

行政院國家科學委員會專題研究計畫 期中進度報告

繁體漢字辨識及唸名作業裡字內成份部件可能扮演的角色

(1/3)

計畫類別：個別型計畫

計畫編號：NSC91-2413-H-002-012-

執行期間：91年08月01日至92年07月31日

執行單位：國立臺灣大學心理學系暨研究所

計畫主持人：吳瑞屯

報告類型：精簡報告

處理方式：本計畫可公開查詢

中 華 民 國 92 年 5 月 26 日

行政院國家科學委員會專題研究計畫成果報告

繁體漢字辨識及唸名作業裡字內成份部件可能扮演的角色(一)

The role of radical in character recognition and naming (I)

計畫編號：NSC 91-2413-H-002-012

執行期限：91年8月1日至92年7月31日

主持人：吳瑞屯 台灣大學心理系

一、中文摘要

拼音文字的研究文獻顯示單字辨識或唸讀歷程裡可以明顯觀察到單字內部被認知處理的次字彙歷程，近十年的國外漢字認知歷程研究者也把焦點放在這個議題上，但目前為止仍沒有明確共識，對臺灣繁體字環境進行此類研究的也顯得不足。本研究擬透過字詞資料庫內容的擴增與分析，完整建立單字內的部件屬性作為研究基礎，並有系統評估國外有關漢字辨識或唸名作業裡字內成份部件扮演角色的實驗研究結果，並設計更周延的實驗企圖回答臺灣繁體字的辨識歷程裡是否涉入明顯可見的成份部件處理歷程，如果答案肯定而且技術上也可行，則企圖透過更進一步的設計，以實驗闡明成份部件訊息被處理的時間歷程。實驗將考慮包含認字作業及唸字作業，測量反應時間，遮蔽式與非遮蔽式促發程序，SOA的操弄盡量包括能被穩定觀察到的最短時間。

關鍵詞：漢字、字彙接觸、辨認、唸讀、促發、部件處理

Abstract

The literature on lexical access shows that sublexical process in word recognition and naming is obviously observed in alphabetic orthographies. The present project is to explore the possibility that radical processing during Chinese character recognition and naming can be reliably observed. If possible, the time course of information processing will also be traced. A database about features of radicals and their connections with characters will be firstly constructed. Then, a series of character recognition and naming RT experiments will be conducted. Masked and Unmasked priming procedure will be adopted. SOA will be manipulated as short as technically

possible.

Keywords: Chinese character, lexical access, recognition, naming, priming paradigm, radical process

二、計畫緣由與目的

認知心理學裡的文字辨識歷程研究

文字閱讀歷程是人類高度進化的產物，這種訊息傳遞方式使得其他生物在進化競爭裡遠遠的望塵莫及，文明由之而生，創造發明也因而加速，生活形式於焉改變，世界所以隨著人類意願而進展。但文字卻非自然產物，而是人類自己創造的符號，不同的文明各有自己的文字，人們利用自己創造的符號架構語意進行溝通，甚至影響思想，因為文字，思考變得洗煉精確，文明也越來越去蕪存菁。戰爭使得較進化的文明透過更精緻的器具及更考究的策略摧毀弱勢的文明，被征服或消滅的文明往往會失去文字。相對於其他生物，人類怎麼會有這種形式的演化，什麼樣的認知系統特性可以支持這種人工符號的創造與處理運作，這無論如何都是心理學裡最重要的議題之一了。

文字的最終目的是表達意義，但在尚未有公認的文字符號之前，人們卻透過發聲器官產生有限的語音表達意思，語意與語音的對應是任意的，伴隨人類的出現而出現，這種自由的決定隨著不同口語的傳承而成為習慣，各有不同。口語成形經過數十甚至百萬年之後開始產生表徵的文字符號，最初的象形文字是表意的符號，後來出現的拼音文字設計卻朝表音的方向發展，表徵的單位從音節，如日、韓本土發展的母語表音文字，到表徵更抽象的音素，如歐陸的拉丁、英德文字，基本的建構元素希求表現語音符合越來越簡約細緻的原則，卻是以依賴由特定地域發展的母語音系統之熟練而聯結語意作為假設的先決條件，不熟練該語音系統的其他文明人們並不能借助於文字符號的任何線索取得語意，這成為推廣此類系統的極大限制。

漢字是古老象形文字歷經數千年沒有被淘汰的不朽文字之代表作，三千年前出現甲骨文的商殷時代，不同方言的各地已各自發展出獨特的象形表意文字，中國境內各地的相互流通與秦皇朝政治上的全盤箝制需求造成文字統一的必要性，但各地語音殊異的不同方言卻是一種不易妥協的現實，就如同現今中國境內仍然存在的眾多方言一樣，表意的甲骨文遂成為統一文字的基礎架構，由於無需依賴語音元素建構意義，而是企圖通盤反映人們認知系統裡的語意世界之主要分類架構，亦即金、木、水、火、土的五行世界觀點，使得統一後的文字，即後來的漢朝文字，數目不足三千，但在拼音文字讀者眼中，已認為在學習上是一種不自然的龐大認知記憶負荷。公元 400 至 600 年間，漢字傳入朝鮮，成為後來書寫韓文的一部份，韓文裡夾雜混合了本土設計的音節符號文字及移植來的漢字，閱讀時，漢字的讀音有兩種方式，一種是仿據中國人的讀音，另一種是依據語意設計讀音。同樣的，漢字也在公元 600 至 800 年間傳入日本，成為後來日本文字的一部份，日文裡也夾雜混合了本土設計的音節符號文字及移植來的漢字，漢字的

讀音也有兩種方式，一種是仿據中國人的讀音，另一種是依據語意的母語音設計讀音。現在留在莫斯科博物館的西夏黑水溝古城遺址文物也顯示，在蒙古滅亡西夏的之前兩百年間，地處古文明帝國波斯與中國通道之間的絲路上之西夏已經大規模的模仿漢字建立自己的文字，但更西邊的波斯人所使用的文字卻是根據波斯母語發展的拼音文字。與中國在歷史上水乳交融的滿、蒙、藏，等的文字，也都漢化。顯然，漢字的設計及發展過程中避開了地區性語音的糾纏牽絆，而往直接聯通語意的目標逼近。從現存至今的中、日、韓文字的發展軌跡來看，漢字數目並無逐漸減少的現象，三千的數目雖然看似龐大，小學生對於漢字的學習卻也沒有太大的困難，顯然，漢字可以看成人類文明創作的一顆極為耀眼明珠，是不同母語的人們都搶著要的寶貝，不通古音的現代華人甚至可以瞭解祖先的古文字而了無障礙，這是遠遠的不同於等閒的一般拼音文字的。中國的漢字在數百年前已增至六萬字以上，但常用流通的字仍只維持在五万左右，三千個漢朝的文字就算在近兩千年後的現在仍然可以反映現代讀物的 99%。雖然近半世紀前，大陸的文字簡化運動使得漢字的外型筆畫簡化，字數也縮減到不足七千，但在臺灣，古老的傳統繁體漢字仍保留一萬三千字不變而正常的運作著，甚至安然渡過二十世紀電腦化革命的嚴酷洗禮而完整無恙。簡體漢字利用了北京語音的讀音特性以較簡字取代較複雜同音字原則，如「里」取代「裡」，難免鬆散了字形與字義的嚴格對應關係，但大體上，仍沒有明顯證據可以顯示簡體字與祖先習慣的繁體字之認知處理上有明顯的不同，臺灣的大學生面對大陸的簡體漢字或大陸的大學生面對臺灣的繁體漢字，在適應上都不會有太大的困難，因為經常出現的單字大部分仍是相同的，語意的掌握完全一樣。

從這樣的分析來看，由於設計原則的不同，拼音文字與表義的漢字在認知處理的過程裡所面對的條件上也自然會有所不同，特別是在閱讀最基礎的單字辨識過程之上，從心理學文獻裡有關探討字彙接觸的論文超過兩萬篇以上，遠多於其他議題，其重要性由此可知。文字認知歷程的認知心理學研究最早大量出現於 1970 年代，主要以美國人對拼音文字裡的英文為主要研究對象所進行的工作，單字辨識歷程就是一個重要的探討議題，為了處理實驗室外學習經驗的可能影響，在現代認知心理學仍未大量流行的年代之前就已經有先遣的研究，從事單字日常生活中出現頻率的計算工作，雖然那時並未有像今日一般方便的計算工具。1970 年代以後至今，用來探討單字辨識歷程的研究共發展出如下的作業：辨字作業、唸字作業、語義判斷作業、眼動循跡作業、及 Stroop 作業，其中尤以辨字作業及唸字作業為最。早期的研究工具以速示器為主，1980 年代後期，電腦逐漸普及，1990 年代以後，速示器的角色被個人電腦取代，更複雜的設備如眼動循跡器、核磁共振等也陸續出現。在 1990 年之前，拼音文字辨識研究的主要議題在語音處理扮演的角色，1990 年之後發展出的議題則包括次字彙(即單字內成份字母串)的處理歷程、語音激發的時間歷程、不同理論模型的更細緻討論等等。

漢字單字辨識歷程研究之過去進展情況

參考並取法拼音文字的研究經驗，探討華文單字辨識歷程的心理學研究最早出現於 1970 年代中期的臺灣，以臺灣所使用傳統的繁體漢字為研究對象，也是從基本的字詞日常出現頻率資料庫的建立做起，研究工具隨著時代的推移大致雷同，但在 1980 年代以後至 1990 年代中期，因應電腦處理漢字的需要，基本的實驗程序控制工具也慢慢的被建立起來，部份的研究者開始比照拼音文字研究的議題進展，引入理論上可以進行比較的有關變項進行探討。1990 年代起，大陸的研究者開始大量的透過與拼音文字研究者的合作或後者的指導從事以簡體漢字為對象的研究，目前世界上不同地區所進行的漢字認知歷程研究，已是大部分以簡體漢字為對象的局面。臺灣早期的主要研究者除了少數仍持續不懈的延續著研究議題之外，其他已淡出對過去熟捻主題或工作的堅持。

漢文單字辨識歷程研究之最近十年發展反映了對拼音文字研究者理論進展的影子，重要的研究走向包括如下議題：

一、語音處理歷程在單字辨識歷程裡扮演的角色或其必要性如何。

這樣的研究探討諸如字詞確認階段或字義提取階段之前是否存在語音處理歷程，若然，其運作細節或時間歷程又如何等的議題。在拼音文字研究的文獻裡，單字辨別確認過程裡出現可觀察到的字內成份字母串的拼音歷程是已無異議的，從字形激發語音到確認單字的時間歷程也逐漸勾勒出來，這樣的研究進展成為漢字文字研究參考或取法的對象。支持漢字在這方面表現與拼音文字沒有重大差異結論的研究如 Perfetti & Zhang (1991)、Perfetti & Zhang (1995)、Tan, Hoosain, & Peng (1995)、Perfetti & Zhang (1996)、Tan & Perfetti (1997)、Perfetti & Tan (1998)、Zhang, Perfetti, & Yang (1999)、Ziegler, Tan, Perry, & Montant (2000)、Spinks, Liu, Perfetti, & Tan (2000)、Pollatsek, Tan, & Rayner (2000)等。持反對觀點的研究如 Wu, Chou, & Liu (1994)、Liu, Wu, & Chou (1996)、Cho & Chen (1998)、Shen (1998)、Weekes, Chen, & Lin (1998)、Shen & Forster (1999)、Wu & Chen (2000)、Wu & Chou (2000)、Chen & Shu (2001)、Chen & Peng (2001)。其中，有熟捻漢字經驗而支持漢字具有類同於拼音文字表現特性的華人研究者不多，以大陸出身的 Tang, L.H. 為主，研究對象為大陸的簡體漢字，但大多的實驗結果卻無法被臺灣、香港、澳洲、美國甚至大陸等不同地區的研究者以繁體漢字及簡體漢字為對象的複驗研究所重複，其中更包括原先與 Tang 合作的華人研究者。這個議題的爭論雖然仍在繼續，但看局面的發展態勢，應該已近尾聲。

二、認知的心理單位或成份部件處理在辨識歷程裡扮演的角色。

探討與拼音文字辨識裡的次字彙(sub-lexical)歷程與相關理論模型進行比

較，漢語字詞有何不同的議題。支持與拼音文字類同，漢字辨識歷程裡存在部件處理的研究包括 Flores d'Arcais, Saito, & Kawakami (1995)的日文漢字研究、Feldman & Siok (1997)、Peng & Wang (1997)、Taft & Zhu (1997)、Feldman & Siok (1999)、Li & Chen (1999)、Li (1999)、Taft, Zhu, & Peng (1999)、Zhou & Marslen-Wilson (1999)等的漢字研究等，d'Arcais, Saito, & Kawakami (1995)的日文漢字實驗採用先呈現單字的部首部份或聲旁部件 SOA 時間後隨整字呈現的促發唸字作業，操弄部首與全字的語意關連、音旁與全字的讀音一致程度、與 SOA(60、180ms)，結論是音旁訊息的激發明顯比部首強，隨著 SOA 增強，音一致時有明顯增進效果，在 SOA 為 60ms 時，語意關連時有增進效果，SOA 增為 180ms 時，義近音同時有增進效果，義異或音異時則有抑制效果，但義異且音異則反無影響效果。Taft 等的研究則透過針對簡體漢字出現頻率資料庫對部件出現頻率的計算與操弄，進行單字判斷與唸讀單純的作業，結論宣稱單字中成份部件的組字能力強時容易被處理，但僅限部件在字的右半位置，以此闡明單字處理歷程裡有能被觀察的成份部件被處理現象，但 Feldman 與 Siok(1997)提出的批評指出其結論裡的位置被功能所混淆，因為右半通常是表音部件，當表義表音的混淆被有效控制後，則位置效果自然消失，只剩下成分頻率效果。Feldman 與 Siok(1999)更進一步採用促發字彙判斷作業，操弄促發字與目標字之部首形似與字義關聯，結果形似義異產生抑制作用，義似產生增快效果，間隔 10 嘗試之嘗試間促發只有在形似義似時產生增快效果。Li (1999)採用認字作業在香港所進行的繁體字博士論文研究則宣稱跟單字層次處理有關的部件處理才有效果。英國，Zhou & Marslen-Wilson (1999)利用大陸學生進行簡體漢字實驗，採用促發唸字作業，操弄 SOA 為 57、100、200 毫秒，促發字中有 1/3 為實驗關注的複合字，與目標字無關但其部件與目標字語義關聯，1/3 則為前述部件本身作為促發字，另外 1/3 為複合字的對照情況。結果被作者解釋為目標單字的內涵部件(而非目標字)之語義關聯字促發字會對目標字之唸名產生增快效果。這個結論與 Tan 等(1997)目標字之同義字之同音促發字會對目標字唸名產生增快效果一文令人有異曲同工之妙，都同樣的違反了一般華人對華文閱讀經驗的直覺。當然，這個結論也明顯與 Li & Chen(1999)的結論互相矛盾。明顯的，這是一個方興未艾的火熱戰場，就算按照拼音文字研究者的邏輯所進行的實驗，支持成份部件在單字辨識裡扮演類似重要角色的漢字研究者們，目前仍沒有共識，能通過嚴格檢驗的穩定實驗結果尚付闕如。在實驗設計與混淆變項的處理方面多存在有可以被擊垮的破綻，就以促發認字或促發唸字作業為例，除了 Flores d'Arcais, Saito, & Kawakami(1995) 以外，其他採用促發程序的實驗中使用的促發項目都是與目標字同樣層次的其他單字，而非單獨的成份部件本身，這樣進行實驗所得的結果解釋會面對意外的邏輯困難，此點容後細敘。Zhou & Marslen-Wilson (1999)的研究設計則明顯可以看出不夠周延，無法避免可能的混淆，雖有考慮加入 100 個填充嘗試企

圖淡化受試者的策略行爲，但因包含了可做爲複合字部件的單獨字做爲促發字的情況，實驗情境仍然難免可能的暗示氣氛，導致受試者注意部件而提醒受試者提高處理複合字內部件的機會，這原是較細緻的實驗設計可以避免的。不太贊成漢字辨識過程中有類似拼音文字研究明顯可見的成份處理歷程觀點的研究較少，目前只有臺灣的研究者有些初步的研究結果，Wu & Liu (1997)採用單純唸字作業，操弄單字頻率、部件本身單獨成字的出現頻率等，結果單字頻率的影響力與過去研究毫無二致，但部件本身單獨成字的出現頻率只對低頻單字的唸出速度有不太可見的影響力，Wu 與合作者(Wu & Chen, 2003)的實驗採用了 Grainger 及合作者的系列研究之邏輯(詳見後敘)，操弄時間歷程(SOA)，探討形似近鄰字做爲促發字的短暫時間遮蔽呈現對對目標字字彙判斷反應時間的影響，發現就算促發字的呈現時間短到 34ms，繁體漢字的受試者已經早就辨認出整字，不像 Grainger 等的法文或荷蘭文 5 字母長度單字辨認實驗，促發字呈現時間在 33ms 至 66ms 之間可以明顯的觀察到成份部件而非整字被處理過的跡象，這似乎顯示，拼音文字需要較長的時間進行次字彙歷程，但漢字辨認歷程卻相對的極快，在現有的實驗技術下，漢字的研究者似乎找不到一個時間點足以提供對單字內部件處理的穩定觀察。這些實驗的研究態勢相當明朗，臺灣的研究者在這方面的研究仍然需要加強。

三、成份部件促發增快效果與形似近鄰字促發抑制效果的研究。

當拼音文字的研究者大量的採用促發程序研究促發字如何增快目標字的唸名，不懂中文字的 Perfetti 甚至透過指導大陸留美研究生把這種邏輯大量的應用在他所不熟悉特性的漢字研究之上，甚至宣稱全世界的文字都應服膺「普遍語音原則」或「語音反射」時，拼音文字字彙辨認的研究戰場上已經默默的在歐陸異軍突起，對法文與荷蘭文之單字辨識及雙語研究有持續投入的 Grainger (1990) 開始注意到目標字的形似近鄰字之頻率效果，並與合作者利用非遮蔽式促發認字作業及操弄較短 SOA 時間的遮蔽式促發認字作業闡明拼音文字單字辨認的形、音訊息激發時間歷程(Grainger & Ferrand, 1994; Segui & Grainger, 1990)。當 SOA 極短(如 32ms)的遮蔽式情況，使得受試者不足以認清整個促發字而只能夠處理其中成份部件時，形似促發字對其後的目標字辨識產生增快效果，但當 SOA 夠長(如 64ms) 到足以認出較高頻的促發字整字時，形似促發字卻對其後的目標字辨識產生抑制效果，SOA 再更長(如 350ms)時，較低頻的形似促發字也產生抑制效果，但此時較高頻的促發字則不再產生任何效果。發現了促發字的形似增益及抑制效果的事實，挑戰了以往利用促發字增快目標字處理以推論成份部件被處理的邏輯，也造就 Grainger 近十年數十篇論文的一系列研究，明顯的反駁並凌駕了美國的同議題研究者。Grainger 後來採用了改良的交互激發模型以解釋其系列研究結果。當大多數的漢字研究者都還沈醉在如何闡明漢字辨認的單字內部件處理

時，筆者投入了最近三年的研究，操弄 SOA，觀察了促發字對目標字的辨識與念讀之影響效果，闡明了比 Grainger 等的研究更穩定的漢字形似近鄰抑制效果(Wu & Chen, 2003)，當 SOA 增加到足夠長，使較低頻的形似近鄰也產生抑制效果時，較高頻的形似抑制效果絲毫沒有減弱的現象，作者以漢字沒有遭遇拼音文字形音共變問題的特性來解釋漢字與法文研究結果之間的差異。此外，值得一提的是，更進一步的複驗實驗之初步結果顯示，在更短的 SOA(34ms)下，也仍然觀察到穩定的形似抑制效果，而非如同法文研究的形似增快效果，這表示漢字比 Grainger 所關心的法文單字(小於 5 個字母長度)有更快的認出速度，除非能把 SOA 操弄到更短的 17ms (以目前的大多數實驗室條件下，這在技術上是仍不穩妥的)，我們仍然看不到穩定的全字形似促發增快效果，意味我們找不到適當的時間點足以提供如同法文一樣可以觀察到字內部件處理的時間歷程。既然找不到可以操作的時間點觀察全字形似促發增快效果，Wu 及 Chen 只好採用類似 d'Arcais, Saito, & Kawakami(1995)的實驗程序，增加單獨呈現字內部件作為促發項目的情況進行部件促發的模擬實驗，這時，目標字的右半成份部件促發項會對目標字的辨識產生增快效果，而對目標字的唸讀產生抑制效果。這一類逐漸累積的研究結果慢慢的襯托出許多未經深思熟慮而貿然追隨拼音文字理論所急就章產出的實驗結果之矛盾局面，面對形似抑制現象的被穩定闡明，過去的很多促發程序實驗之研究者，將會很難說明他們為什麼從來都沒有注意到形似促發抑制的可能性，現在又將如何重新解釋他們過去得到的結果，特別是過去強烈支持漢字辨識歷程一定涉入語音處理歷程觀點的研究者。

Grainger 等人(Ferrand & Grainger, 1992; 1993; 1994; Grainger, 1990)針對法文及荷蘭文的拼音文字所進行的這一系列字彙辨識研究，闡明了拼音文字字彙辨識過程裡的辨識前語音激發時間歷程，這些顯示時間歷程的研究結果支持了長期以來拼音文字辨識議題裡的兩項重要主張： 1). 在還沒有完成單字辨識前，就顯示了語音處理歷程的跡象； 2). 這種形似或語音相似的促進效果自然是表現在部分字母串的處理之上，亦即因為包含目標字中部分字串的促發項目可以增快或促進目標字的辨識或唸名，所以因此推論字彙辨識前的過程裡有處理部份字串的次字彙(sub-lexical process)歷程。如果 SOA 夠長，使得受試者已認出整個單字，那麼形似抑制的效果就會出現(如 Grainger & Ferrand, 1994; Segui & Grainger, 1990)。這樣的研究發現對過去的累積文獻有畫龍點睛的效果。

過去，探討漢字辨識歷程的研究者也多師法拼音文字研究的類似方法與「相同」觀點，多數的研究者一直企圖找出證據支持這兩個主張，這變成最近十數年來漢字研究的兩個主要研究方向，已如前所示，其一為主張漢字辨識前或語意提取前存在的語音處理歷程，或稱語音轉錄，如 Cheng (1992), Cheng & Shih (1988), Fang, Horng, & Tzeng (1986), Hue (1992), Tan, Hoosain, & Peng (1995), Tzeng &

Hung (1980), Tzeng, Hung, & Wang (1977)等的研究，這些都是 Perfetti 與 Zhang (1995)準備提倡普遍語音法則(universal principle of phonology)時所引用的漢字研究論文。以及與其實質上可能也相互有關的另一主張，就是字彙辨識前的成分部件處理歷程，如 Flores d'Arcais, Saito, & Kawakami (1995)的日文漢字研究、Feldman & Siok (1997)、Peng & Wang (1997)、Taft & Zhu (1997)、Feldman & Siok (1999)、Li & Chen (1999)、Li (1999)、Taft, Zhu, & Peng (1999)、Zhou & Marslen-Wilson (1999)等的研究。值得注意的是，看起來很大部分的知名漢字認知研究者似乎都接受拼音文字研究所得的結果，直接運用在漢字研究上，而不太用非常慎重及嚴格的標準檢驗這些漢字研究所得的結論。關於這方面的討論可以由 Liu(2003)的論文裡清楚的看到這樣的研究態勢與評論。

如前文所述，Wu 與合作者(Wu & Chen, 2000; Wu & Chou, 2000)曾質疑 Perfetti 等(Perfetti & Zhang, 1991; Perfetti & Tan, 1998; Tan & Perfetti, 1997)的研究，並以更周延的設計反駁了後者對漢字語義提取前存在語音處理的主張，香港及中國大陸的後續研究者(Chen & Shu, 2001)更以與 Perfetti & Tan (1998)完全同樣的材料與設計再行檢驗，也得到與 Wu 等類似的結果。自此，漢字辨識前或語意提取前存在語音處理歷程的主張就失去了最強有力的支持證據，其他的實驗證據如 Cheng(1992)、Shih & Cheng (1988)、Fang 等(1986)、Hue (1992)等等也都存在著難以說服人的矛盾之處，筆者計畫留待以後為文仔細交代。

一如前面提及的 Wu & Chen (2003)論文，Wu 等從 Grainger 等的法文研究取得靈感，以更周延的方式針對台灣的繁體漢字也進行了系列的字彙辨識及唸字研究，同樣採用促發程序，操弄形似變項，相對字頻及促發項呈現時間(SOA)，但在同樣短的 SOA 內，卻沒有觀察到類似法文拼音文字實驗的字彙辨識前部件處理時間歷程。這樣一來，似乎蘊含著漢字辨識前存在可觀察到部件處理歷程的說法或證據，已經面臨極為嚴酷的挑戰。顯然，過去主張漢字辨識前存在可觀察到部件處理歷程說法的支持證據或實驗必須重估並接受嚴格檢驗了。這是本三年計畫方案的研究目標。

本研究計畫目標

根據前文的剖析及過去的研究經驗，基於以下一些理由，本研究計畫擬針對繁體字辨識的研究議題提出一個三年計畫，詳細探討部件處理在單字辨識歷程裡可能扮演的角色。

1. 部件處理在單字辨識歷程裡扮演的角色之議題目前仍無定論，但大體看起來很像拼音文字理論進展的影子。就如同對語音歷程的觀點之探討一樣，研究

者很容易一開始就沿用拼音文字研究者的理論模型或觀點作為類似預設立場進行探討，為了發表論文方便，經常過度重視相同之處卻輕易忽略相異之處，造成後來研究者在理論推進上的不便。嚴重者甚至成為一種學術上不太負責的示範。

- 2.目前探討漢字認知歷程研究的論文及研究者已有大部分採用簡體字進行實驗的趨勢，採用繁體漢字進行認知心理學實驗的研究者人數益形減少。但繁體字的部件與全字關係，有可能與簡化了的簡體字有些不同，雖然目前仍未有人闡明。台灣差不多已是繁體字目前存在世界上的最後一片淨土，繁體字不只是歷史的見證，也可以說是祖先留下來的，過去現在與將來都能夠代表台灣的獨特文字，十多年前，漢字認知歷程研究都是台灣出身的研究者在拼音文字的世界裡以祖先留下來的繁體字遺產闖出的天下，可惜當年在這方面創立名號的很多前輩高人多已不再繼續發揚光大或綿延當時的工作，導致厚實的基礎日漸不足，後輩的接手委實已有些困難，實在令人唏噓。慢慢的，台灣的文字難道要在世界的舞台上逐漸喪失發言權嗎。我們不做，將來就不會有人替臺灣的文字講話了。
- 3.當我們獨創的掌握了繁體字的強固形似抑制特性及相關實驗技術之後，我們有別的研究者所沒有的獨特經驗及視野進入大家湧入的競技場，過去，我們曾經不正式的嘗試了進行這方面的實驗，也有了一些初步的經驗(參考 Wu & Liu, 1997)，這次，我們將正式帶著裝備介入。

研究方法、進行步驟及執行進度。

第一年

主要工作是繁體漢字資料庫細節功能的增添。過去的十多年，計畫申請人發展出配合漢字認知心理學研究所需的電腦化字詞屬性資料庫，包括出現頻率、單字及詞層次的屬性等等。第一年將完整的將該資料庫延伸至包括單字成份部件屬性，臚列工作細目如下：

- 1.分析單字成份部件的語義、語音屬性。
- 2.分析部件屬性與全字字形、語義、語音之間的關聯。
- 3.計算包含同一部件的形似近鄰字群屬性、語音一致性、語義關聯性、猜測全字

讀音之訊息不確定量等，以供正式從事實驗時進行獨變項操弄之用。

- 4.開始設計從資料庫裡挑選刺激材料的演算程式，同時進行不正式的先遣實驗。

第一年研究的目前進展情況及部份結果摘述於後文。

第二年

進行正式的實驗。參考目前這個議題的文獻進展情況及研究者累積的個人研究經驗及執行效率之考量，將通盤性的涵蓋如下的實驗規劃藍圖：

1.主要的實驗作業將包括辨字與唸名作業，也會包括全字促發與單獨部件促發程序。

2.可能考慮操弄的獨變項如下：

部件層次屬性

部件類型:部首或聲旁

type frequency vs token frequency

部首之表義程度

聲旁之表音程度

其他相關屬性

單字層次屬性

出現頻率

字義明確 vs 字義模糊

部件與單字關係

部件與全字是否語義關聯

語音規則性

時間歷程的操弄

SOA，如技術上能克服其他伴隨問題，就包括最短為 17ms 的極限情況。

3.近程實驗

檢驗現有文獻裡較重要的全字促發實驗。由於技術上的難度，現有文獻裡的大部分促發實驗都採用全字促發程序，如前文所述，全字促發程序與單獨部件促發程序會造成不一樣效果，理論解釋也可能不同。因此，這些實驗的檢驗會考慮參酌添加單獨部件促發程序。

4.根據過去經驗，大約一年可以進行四至六個實驗設計比較複雜或需求受試者人數較多的實驗。申請人將在這樣的條件考量下，以實際可行的最有效率方式，將上述獨變項的操弄盡可能配合問題的需要融入四至六個實驗之設計中。為了考量論文內容的獨創性以利發表，如果第一年的部件屬性資料庫分析進行情況可以配合，也可能先延續 Wu & Liu (1997)的論文之後續工作，Wu 與 Liu (1997)的論文裡，受試間的比較三種唸名作業：平常的單字唸名作業，呈現可以單獨成字的部件之念名作業，及呈現單字但令受試針對其內含右半部件進行唸名的唸名作業。關注的焦點是受試內的比較單字頻率及部件頻率之相對影響效果，結果發現，相對於呈現單獨成字的部件之念名作業，及呈現單字但令受試針對其內含右半部件進行唸名的唸名作業而言，平

常的單字唸名作業都幾無受到部件出現頻率的影響。從另一個角度來看，呈現單字但令受試針對其內含右半部件進行唸名的唸名作業反而受到包含該部件之單字的出現頻率之干擾，單字出現頻率越高，影響越大。似乎顯示，單字的全字處理不太受到單字內部件的影響，但在單字之內企圖處理其中部件，卻一定會受到包含該部件之單字的影響。該實驗的程序面臨了一些未及處理周延的混淆因素，其中的一個問題是，由於當時技術上未能完成整個部件資料庫的計算，因此權宜的把部件單獨成字的單字出現頻率視為部件出現頻率，亦即 **token frequency** 與 **type frequency** 混用。當第一年的部件資料庫之計算有點眉目時，我們當然迫不及待的企圖馬上以更周延的方式處理 Wu & Liu (1997)的問題。

第三年

- 1.重復檢驗實驗的穩定性。
- 2.檢驗不同理論說法的較關鍵實驗。
- 3.因應前述所得結果的狀態，再行設計關鍵的後續實驗。

預期成果。

提出實質的發現及可以穩定重複的實驗結果，闡明繁體字或漢字辨認歷程裡的部件扮演角色。

三、第一年研究的目前進展情況大要

研究過程首先針對每個繁體漢字(共 13000 餘字)進行單字成分部件的拆解登錄,同時也判斷並登錄部首與全字的語義關聯程度,部首與全字語義關聯程度的判斷與登錄工作係由精通中文的大學畢業生研究助理分成三次從事,第一次是以「有關或無關」作為判斷標準,第二次是以從有關至無關之間進行主觀的五點量表評量,第三次是利用字典記載針對全字的意義再做更進一步的意義之分析,分別記錄各次結果。這些單字成分部件的拆解及部首與全字語義關聯程度的判斷等的工作都只能透過研究人員手工運作,然後再由研究者設計程式結合 Wu & Liu (1987)所發展的漢字詞出現頻率資料庫,進行重複運算,以求得部件的屬性資料庫。目前已完成絕大部分的程式設計工作,並開始執行程式以得到初步結果。由於前述單字成分部件的拆解及部首與全字語義關聯程度的判斷等的工作都係由研究助理逐字進行判斷並以人工操作鍵入,數量龐大(繁體漢字共約 13000 字),出錯難免,所以利用這種初步得到的結果,可以檢驗出單字成分部件的拆解登錄階段中仍然存在的錯誤並進行修正,也同時補足程式設計的除錯過程,之後再開始執行程式,如此反覆交互迭代進行,整個過程的錯誤就會逐漸減少。

目前所得到的部件資料庫雛型包含了像如下以部件「少」為例的各項屬性:

部件: 少

百萬單字內此部件之出現頻率: 1384

包含此部件之鄰字數: 20

步 沙 省 妙 吵 劣 抄 砂 紗 秒 雀 鈔 炒 眇 秒 鈔 鈔 鈔 鈔 鈔

不考慮音調之不同鄰字拼音及其百萬單字內出現頻率:

ㄅ	495	尸	320	尸	231	冂	168
彳	108	勹	40	勹	22		

跨鄰字平均拼音訊息量: 2.3493

百萬單字內字頻大於 0 之鄰字及其屬性

(依序包括單字、讀音、字頻、筆畫、同音之其他鄰字字頻和、不同音之其他鄰字字頻和、
鄰字同音優勝比=同音之其他鄰字字頻和/不同音之其他鄰字字頻和、
部件在單字中位置:2 上 3 下 4 左 5 右 6 中間 7 外圍 8 部件單獨成字)

步	ㄅ	495	7	0	889	0.0000	3
沙	尸	251	7	69	1064	0.0609	5
省	尸	231	9	0	1153	0.0000	2
妙	冂	134	7	34	1216	0.0272	5
吵	彳	46	7	62	1276	0.0463	5
劣	勹	40	6	0	1344	0.0000	2
抄	彳	35	7	73	1276	0.0541	5

砂	尸 彳	35	9	285	1064	0.2113	5
紗	尸 彳	34	10	286	1064	0.2119	5
秒	冫 一 彳 丿	32	9	136	1216	0.1006	5
雀	勹 口 世 丿	22	11	0	1362	0.0000	2
鈔	彳 幺	16	12	92	1276	0.0673	5
炒	彳 幺 丿	11	8	97	1276	0.0706	5
眇	冫 一 彳 丿	2	9	166	1216	0.1201	5

其中，鄰字平均拼音訊息量是一種測量指標，可以反應特定部件攜帶的訊息足以正確猜測包含該部件之單字發音的機率。這會是一種比 Fang 等(1986)所發展出的 consistency 更好且更穩定的測量指標，該 consistency 測量的概念其實面對了相當程度的邏輯困難。Fang 等(1986)的報告揭示了唸名作業裡的一致性效果，Hue(1992)採用了同樣測量指標及作業，並額外操弄了目標字的字頻，結果發現只有低頻字才出現一致性效果，高頻字並不出現一致性效果。根據 Hue(1992)的論文結果進行仔細思考，可以隱含 Fang 等(1986)的實驗結果是不能被穩定重複的推想，因為 Fang 等(1986)的實驗材料沒有處理字頻因子，理當絕大部分包括高頻字，如果 Hue(1992)的結果穩定，Fang 等(1986)就不會得到穩定的一致性效果。但 Hue(1992)的結果卻也是邏輯上無法周延的，因為 consistency 的計算指標並沒有考慮到所牽涉鄰居單字的個別出現頻率。包含特定聲旁的常見字與罕見字在 Fang 等(1986)與 Hue(1992)的 consistency 指標之計算過程裡的權重都被當成是一樣的，然而例如前述罕見字「眇」的讀音如果也包含在部件「少」的 consistency 計算裡，就顯然不一定能反映受試者認知系統裡的狀態。如果一個單字的發音需要考慮到近鄰字的發音，那麼低頻字的發音歷程裡理當考慮到較高頻鄰居字的讀音，這符合 Glushko(1979)的想法，但高頻字的發音如何會需要去算計到較低頻鄰居字的讀音呢？高頻字的發音既然還沒有算出，那麼低頻鄰居字的讀音又如何據以進行計算呢？這就是 Fang 等(1986)與 Hue(1992)所採用的一致性指標在概念上的弔詭之處。顯然，Hue(1992)在高頻字上不出現一致性效果也可能與這種一致性指標不能穩定的反映心理真實性有關。

繁體漢字超過 13000 字，其中大約有 8000 字是罕見字，大學生看過其中的一些部分，小學生眼中的罕見字數目更多，相對而言，看過的更少，按所謂 consistency 指標計算所根據的邏輯推演，大學生認知系統中的 consistency 顯然會大不同於小學生，兩者之間波動極大。反映文字整體表音特性的指標，照理不應有這麼大的不穩定影響因素，它應該是可以讓不同的文字之間，如繁體漢字與簡體漢字之間，或漢字與英文之間，進行形式上的公平比較才合理。本文擬採用的表音不確定性計算指標就可以滿足這種公平比較的要求，而且這種計算在心理學文獻裡也早就有極為穩固的說服力，茲摘述其根據如下：

根據 Carlson (1968)對訊息理論(information theory)的描述，可定義一個溝通系統內的訊息傳遞量，類似的作法也可以運用在行為的研究裡，對刺激出現機率與反應出現機率進行訊息量的量化定義(見 Luce, 1960)。依此，Coombs, Dawes, & Tversky (1970) 定義事件 x 的訊息量， I_x ，的計算條件如下：

1. 若 $p(x) = 1$ ，則沒有可傳遞的新訊息，所以訊息量(information)，或稱不確定性(uncertainty)，為 $I_x = 0$ 。
2. 若事件 x 出現機率小於事件 y 的出現機率，則 $I_x > I_y$ 。
3. 若兩事件互為獨立，則各事件之對應訊息量為加成的(additive)。

以此可定義出事件 x 之訊息量如 $I_x = -\log p(x)$ bits。若隨機變項 x 有多個不同變量 x_i ，亦可推演 x 的平均訊息量或平均不確定性，如

$$U(x) = -\sum_i p(x_i) \cdot \log p(x_i)。$$

同理，條件 A 下的 x 之條件平均訊息量或條件平均不確定性可表示如式 $U(x|A) = -\sum_i p(x_i|A) \cdot \log p(x_i|A)$ 。

因此，當面對一個包含部件「少」的不熟悉讀音之單字，可以有大約 2.35bits 訊息量的不確定性可以進行猜測，亦即大約有 5 個不考慮音調的可能不同讀音足供猜選。同樣的根據訊息理論計算，在該單字被限定於包含左右兩半形部件結構的條件下，得到其發音之條件平均不確定性為 1.44，亦即大約有 3 個不考慮音調的可能不同讀音足供挑選。我們可以用這種方式整個跨部件的通盤計算部件預測單字發音的表音能力，顯而易見的，如果連表義的部首都包括進來，中文的部件表音不確定性是很高的，亦即表音能力極低，遠遠的不如拼音文字，依直覺推想，是不值得讓繁體漢字，甚至簡體漢字，的讀者利用這種線索來發音的。就算同樣一套的漢字符號系統，以不同方言作為母語的使用者之發音，如北京話、廣東話、閩南語、客家語、蘇州話、日本、韓國、甚至已經消失的西夏等等，所各自據以計算出來的整體表音不確定性，也都顯然會遠大於拼音文字，而且，針對同一個特定漢字，不同方言之間也不會得到一樣的計算結果，這將清楚闡明漢字辨認後發音過程的特徵。詳細而正確的計算結果將在部件屬性資料庫全部完成之後提出來。

四、計畫成果自評

如前所示。

五、參考文獻

- Carlson, A. B. (1968). *Communication System: An Introduction to Signal and Noise in Electrical Communication*. New York: McGraw-Hill.
- Chen, B., & Peng, D. (2001). The time course of graphic, phonological and semantic information processing in Chinese character recognition: 1. [Chinese]. *Acta Psychologica Sinica*, 33, 1-6.
- Chen, H.C., & Shu, H. (2001). Lexical activation during the recognition of Chinese characters: Evidence against early phonological activation. *Psychonomic Bulletin & Review*, 8, 511-518.
- Cheng, C. M. (1992). Lexical access in Chinese: Evidence from automatic activation of phonological information. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 67-92). NorthHolland: Elsevier.
- Cheng, C. M., & Shih, S. I. (1988). The nature of lexical access in Chinese: Evidence from experiments on visual and phonological priming in lexical judgment. In I. M. Liu, H. C. Chen, & M. J. Chen (Eds.), *Cognitive aspects of the Chinese language* (pp. 1-14). Hong Kong: Asian Research Service.
- Cho, J. R., & Chen, H. C. (1998). Phonology and orthography in recognizing Hanja in skilled and less-skilled readers. [Korean]. *Korean Journal of Experimental & Cognitive Psychology*, 10, 17-35.
- Coombs, C. H., Dawes, R. M., & Tversky, A. (1970). *Mathematical Psychology: An Elementary Introduction*. Englewood Cliffs, N. J.: Prentice-Hall.
- Fang, S. P., Horng, R. Y., & Tzeng, O. J. L. (1986). Consistency effects in the Chinese character and pseudo-character naming tests. In H. S. R. Kao & R. Hoosain (Eds.), *Linguistics, psychology and the Chinese language* (pp. 11-21). Hong Kong, China: University of Hong Kong Press.
- Feldman, L. B., & Siok, W. W. T. (1997). The role of component function in visual recognition of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 776-781.
- Feldman, L. B., & Siok, W. W. T. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory & Language*, 40, 559-576.
- Ferrand, L., & Grainger, J. (1992). Phonology and orthography in visual word recognition: Evidence from masked non-word priming. *Quarterly Journal of Experimental Psychology: Human Experimental Psychology*, 45A(3), 353-372.
- Ferrand, L., & Grainger, J. (1993). The time course of orthographic and phonological code activation in the early phases of visual word recognition. *Bulletin of the Psychonomic Society*, 31, 119-122.

- Ferrand, L., & Grainger, J. (1994). Effects of orthography are independent of phonology in masked form priming. *The Quarterly Journal of Experimental Psychology*, *47*, 365-382.
- Flores d'Arcais, G. B., Saito, H., & Kawakami, M. (1995). Phonological and semantic activation in reading Kanji characters. *Journal of Experimental Psychology*, *21*, 34-42.
- Glushko, R. J. (1979). The organization and activation of orthographic knowledge in reading aloud. *Journal of Experimental Psychology: Human Perception and Performance*, *5*, 647-691.
- Grainger, J. (1990). Word frequency and neighborhood frequency effects in lexical decision and naming. *Journal of Memory and Language*, *29*, 228-244.
- Grainger, J., & Ferrand, L. (1994). Phonology and orthography in visual word recognition: Effect of mask homophone primes. *Journal of Memory and Language*, *33*, 218-233.
- Hue, C. W. (1992). Recognition processes in character naming. In H. C. Chen & O. J. L. Tzeng (Eds.), *Language processing in Chinese* (pp. 93-107). Amsterdam: North-Holland.
- Li, H. (1999). Recognition of Chinese characters: A radical-based approach. (character recognition, orthography, radicals). *Dissertation Abstracts International: Section B: the Sciences & Engineering*, *59*(11-B), 6086.
- Li, H., & Chen, H. C. (1999). Radical processing in Chinese character recognition: Evidence from lexical decision. *Psychologia: an International Journal of Psychology in the Orient*, *42*, 199-208.
- Liu, I. M. (2003). Introduction to Chinese character/word processing. *Chinese Journal of Psychology*, *45*, 1-9.
- Liu, I.M., Wu, J.T., & Chou, T.L. (1996). Encoding operation and transcoding as the major loci of the frequency effect. *Cognition*, *59*, 149-168.
- Luce, R. D. (1960). The theory of selective information and some of its behavioral applications. In R. D. Luce (Ed.), *Developments in Mathematical Psychology*. Glencoe, Ill.: The Free Press.
- Peng, D, & Wang, C. (1997). Basic processing unit of Chinese character recognition: Evidence from stroke number effect and radical number effect. [Chinese]. *Acta Psychologica Sinica*, *29*, 8-16.
- Perfetti, C. A., & Tan, L. H. (1998). The time course of graphic, phonological, and semantic activation in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 101-118.
- Perfetti, C. A., & Zhang, S. (1991). Phonological processes in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, and*

- Cognition*, 17, 633-643.
- Perfetti, C. A., & Zhang, S. (1995). Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 24-33.
- Perfetti, C. A., & Zhang, S. (1996). The universal word identification reflex. *The Psychology of Learning and Motivation*, 33, 159-189.
- Pollatsek, A., Tan, L. H., & Rayner, K. (2000). The role of phonological codes in integrating information across saccadic eye movements in Chinese character identification. *Journal of Experimental Psychology: Human Perception & Performance*, 26, 607-633.
- Segui, J., & Grainger, J. (1990). Priming word recognition with orthographic neighbors: Effects of relative prime-target frequency. *Journal of Experimental Psychology: Human Perception and Performance*, 16, 65-76.
- Shen, D. (1998). The role of phonological information in reading Chinese. *Dissertation Abstracts International: Section B: the Sciences & Engineering*, 58, 6254.
- Shen, D., & Forster, K. I. (1999). Masked phonological priming in reading Chinese words depends on the task. *Language & Cognitive Processes*, 14, 429-459.
- Spinks, J. A., Liu, Y., Perfetti, C. A., & Tan, L. H. (2000). Reading Chinese characters for meaning: the role of phonological information. *Cognition*, 76, B1-B11.
- Taft, M., & Zhu, X. (1997). Submorphemic processing in reading Chinese. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 761-775.
- Taft, M., Zhu, X., & Peng, D. (1999). Positional specificity of radicals in Chinese character recognition. *Journal of Memory & Language*, 40, 498-519.
- Tan, L. H., Hoosain, R., & Peng, D. L. (1995). Role of early presemantic phonological code in Chinese character identification. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 43-54.
- Tan, L. H., & Perfetti, C. A. (1997). Visual Chinese character recognition: Does phonological information mediate access to meaning? *Journal of Memory and Language*, 37, 41-57.
- Tzeng, O. J. L., & Hung, D. L. (1980). Reading in a nonalphabetic writing system: Some experimental studies. In J. F. Kavanagh & R. L. Venezky (Eds.), *Orthography, reading and dyslexia*. Baltimore: University Park Press.
- Tzeng, O. J. L., Hung, D. L., & Wang, W. S.-Y. (1977). Speech recoding in reading Chinese characters. *Journal of Experimental Psychology: Human Learning and Memory*, 3, 621-630.
- Weekes, B. S., Chen, M. J., & Lin, Y.-B. (1998). Differential effects of phonological priming on Chinese character recognition. *Reading & Writing*, 10, 201-222.

- Wu, J. T., & Chen, H. C. (2000). Evaluating semantic priming and homophonic priming in recognition and naming of Chinese characters. [in Chinese] *Chinese Journal of Psychology*, 42, 65-86.
- Wu, J. T., & Chen, H. C. (2003). Chinese orthographic priming in lexical decision and naming. *Chinese Journal of Psychology*, 45, 75-95.
- Wu, J. T., & Chou, T. L. (2000). The comparison of relative effects of semantic, homophonic, and graphic priming on Chinese character recognition and naming. *Acta Psychologica Sinica*, 32(Suppl.), 34-41.
- Wu, J. T., Chou, T. L., & Liu, I. M. (1994). The locus of the character/word frequency effect. [in Chinese] In H.W. Chang, J.T. Huang, C.W. Hue, & O.J.L. Tzeng(Eds.), *Advances in the study of Chinese language processing*. (Pp. 31-58). Taipei: National Taiwan University.
- Wu, J. T., & Liu, I. M. (1987). *Frequency counts of Chinese characters, words, and syllables*. National Science Council Technical Report: NSC75-0301-H002-024.
- Wu, J. T., & Liu, I. M. (1997). Phonological activation in pronouncing characters. In H.C. Chen (Ed.), *The Cognitive Processing of Chinese and Related Asian Languages*. (Pp. 47-64). The Chinese University Press.
- Zhang, S., Perfetti, C. A., & Yang, H. (1999). Whole word, frequency-general phonology in semantic processing of Chinese characters. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 25, 858-875.
- Zhou, X., & Marslen-Wilson, W. (1999). The nature of sublexical processing in reading Chinese characters. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, 25, 819-837.
- Ziegler, J. C., Tan, L. H., Perry, C., & Montant, M. (2000). Phonology matters: The phonological frequency effect in written Chinese. *Psychological Science*, 11, 234-238.