engineering

National Central University

chia@csie.ncu.edu.tw

吳文斌 Wen-Pen Wu

國立中央大學資訊工程學系

Department of Computer Science and Information Engineering

National Central University

995202021@cc.ncu.edu.tw

摘要

有越來越多人把中文當作第二外語學習，本文目的是使其輕鬆識字。在常用字方面，形聲字約佔 60%，為此我們提出以聲符部件為主的漢字識字教學，再匹配適合的發音規則幫助漢字的學習。本系統類似於以字帶字識字教學，達到教導少量部件與少量部首，卻認識大量的字。最後與中文專家合作，從中研院漢字構形資料庫，分析出 1453 筆常用部件配合形音義及字詞教材，建立以聲符部件為主之漢字識字線上學習系統，並且由發音強度、筆畫數及其延伸字出現頻率作為教學順序。前 400 個部件與其延伸字時，對於一般文章，可達到 60% 以上的識字率。前 800 個部件就可達到 9 成左右的識字率。

Abstract

An increasing number of people learn Chinese as second language in the world. About 60% of Chinese characters are picto-phonetic compounds which are composed of a phonetic component (PC) and semantic component. Therefore, one can make a guess at a character's pronunciation and meaning from its phonetic and semantic component for a new character. For this reason, we propose an order of phonetic components based on pronunciation strength, frequency and number of strokes for efficient learning with proper pronunciation rules and graph recognition. We adopt stem-deriving instructional method which extends each phonetic component with different radical component to derive new picto-phonetic compounds of similar pronunciation. Via simulation, the top 400 phonetic components and their

picto-phonetic extensions are enough for the recognition of 60% characters in general articles; and top 800 phonetic components can help recognition of 90% characters of general news articles.

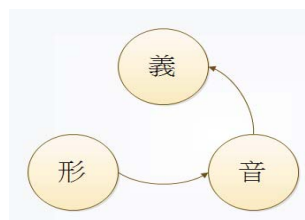
關鍵詞：形聲字，聲符部件，部件，以字帶字

Keywords: picto-phonetic compounds, phonetic component, component, stem-deriving instructional method.

一、緒論

這些年由於中國市場的崛起，加上華人第二代或第三代在海外的擴展，在台灣，由於外籍配偶人數的增加，使學習中文的人數增加。漢字的字形繁瑣，初學者難以掌握。其中最主要的原因在於漢字是圖形文字 (pictograph system)，圖形文字大部分為從字中表示出字的意思，若字本身形狀與音的連結度不高，就無法像英文等拼音文字(alphabet system)一樣，掌握了其拼音規則，就有基本能語文能力。再者，因為很難從一個漢字中得知其發音，通常我們還必須以漢語拼音(Hanyu pinyin)或是注音符號(Chinese phonetic symbols)等拼音輔助，才可知道每個漢字的發音。

以往中國有著相當比例的文盲，與現今對於海外人士的第二代、第三代，或是臺灣地區的新移民（新移民指的是外籍新娘或是外國人士），這些人由於平時都有機會接觸使用中文的人，自然有一定程度基本口語，卻可能因為不識字亦或是識字程度不高而無法閱讀。這些人在平常生活中，雖已有音跟義中間的連接關係，卻缺少與形之間的連接，因此還是看不懂中文字（圖一）。以圖一來說，認識一個新的中文字，是需要形音義三者都結合，才等於認識了一個新的字。



圖一：語言學習形音義關係

另一方面，英文在認識一個新單字時，由於形與音之間有自然發音的連接，只需要記得其讀音與意思，即可學習新的語彙。相較於華語在學習一個新字時，首先要先建立形跟音中間的連結，再來要音跟義的連結，當形音義三者結合起來之後，才學習了一個新的字，可以說華語新字學習的成本，遠較英文單字的學習成本高。因而本篇論文主要目標是幫助漢字學習者強化形與音的連接關係，讓已經有基本口語能力的使用者可以輕鬆識字，讓人在念出字的讀音之後，透過其本身已有之音與義的連結，即可了解其詞句所包含的意思。

漢字分成六大類[1]據統計資料，教育部訂常用字 4783 個，其中形聲字佔 3026 個，佔總常用字中的六成以上。這麼多的形聲字在構字上，多採用[1+1]的方式，

也就是一個部首部件加上聲符部件，對於此種現象，我們可以從聲符部件出發，並擬定一套教學順序，並且建立以部件為主之漢字識字線上教學。

雖說部件教學意義重大，也可降低學習的成本，但卻未受到現有教學的重視，市面上部件介紹零散、不成系統；部件教學過於隨意、缺少完整計劃性[11]。因此本篇論文主要目標是希望用部件教學的角度，幫助漢字學習，並實際設計了一套以部件為主的漢字學系系統，更進一步的強化形與音的連接關係，讓已經有基本口語能力的使用者可以輕鬆識字。

本文第一部分著重於聲符部件的順序，總計 1453 個常用部件，與其延伸字即可涵蓋大部份教育部訂定的常用字，依照此種部件排序做為教學順序，在學習前面的字時，其涵蓋延伸字的學習曲線相對其他順序教學，有較高的投資報酬率。

第二部分在於強化形跟音中間的連結，並做成一套形音義三者結合，且以聲符部件為主的中文識字線上教學。在形的方面，我們提出了相似構字矩陣，可以一次教導多個相似構字的字或部件；在音的部分，我們對於每一個聲符部件針對聲母韻母呈現視覺化轉音規則，及匹配每個聲符部件的形聲字關聯規則；最後提供每個字基本屬性，並且參考邱博瑋 2012[8]論文裡繪製代表其字意義的圖形。

理想上，前 400 個部件之教學，其整體識字率已達到六成以上，對於市面上的一般文章，在學習了前面 800 個部件，整體識字率更高達九成以上，本文雖未能包含實地教學的實驗，但從模擬實驗中得到一些數據參考。

二、相關研究

中央研究院資訊科學研究所文獻處理實驗室從 1993 年開始，陸續建構古今文字的源流演變、字形結構及異體字表，做為記錄漢字形體知識的資料庫，也就是漢字構形資料庫[5]。

為了了解漢字中形聲字發音規則的轉變，我們必須知道每一個形聲字的聲符為何。為此在 2010 年張嘉惠教授與李淑瑩等人於 ROCLING 2010[3]，提出以最佳化及機率分佈法去判斷漢字聲符，他們應用中研院文獻處理實驗室所建立的「漢字構形資料庫」，建立形聲字標記系統，並由中央大學中文所研究生與教授，以人工標記漢字構形資料庫中 14598 有注音標示的漢字是否為形聲字以及其聲符部件。

在 ROCLING 2011[2]，張嘉惠教授與林書彥等人提出聲符部件排序與形聲字發音規則探勘，對於形聲字的發音規則，找出高支持度與高信賴度的規則。另外，依據部件發音強度、延伸字出現頻率與筆畫數三種因素，比較線性加總、幾何平均與調和級數三種排序方式，發現幾何平均的方法，其延伸字涵率相較其他兩種成長更快。幾何平均公式如下：

$$Score_1(PC) = \frac{PC(\text{延伸字出現頻率}) * PC(\text{發音強度})}{PC(\sqrt{\text{筆劃數}})} \quad (1)$$

主要支持我們使用以部件教學的研究為高嘉慧在 2011 年[4]，此論文比較了

傳統分散式教學，與以部件為主的以字帶字識字教學。傳統分散式教學是以主題課文為主學習，學習的生字多為課文內所帶出的生字。而以部件為主的以字帶字識字教學，則是教導部件與部件組字規則的一種教學方法。經由教學實驗，得到一個重要結論，當使用以部件為主的以字帶字識字教學時，受試者在識字率方面會更有效率的成長，尤其是對低口語能力的人更是顯著。

三、問題描述

為驗證聲符部件對於漢字學習的幫助，本文實際設計一個以聲符部件為主的漢字識字線上教學平台，從聲符部件的教學順序，到聲符部件與其延伸字的組合關係、延伸形聲字的發音規則，加強形跟音中間的連接關係。

3.1 部件排序

在 ROCLING2011 所提出的部件排序是依據發音強度、頻率、及筆劃三個因素來組合，存在的問題是筆畫數的權重太過於高，導致筆畫少的部件容易被排序在教學順序的前面。如 丿 ㇇ ㇇ 等部件。(實驗教學字[10]的一部分)，初學漢字的人可能會因為學習太細碎的部件，而喪失了字形的結構性，另外可能因為筆劃數目太少，會讓人覺得切割太細，同時也可能讓人覺得此類部件不應該為聲符部件進而誤導使用者，所以這類型的部件希望使用者有一定的口語能力之後，在去做學習。另一個問題則是延伸字的頻率同時包含許多的非常用字，學習此類部件，真正應用情形不高。對於這兩個方面，我們對原本所提出的方法。重新定義一個聲符重要性分數計算如下：

$$Score_2(PC) = \frac{PC(\text{常用延伸字出現頻率}) * PC(\text{發音強度})}{PC(\sqrt{\text{筆劃數}})} \quad (2)$$

對於公式(2)與公式(1)排序的差異，我們比較排名前一百的部件如圖二(表格橫著看，以公式(2)為例，第一名是「包」，第二名是「分」以此類推)，公式(2)新增及減少之部件如圖三所示。我們可以發現，新增的部件多半是個體已有相當的結構性，而且都為一個完整的字，相反的減少的部件，大都是筆劃少並且結構性較低部件的部件，單獨難成為一個完整的字。

分	包	丁	古	方	令	非	巴	莫	各	包	分	非	古	莫	方	令	丁	巴	尚
尙	交	工	尙	台	亥	亡	可	寸	反	各	尙	交	亥	台	可	尙	卑	票	乍
亢	屯	干	乍	比	加	它	由	刃	于	工	反	青	加	登	亢	尙	屯	者	甫
卑	牙	白	儿	兪	且	青	皮	良	甫	它	亡	比	由	寸	旁	良	其	章	白
文	正	二	者	皇	票	召	甲	圭	申	且	倉	千	皮	婁	甫	辰	辟	圭	牙
元	其	甫	辰	乃	十	弗	斤	少	登	正	采	寮	龍	召	甲	同	申	文	易
半	旁	主	也	子	朱	今	斤	厶	共	于	朱	賓	每	里	共	同	申	文	元
采	倉	同	寺	章	卜	丿	里	氏	婁	半	朱	賓	主	區	爭	斤	扁	少	廷
至	句	易	每	占	丕	兆	支	辟	寮	果	朱	賓	主	區	爭	斤	扁	少	廷
句	爭	兪	勺	果	耳	并	廷	夫	扁	句	音	句	句	敖	高	耳	宗	軍	并

圖二：聲符排序公式(1)左與公式(2)右前 100 部件之比較

龍 賓 堯 舜 僉 區 奇 音 敖 禹 宗 軍
 夬 勺 支 ノ ト 厶 子 也 乚 十 乃 二

圖三: 聲符排序公式(1)與公式(2)前 100 部件：新增（上）及減少（下）的部件

3.2、設計以部件為主的漢字識字學習

在這一節中，我們希望設計出一套有效率的以部件為主的漢字識字學習，形的方面使用相似複合字矩陣；音的方面為形聲字發音規則。

3.2.1、相似複合字矩陣

由於形聲字多為聲符與部首所組成的複合字，因此我們想要找尋相似複合字矩陣（如圖四），希望藉由此種矩陣，讓使用者了解形聲字由部首及聲符組字的大原則，藉由所教導的聲符部件與一些部首去做結合，而產生不一樣的複合形聲字，達到以字帶字的學習效果。

	手 shǒu	水 shuǐ	火 huǒ
包 bāo	抱 bào (hold)	泡 pào (loose and soft)	炮 pào (a big gun)
操 cāo	操 cāo (to grasp)	澡 zǎo (a bath)	燥 zào (dry)
堯 yáo	撓 nǎo (to hinder)	澆 jiāo (to pour)	燒 shāo (to burn)
齊 qí	擠 jǐ (to squeeze)	濟 jì (various)	

圖四: 相似構字矩陣

過去網路上也有類似的相似構字矩陣[4]，但是由專家標記產生的方式，很難對每一個聲符部件產生對應的矩陣。我們希望自動產生相似複合字矩陣。我們使用兩個步驟的 K-NN 演算法。

第一步為，找出與給定聲符的相似聲符，再由這些聲符的延伸字中找出共同的部首，建立而成矩陣，其相似複合字矩陣的建構方法，首先為找到 K 個相似的聲符部件，我們所使用的相似度公式 Jaccard，計算兩個聲符 x 與 y 延伸字的共同部首比例：

$$\text{Similarity}(x,y)=\text{Jaccard}(M_x, M_y),$$

$$\text{Jaccard}(A,B)=\frac{A \cap B}{A \cup B}$$

$$M_x = \cup_{w \in W(x)} RC(w)$$

其中 W(x)是指聲符 x 的延伸字，而 RC(w)指的是形聲字 w 的部首。

	刀	口	手	人	食
包	刨	咆	抱		飽
梟		噪	操		
堯			撓	僥	饒

圖五：由 K-NN 演算法輸入聲符找出其相似聲符範例

我們以聲符「包」、「梟」及「堯」為例，三者之間的 Similarity 相似度計算分別為：

$$\text{Similarity}(\text{包}, \text{梟}) = \frac{\text{口, 手}}{\text{刀, 口, 手, 食}} = \frac{2}{4}$$

$$\text{Similarity}(\text{包}, \text{堯}) = \frac{\text{口, 手}}{\text{刀, 口, 手, 人, 食}} = \frac{2}{5}$$

對於每一預備教導的聲符部件，我們先計算其與每個部件相似度，並且找出最相似的前三名部件，並只考慮構字方式相同的延伸字。當相似部件分數一樣，我們會採用之前經過排序的部件，把排名前面的部件做為我們所要使用的部件。

第二步我們給定聲符部件的延伸字部首，去除不能同時與前三名部件和預備教導的聲符部件構成形聲字的部首，構成所要的相似構字矩陣，以圖六為例。

	刀	口	手	人	食
包	刨	咆	抱		飽
梟		噪	操		
堯			撓	僥	饒

圖六：由 K-NN 演算法刪除不要的部首範例

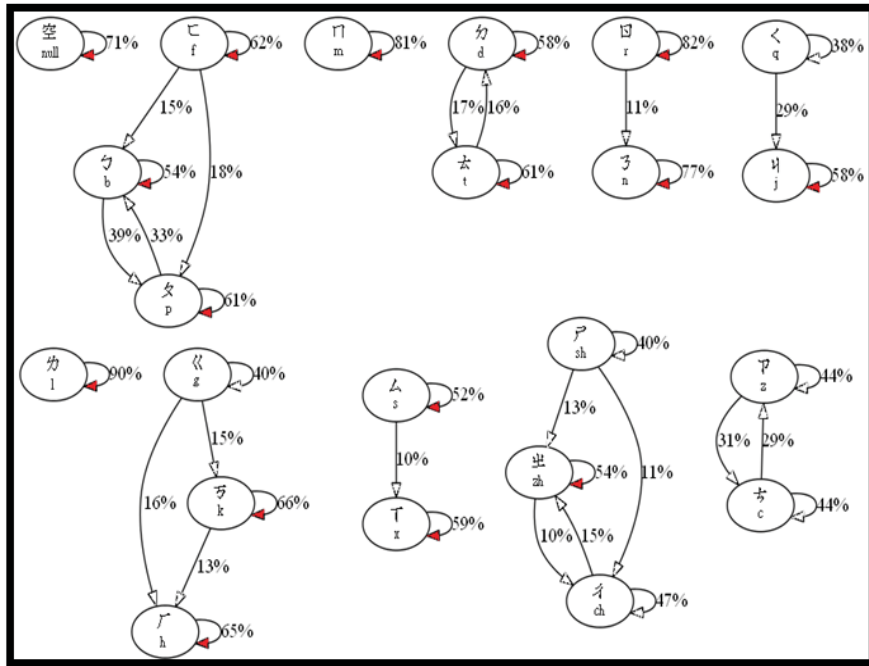
3.2.2 聲符與形聲字之轉音機率

在 Chang 等人的論文中所探勘的形聲字發音規則中，找出來規則大多數屬於沒有轉音的發音規則，同時對於這些發音規則，我們很難對其做視覺化。針對於這個部分我們提出了視覺化聲符與其延伸字的轉音機率。

在這裡，我們目標是希望能輕易的觀察出，聲符與其延伸字中間的轉音機率。為此，我們先把所有聲符與其延伸字的關係統計出來，我們使用聲符「包」與聲符「巴」當我們的例子。我們把聲符的聲母為「ㄅ」與延伸字的聲母也都為「ㄅ」的字全部統計出來，相同的我們把聲符的聲母為「ㄅ」與延伸字聲母也都為「ㄅ」的字也全部統計出來。

我們使用 Graphviz 的 Api，將轉音機率大於 10%的聯結考慮進來，即可得到最接近平面圖的圖形。藉由此種圖，我們可以很輕易觀察到每個 Node 間的轉音關係。當把所有 Node 與 Node 間的轉音機率都輸入之後，則我們可以得到圖七一樣是以聲母為例，深色箭頭代表轉音機率 50%以上，白色箭頭代表轉音機率 50%以下。由圖七所示，這些轉音關係與小學注音歌謠教導之注音有相似之處，

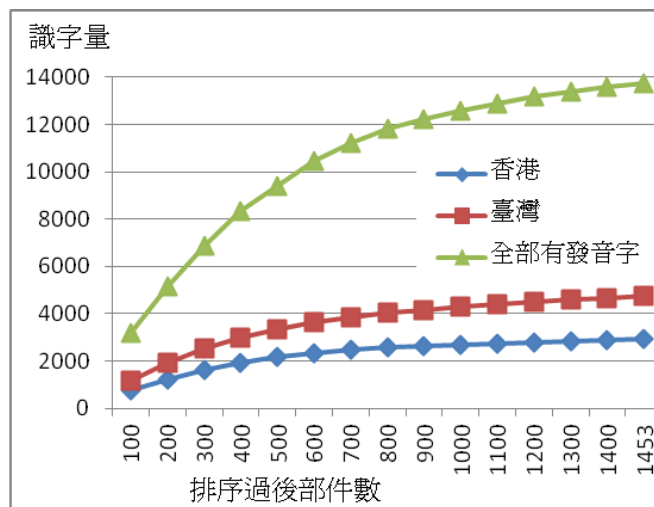
也有相異之處：如ㄅ、ㄆ、ㄇ、ㄏ、ㄏ、ㄏ不僅是發音相近，也是轉音機率高的相似聲母；然ㄇ、ㄏ則自成一組，本身發音强度高，轉移至其他發音的機率相當低；而ㄌ、ㄋ則是在注音歌謠中未顯示出來的。



圖七：聲符聲母與其延伸字的聲母轉音機率圖

四、策略模擬驗證實驗

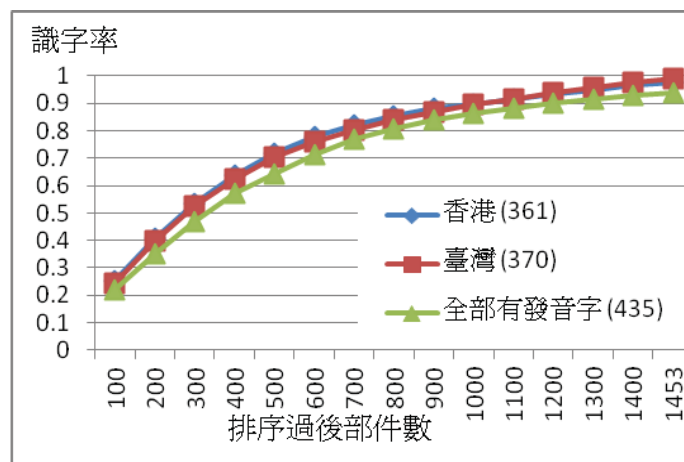
為了驗證我們所提出來的部件排序在學習上的助益，我們以學習曲線成長的情形來做為比較，學習曲線指的是累計與學習聲符部件有相似發音的延伸字數。在這裡我們用數種資料集來做評估，並用單字取向及文章取向兩種不同的方法，來計算識字率，識字率是學習曲線的數值除以資料集裡面的總字數。對於相同的單字我們只會計算一次的單字取向，與相同的單字會被重複計算的文章取向。



圖八：三種不同資料集以單字為主的學習曲線成長情形

在以單字取向的學習曲線中，我們使用三種資料集，第一種為香港地區在 1987 年[9]所訂定分成 6 個等級的常用 3000 字，第二種為中華民國教育部所制定的 4783 個常用字，最後一種為我們所使用的，漢字構形資料庫裡面所有有標記發音的字，總計 14598 個字。

圖八為三種資料集的學習曲線，X 軸為部件排序，Y 軸為累積的延伸字，我們可以發現，不管在哪個資料集裡面，其學習曲線一開始都成長得很快。我們對圖八做正規化，也就是將累計的延伸字數除上資料集裡的總字數，得到識字率，如圖九。經由圖九我們可以發現其識字率曲線對三個資料集的成長幅度是差不多的，並且大約 400 個部件(361、370 及 435)時可以達到識字率 60%以上。

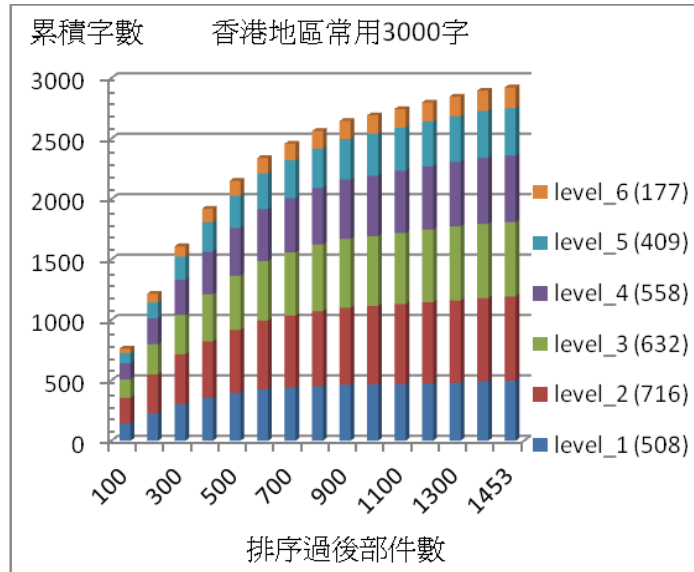


圖九：正規化以單字為主的學習曲線

深入分析聲符部件延伸字與傳統以主題為主的課文教學方法的差異，我們以柯華葳與吳敏而等人對香港常用字集所做的六個分級，在延伸字中的分佈情形，做成加強版的圖形，如圖十所示，右邊 level 的數字代表等級內有多少個字。我們發現同一聲符部件的延伸字，均勻地散佈在六個等級，這表示傳統上被視為較為困難的字，只要把其聲符部件拆解出來，其實可以降低這個字的困難度。某種程度上，這也解釋了『以字帶字』教學法對低口語者為什麼成效較佳。一般說來能力較佳的學習者，比較能自我發掘形聲字以部首及聲符部件構字的關係，並且簡化其發音的記誦，因此對於識字學習上有其優勢。而『以字帶字』教學將這層關係透明化，使得低口語者掌握識字的要訣，達到輔助教學的目的。

不過以聲符部件為主的教學策略，對於非形聲字仍有不足之處，對香港常用字集來說，有 26 個字是無法被我們的教學方法所涵蓋的，如表一所示，表二則是臺灣教育部所訂定常用字裡面不能被涵蓋的 40 個非形聲字。這些聲符部件所無法涵蓋的部分，在教學時是需要被特別拿出來教學的。

接著我們在以文章取向的學習曲線中，準備了市面上 200 字的小短文 7 篇，400 字的小短文 14 篇，500 字小短文 14 篇，600 字小短文 20 篇，皆為小學生優良作文，並從此種資料集中觀察其識字率的成長情形，並且與目前南一版國小課本前三個年級(共六冊)的教學順序所得之識字率作一個比較。



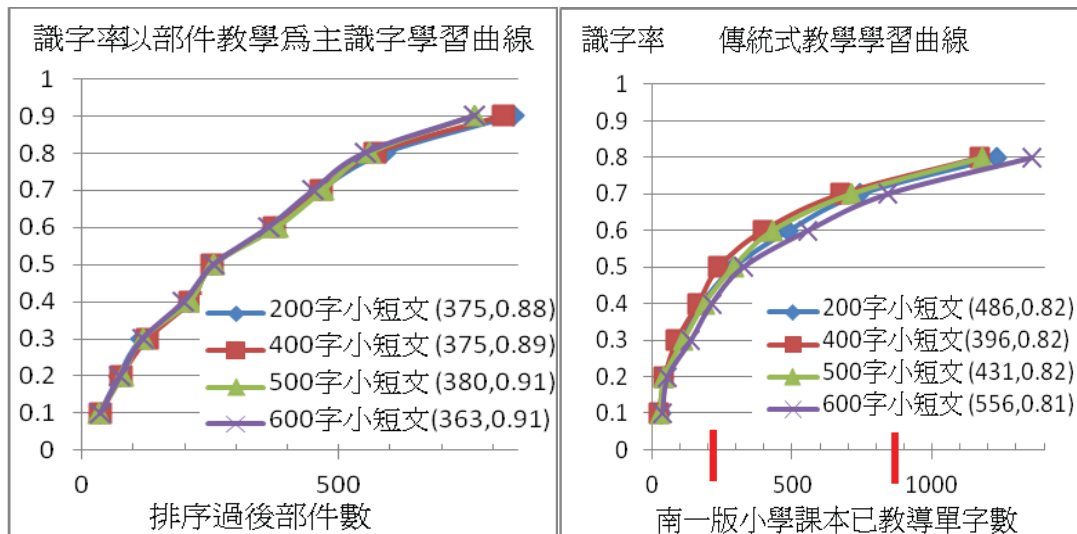
圖十：香港常用 3000 字以單字為主的學習曲線成長詳情

表一：1453 個聲符部件無法涵蓋的香港常用字

竹	已	年	事	飛	候	互	森	晶	囊
惱	傷	浸	畢	惰	龜	佩	承	絕	凸
秉	侵	姦	喪	鼎	磊				

表二：1453 個聲符部件無法涵蓋的臺灣常用字

鼎	韭	龜	飛	沉	孑	孑	候	梟	森
潛	楞	絕	已	侵	佩	弔	姦	浸	傷
畢	喪	卅	承	囊	事	互	年	竹	晶
珮	丫	凸	甩	璫	惱	愣	惰	磊	秉



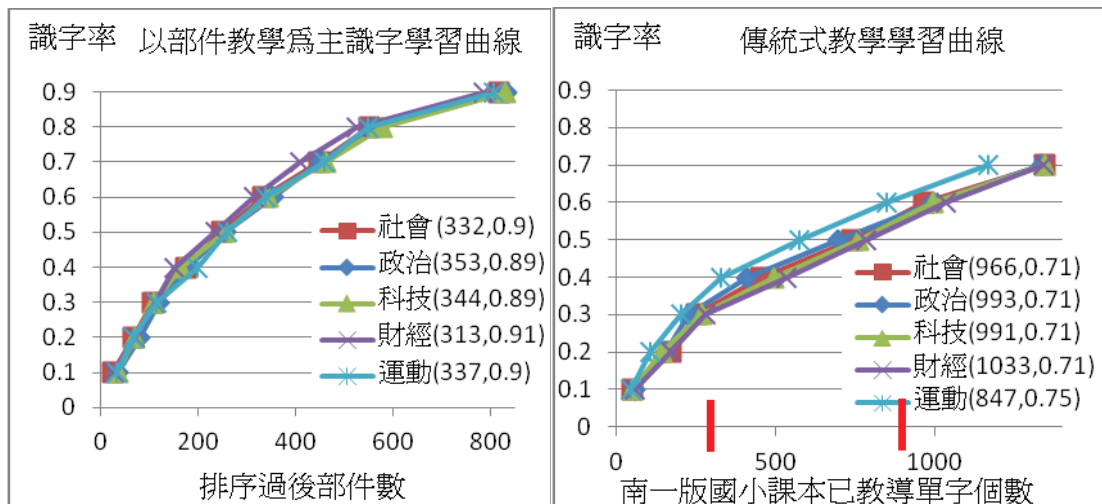
(a) 以部件教學為主

(b) 以傳統教學為主

圖十一：市面上小短文學習曲線

圖十一(a)括弧內兩個數字一樣分別為識字率 60%時所需要教導聲符部件數，與教導了 800 個聲符部件時的識字率，會選用 800 個聲符部件是希望在三年內可以交導完畢。觀察圖十一我們可以發現不管文章字數多寡，識字率 60%所要需的聲符部件個數是差不多的(363~380)。而以南一版六冊 1400 個字的教學順序，也可計算相似的學習曲線。圖十一(b)中橫軸兩條紅線為年級的分界線，意思為學習超過了線之後就為下一個年級。經由觀察兩圖可發現，我們所採用的以部件為主的識字教學，其成長曲線優於傳統式的教學方法，並且我們提出來的方式其識字率可維持在九成左右的水準，高於南一版前六冊的八成二識字率。

最後我們對於市面上的新聞，也採用類似的方式去看其學習曲線成長為何，了解以聲符部件為主之華語教材的適用性。在這邊我們選用五個不同類型的新聞，如圖十二所示，並且圖中括弧內兩個數字分別為識字率 60%時所需要教導聲符部件數，與教導了 800 個部件時的識字率。經由觀察兩張圖之後，可以發現我們提出來的方式一樣優於傳統式的教學，並且很快速的就可達到 60%以上的識字率；傳統式的教學要在第三年之後才可達到相同的識字率水準，此外在 800 個部件與 1400 個單字，我們的識字率一樣可維持在 9 成左右，而傳統式的教學卻只能在 7 成左右。



(a) 以部件教學為主

(b) 以傳統教學為主

圖十二：新聞類學習曲線

五、線上學習介面

以聲符部件為主的漢字教學線上系統分成了形音義三大部分，形的方面為「組字練習」，音的部分有「轉音機率」和「發音規則」，義的部分是「基本屬性」。

5.1 形：組字練習

在介面設計上，雖可以直接呈現相似構字矩陣，但為避免過多資訊，我們以「組字練習」方式，讓使用思考漢字的構成。另外在這裡，我們希望與使用者有

著一些互動，所以使用踩地雷的方式呈現給使用者，在點擊交織地方的同時，才呈現是否有構字，以「包」為例全部點擊完後如圖十三所示。



圖十三：組字練習介面

5.2 音：轉音機率、發音規則

在「轉音機率」的介面裡，我們對於每個所教導的部件或是字，把其聲母與韻母個別拆開，並個別繪製轉音機率圖，依照當前的聲母或韻母，對其延伸字呈現其轉音機率。以聲符「包」為例，在漢字學習系統中介面如圖十四。



圖十四：聲符之聲母與韻母之轉音機率



圖十五：基本屬性介面

5.3 義：基本屬性

對於字義的部分，在「基本屬性」中的介面中，我們呈現的內容包括字義、字音、部件拆解、以及字形來源的圖解。以圖十五為例，圖解來源參考邱博璋 2012 年[8]的論文裡繪製代表其字意義的圖形，使用者如果以象形圖示方式來記憶部件，鮮少出現錯誤或是想不起來的情形[10]。

六、結論與未來工作

本篇論文主要分成兩大部分，在第一部分我們重新定義了部件的排序，使得其更接近教學方向，經由與 ROCLING 2011[2]比較，可以發現延伸字在常用字方面，有很大的成長，並且降低了非常用延伸字的部分。在模擬實驗方面，選用不同資料集去做識字率的呈現，成長曲線遠大於一般傳統式教學，也發現在學習了前 400 個部件之後，已經有基本能力應付各種不同的資料集，並且識字率皆可達到了 6 成以上；在學習了前 800 個部件甚至達到 9 成的水準，幾乎可涵蓋整篇文章；另一部分，我們為使用者打造以部件為主之漢字線上學習系統，在形的方面我們提出相似複合字矩陣，可經由教導少數的部首與聲符達到大量識字的目的，在音的方面則有視覺化形聲字轉音機率，與匹配每個教導部件的形聲字發音規則，藉由選出高 support 與高 confidence 的發音規則，強化形跟音中間的連接關係，並且達到以字帶字的教學效果。

然而，仍然有許多地方尚待我們改進的，目前只有在針對單字的識字教學，還欠缺延伸詞的搭配，另外破音字沒有考慮。發音規則對形聲字的涵蓋可以轉成 set covering 的問題，希望可以找到少數的發音規則卻可以還蓋大量的形聲字，另外在本系統中，對每一個字拆解成部件之方式係依照漢字構形資料庫所提供之

拆解方式，但有些部份會拆解的太過瑣碎，與一般使用者認知的拆解不同，也是可以改善之處。此外目前本系統尚屬測試階段，未來還需加強和字線上學習系統的介面，或是推廣到手機上面，使之變成有趣的 App 供使用者下載，才能進行實地測試，邀請受試者的參與。最終目的，是希望正在為學習中文字發音而苦的使用者，能更快掌握漢字的發音。

參考資料

- [1] 許慎撰、段玉裁注《說文解字注》，台北藝文印書館，1988年。
- [2] 張嘉惠、林書彥《聲符部件排序與形聲字發音規則探勘》，ROCLING 2011。
- [3] 張嘉惠、李淑瑩、林書彥等《以最佳化及機率分佈判斷漢字聲符之研究》，ROCLING 2010。
- [4] 高嘉慧《識字教學法與口語詞彙能力對新移民女性中文識字學習之影響》，2010年，中央大學碩士論文。
- [5] 中研院文獻處理實驗室「漢字構形資料庫」網站。
- [6] 莊德明、謝清俊《漢字構形資料庫的建置與應用》，漢字與全球化國際學術研討會，台北，2005年。
- [7] 莊德明、鄧賢瑛《文字學入口網站的規畫》，第四屆中國文字學國際學術研討會，山東煙台，2008年。
- [8] 邱博璋《說文解字數位編輯規劃研究—以繪圖、檢索與排版為探討對象》，2012年，中央大學碩士論文。
- [9] 柯華葳、吳敏而等《國民小學常用字及生字難度研究——六年級》。台北：台灣省國民學校教師研習會編印，1990。
- [10] 賴富美《部件識字教學對國小學習障礙學生識字學習成效之研究》，2008，台東大學碩士論文。
- [11] 盛繼豔《華文教學中漢語的部件教學》。
- [12] 費錦昌《現代漢字部件探究與語言文字應用》，1996。
<http://wuxizazhi.cnki.net/Article/YYYY602.004.html>