

《漢字語法規則》概略*

麥傑 (James Myers)

國立中正大學

<http://personal.ccu.edu.tw/~lngmyers/>

2019/7/11

(陳欣徽—譯 2020/1/6)

0. 前言

《漢字語法規則》(Myers, 2019) 顧名思義即闡述漢字具有一套在心理層次上相當活躍且高產的語法，與手語和口語的構詞規則和音韻規則有著許多驚人的相似之處。本文依序在每一小節中分別簡要說明《漢字語法規則》各章所提出的要點。第一節概述了過去漢字的分析結果，這些結果揭示了漢字的「類語法」特性。第二節針對漢字具有類似於詞綴化、複合詞化和重複等構詞機制加以論證。第三節則探討漢字也具有音韻規則（一種靜音形式的音韻規則，如手語），這種音韻規則描述了筆劃和整體文字符號形狀中的抽象形式規律。第四節則為前述規則的能產率提供語料庫的研究證據。第五節提出了實驗證據。第六節先探究漢字語法的可能解釋，接著說明其概念如何加以落實在理論語言學之外的領域。

1. 漢字語法：概念

在本節中，我將回顧漢字的本質與語法的本質，接著將兩者相互連結。**漢字**是一種在包含中文在內的漢語語言中，幾乎總是以單一語音詞素（亦即單一音節）呈現的書寫符號。這些符號在非漢語語言中，也經過了長時間的適應、改編和修正，例如日文的漢字、韓文和越南的喃字（請見 Handel, 2019）。在本文中，我主要將討論重點聚焦在「正體字/繁體字」的系統，也就是目前為止歷史最悠久的中文拼字系統，而現在僅使用於香港、澳門與臺灣等地。本文後面也會將「簡體字」的系統納入討論範圍。簡體字由中國開始發展，現在已廣泛使用於全世界其他國家，如新加坡。

儘管漢字的語言譜系相當古老，但漢字並未因此而形成一套固定文字的檢索表，新的漢字反而以成倍的速度持續增加，如圖 1 所示。漢字數量會持續成長的現象是有可能的，因為漢字是系統性構成的符號，且這種系統性的符號建構方式是深受至今仍有巨大影響的《說文解字》（西元前一百年）所認可的。在所有的漢字類型中，「形聲字」是目前最常見（超過 80%）、也是使用長達 3,000 年的類型（黃德寬，2003）。如表 1 所示，形聲字主要由一個帶有意義的「部首」和一個語音成分的「部件」所組成。第二個最重要的漢字類型是「會意字」，此類型的漢字字義與該漢字內部組成的部件成分有關。透過數個相同部件所組成的漢字之類型稱為「疊體字」，疊體字雖然在傳統分類上屬於會意字的一種，但這類漢字具有自己的有趣屬性，我們將會在後面的文章中回頭談論這類漢字。

* 如欲參考更多的相關細節，請參閱 Myers (2019) [www.routledge.com/9781138290815;http://personal.ccu.edu.tw/~lngmyers/CharGram.htm]。在此誠摯致謝各場講座觀眾、讀者、實驗室助理的各項協助與建議，以及科技部研究計畫的支持 (NSC 97-2410-H-194-067-MY3, NSC 97-2410-H-194-067-MY3, MOST 103-2410-H-194-119-MY3, MOST 106-2410-H-194-055-MY3)。

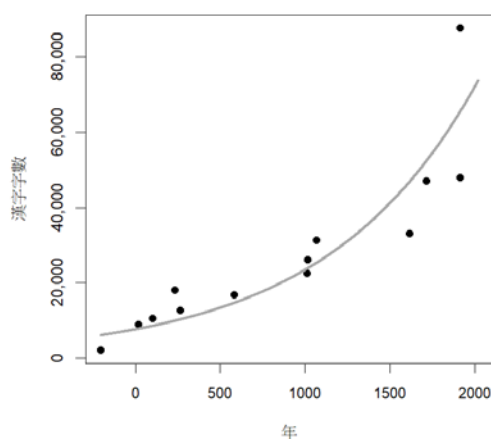


圖 1. 漢字字典規模 (資料來源：Yip, 2000, 第 19 頁)

表 1. 重要漢字類型與釋例

傳統分類	例子
形聲	媽 = 女 + 馬
會意	明 = 日 + 月
↳ 疊體字	林 (從「木」)

我們如何決定這些觀察結果 (以及本文後續更多的討論內容) 是否暗示漢字的形成存在一套真實的語法? 語言學家普遍期望一套語法應具有心理真實性、創造性和抽象性。例如, 我們之所以識別出聾人的自然手語具有語法, 是因為流利的手語使用者仰賴系統化的心理活動 (如 Emmorey, 2001), 他們的語言知識具有創造性, 有別於死記硬背的記憶方式 (如 Berent 與 Dupuis, 2018), 而這些知識的抽象性展現在三個方面: 第一, 手語非完全簡化為象似性的符號 (如 Frishberg, 1975); 第二, 手語涉及非模態的 (非單純在發音方法或語言感知層面) 心理表徵 (手語音韻研究已公認此現象; 如 Brentari, 2011); 第三, 手語不僅包含與其他認知截然不同的心理模組, 而且其本身是由各個分門別類劃分清楚的子模組所組成 (如 Padden 與 Perlmutter, 1987 在手語音韻學與手語構詞學的研究)。

書寫系統也同樣存在語法嗎? Bloomfield (1933, 第 21 頁) 最著名的一段話駁斥了這樣的說法: 「書寫系統不是語言, 而是一種僅止於透過視覺標記來紀錄語言的形式而已。」的確, 不像口語和手語, 閱讀與書寫相形之下是「不自然的」語言形式, 在這樣的前提下, 閱讀與書寫無法普遍通用, 也需要仰賴明確的教學引導。然而, 無法否認的是, 閱讀和書寫也喚起了真實的心理知識, 這些知識得以有效擴展那些經由教學引導的內容 (如 Pacton 等, 2001)。這樣的心理知識同樣也是抽象的、非象似性的 (意即語音或語義的半獨立狀態; Householder, 1971)、非模態的 (語文素養有賴於書寫與閱讀的連結; 如 Naka, 1998), 以及模組化 (視覺字形處理區 (Visual Word Form Area) 是大腦閱讀網絡的關鍵部位; Dehaene 與 Cohen, 2011)。

語言學家同樣也觀察語法分析工具應用於拼字系統的分析效果。例如希臘文的字素 (grapheme) /s/ 有兩個可預測的「同位字」(allographs): 字尾的 <ς> 和可置於任何位置的 <σ>。另外, 筆誤反映了一種超字素等級的心理表徵, 就像是音韻韻律 (prosody), 最常見的拼寫錯誤如把 <Philippines> (菲律賓) 寫成 <Phillipines>。如例 (1) 所示, 字母的書寫規則也會依循音韻規則和語音限制等方式相互替換 (根據 McCawley, 1994

的分析；也可參閱 Evertz, 2018)。

- (1) a. $y \rightarrow i / _ X$ carry ~ carrier spy ~ spies dry ~ dried
 b. *ii carry ~ carrying spy ~ spying dry ~ drying
 c. 最少三個字母的實詞 bee vs. be inn vs. in buy vs. by two vs. to
 d. 前三項的其他交互形式 die (*dy) ~ dying lie (*ly) ~ lying

因此，不出意料，許多研究者認為漢字同樣也具有語法規則。在這些充滿雄心壯志的研究中，Wang (1983) 採用了 Chomsky 與 Halle (1968) 的形式語言學作為理論架構，解釋漢字結構和其替換形式的各種觀察結果（其中許多部份可從我的研究證據中得到印證，將呈現於後面的文章中）。其他語言學家則採取更為量化統計的方式來分析漢字組成成分，以 Stalph (1989) 和 Kordek (2013) 最為經典。而 Sproat (2000) 在他的研究中將中文納入他的論點中，並且指出所有的書寫系統都具有「常規語法」(regular grammars)，其特色是可經由簡易鏈模型來加以表達。Ladd (2014) 奠基於先前的研究基礎上進一步指出漢字呈顯二元結構 (Hockett, 1960)。

本人採取的研究方法與 Wang (1983) 最為相近，但在理論基礎上我盡可能降低使用形式理論的比重，運用各種數據資料來支持每一個主張，並且遵循口語與手語的詞彙語法模型，採用彼此獨立但相互作用的三套模組，分別是：「漢字構詞」(character morphology，在意義或發音上具有詮釋功能的結構操作方式)、「漢字音韻」(character phonology，在無法詮釋的結構中所呈現的規律性，類似手語音韻的無聲形式)與「漢字語音」(character phonetics，心理生理影響程度的梯度現象)

2. 漢字構詞學

漢字的「構字語素」(character morphemes) 通常是顯而易見的，如同表 1 的例子所示，然而不是每一個漢字的構字語素都這麼容易被拆解出來。僅就繁體字的構字語素而言，在詞彙上具有辨義功能的構字成分數量其估算範圍不一，從 249 個 (Lui 等, 2010) 到 667 個 (Morioka, 2008) 等數量皆有。例 (2) 便顯示了一個內部構字成分出現語義模稜兩可的有趣現象（您需要用您敏銳的眼睛來看這個問題）。當然，口語與手語也具有類似這種語義含糊或具有歧義的準語素現象，像是廣為人知、與鼻音有關的/sn/，在英文字彙中發展出與鼻子有關的詞彙，如 sneeze (打噴嚏)、snore (打呼)、snot (鼻涕) 和 sneer (冷笑) 等。

- (2) 鬪 = 鬥 + 斲

漢字的生成方式和口語、手語相似，主要透過三種構詞機制而成：加綴、複合與重複。例 (3) 列出一些加綴機制的公認屬性，並且以英文 greenish 和 greenhouse 分別說明「加綴」與「複合」的構詞機制。

- (3) a. 黏著語素 (-ish 無法獨立成詞，而 house 可以獨立成詞)
 b. 封閉性詞類 (詞綴條目是固定的，而新的詞根語素則容易被創造出來)
 c. 語義淡化 (semantically bleached: Sweetser, 1988) (-ish 僅具有抽象功能，與 house 不同)
 d. 位置固定 (-ish 只出現在詞幹最後的位置，而 house 的位置可前可後，例如 houseboat 與 greenhouse)

- e. 語音形式弱化（*-ish* 在 *gréenish* 中為輕音節，而 *house* 在 *gréenhòuse* 中則為次重音節）

形聲字的語義部首也具有上述五種屬性。這些部首的條目是固定的（成書於西元 1716 年的《康熙字典》僅以 214 個符合當代傳統標準的部首作為分類標準）。而誠如例（4）所示，一些語義部首完全無法單獨出現（部分語義部首為「黏著形態變體」(bound allomorphs)，將於文後簡要討論）。部首同樣展現出語義淡化的現象，如例（5）所示。（貝殼在古代中國被作為貨幣使用；可參閱 Handel, 2018）。

(4) 丿 冫 灬 廴 彳

- (5) a. 貝 賬
b. 口 嗎

雖然語義部首的位置不完全固定，但這些部首的確具有強烈的位置優先性。圖 2 呈現出語義部首的位置優先性現象（本研究中的 6,607 字漢字主要來自於中研院現代漢語平衡語料庫[Chen, K. J.等, 1996]，並在維基共享資源「漢字分解」(Chinese Character Decomposition) 的幫助下解構漢字 [維基共享資源 Wikimedia Commons, 2017]；漢字類型的相關資訊主要以維基詞典, 2019 為主)。圖 2 清楚呈現出大多數的部首出現的位置並非毫無章法，而是具有特定傾向。特別值得注意的是圖中左側的黑色區塊，這個位置就是部首最常出現的位置。這樣的現象和口語的詞綴也會偏向某一側位置的現象類似（例如，英語中的後綴和納瓦霍語中的前綴；Matthews, 1991）。位置變化也是系統性的，像是位於左側位置的部首會以底部位置作為第二選擇（即圖中所見，黑色區域大部分被淺灰色覆蓋）。

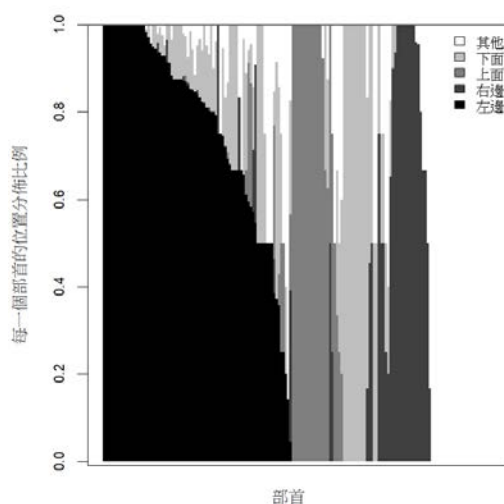


圖 2. 語義部首位置

最後，語義部首也呈現了「形式簡化」的現象，這可視為漢字音韻學的其中一個層面。有些簡化現象也會發生在非部首字上，只要這些非部首字出現在部首偏好的位置。例（6）列出了四種規律的形式簡化歷程，在每一行的第一個例子都包含一個語義部首。例（6d）的「延展法」指的是擴展右下方的筆畫（同時收縮左側的組成主體），這樣的形式簡化過程將隨後證明與漢字音韻學的某一方面是一致的。

- (6) a. 對角線法 (diagonalization) 土：型～地 工：紅～功
 b. 點散法 (dotting) 火：燙～爛 禾：鉢～和
 c. 收縮法 (shrinking) 雨～電 尚～當
 d. 延展法 (stretching) 風～颱 支～翅

除了上述這些規律的音韻歷程外，許多語義部首也經過「語素變體特異化」(idiosyncratic allomorphy) 的程序，這道程序也與特定位置相互關聯，如例(7)所示。此外，如表2所示，語素變體特異化也可能關乎到一個組成成分是否為詞綴(如，在形聲字中的語義部首)而非詞根(如，在語義複合字中)的判斷依據。誠如上述所見，左側位置應是語義部件的預設位置。

- (7) a. 左： 心：忘～忙
 b. 右： 刀：剪～刻
 c. 上： 艸～花
 d. 下： 火：燈～照

表2. 部首語素變體特異化在不同位置的比率(括號內為每一項的總數)

部首位置	形聲字	會意字	其他
左	46% (3086)	41% (126)	100% (3)
右	43% (289)	22% (60)	20% (10)
上	49% (578)	18% (114)	9% (44)
下	13% (410)	8% (125)	18% (44)

與加綴法相反，複合法反映的是無限制的串聯(Jackendoff, 2010)。構字方法所產生的結果包括在複合詞中位置依賴性的語素變體特異化不足，如上表2所示。語用學在此起了作用，如例(8)所示，漢字內組成成分的位置顯然以真實世界的狀態為參考依據。

- (8) a. 畚(水蓋過農田)
 b. 尿(水在身體下方)
 c. 杲(亮) 杳(暗)

最後，字符的重複在功能上與語義複合不同，關鍵在於複製本身具有語義，如例(9)所示，這樣的語義範圍與視覺所見的手語或口語的重複方法類似(Behr, 2006)。

- (9) a. 豐富化：多 品 蟲
 b. 強化： 晶 炎
 c. 弱化： 弱

字符重複也與手語和口語一樣具有一個顯著的形式特徵(McCarthy 與 Prince, 1998; Sandler 與 Lillo-Martin, 2006)：能夠重複的字符僅限於少數特定的字形，如表3所示。單一字符重複四次所形成的2×2方塊字是最稀少的類型，此類型的發生應是從原本的結構透過橫向複製或是縱向複製而成，仍在辯證中(可參考「爾器毘」等字)。線性三連體(如「三川靈龕」等字)只會同時出現在漢字組成成分內(意即，這樣的現象完全不具有構詞性)。

表 3. 漢字字符重複類型表

字形	計算 (%)	例子
橫向	47 (43%)	比林單朋雙
三角	39 (35%)	品森蟲晶轟
縱向	20 (18%)	多哥肉昌炎
方塊	4 (4%)	彘犴焮翳

3. 漢字音韻學與漢字語音學

漢字音韻學有別於漢字語音學，並且遠遠超過模式二元性的基本事實（存於解釋字符結構的規則性）。「漢字韻律」（character prosody）或整體漢字結構所扮演的中心角色證實這些要點。該結構的「韻律模板」（prosodic template）同時具有橫向與縱向的二元性，其中左側與頂部位置為弱（有助於簡化），而唯一強勢的位置（有助於顯著化）則是右下方。這些參數相互作用的結果所形成的結構如例(10)所示（沿垂直軸對稱顯示）。除了那些歸因於自身二維性質所形成的獨特屬性外，這套模板在形式上和功能上都與口語和手語的韻律組成成分相似，尤其是在韻腳上更是如此。

$$(10) \begin{bmatrix} W & \\ W & S \end{bmatrix}$$

不同漢字的模板式分析結果如例(11)所示。例(11a-b)呈現遞歸現象，例(11c)呈現左邊和頂部成份的位置分析，而例(11d)則顯示「例外」的韻律結構，這些結構是少數幾個在右邊或底部位置簡化的語義部首（參見：可將英文規律重音的 *búttón* 與例外的英文字尾重音字 *batón* 相比）。

$$(11) \begin{array}{lll} \text{a.} & \text{甫} [S] & \text{浦} [W \ S] & \text{蒲} \begin{bmatrix} W \\ [W \ S]_s \end{bmatrix} \\ \text{b.} & \text{烏} [S] & \text{寫} \begin{bmatrix} W \\ S \end{bmatrix} & \text{瀉} \begin{bmatrix} W \\ [W \\ S]_s \end{bmatrix} \\ \text{c.} & \text{厭} & \text{居} & \text{底} & \text{扉} & \text{痛} & \text{處} & \begin{bmatrix} W \\ W \\ S \end{bmatrix} \\ \text{d.} & \text{刻} [S \ W] & \text{或} [S] \ W & \text{熱} \begin{bmatrix} S \\ W \end{bmatrix} & \text{或} \begin{bmatrix} S \\ W \end{bmatrix} \end{array}$$

因此，韻律模板有助於解釋何以規律的簡化現象只會發生在左邊（對角線法和點散法）和頂部（收縮法），且這兩個位置都是弱勢的位置。漢字韻律甚至可對應到「音重」的概念（weight，類似吸引重音的音節帶有長母音）：如果沿著橫軸的高位組成成分都是輕音節，而沿著縱軸的成分都是重音節，那麼我們就可以解釋為何這些成分偏好在左邊位置（弱 **Weak**）或底部位置（強 **Strong**）（可參閱前文的圖 2）。圖 3 展示對語義部首「言」的分析，如在「說」和「警」這兩個漢字的不同位置。韻律模板顯然也能充分解釋重複字形的現象：三角形重複填滿了整個模板，而橫向和縱向的重複則填充其中一部分。

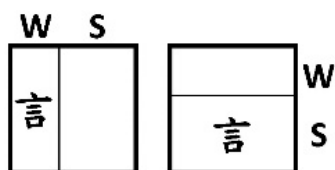


圖 3. 部首的維度和位置之間的關係圖

漢字韻律同樣也在筆畫中扮演相當重要的角色。如同平常看到的基本筆畫清單（如 Wang, 1983；Peng, 2017），可以被解構為一小組參數，如同口語和手語的音韻規則。這些參數可以描述線性方向，如表 4 所示。有些則時加工筆畫，像是彎曲或上鉤，如表 5 所示。複合筆畫則依照基本筆畫的順序加以分析，如表 6 所示。

表 4. 基本筆畫的線性方向參數

筆畫	軸線	方向	例子
丶	無	(預設)	太
一	橫向	(預設)	十
丨	縱向	(預設)	十
㇇	正對角線	(預設)	木
㇆	反對角線	向左落下	千才
㇇	反對角線	向右提起	子

筆畫圖示擷取自 https://commons.wikimedia.org/wiki/Category:CJK_strokes

表 5. 基本筆畫的加工參數

筆畫	軸線	彎曲	上鉤	例子
丿	縱向	是	否	川
𠃉	縱向	否	向左	丁
レ	縱向	否	向右	艮
フ	橫向	否	向下	ハ
㇇	正對角線	是	向上	戈
㇆	縱向	是	向左	彳

筆畫圖示來源與表 4 相同。

表 6. 基本筆畫在複合筆畫中的順序

複合筆畫	基本筆畫	例子
𠃉	一 丿	刀
厶	ノ 一	公
㇇	フ 丿	𠃉
㇇	丨 一 丿	弓
㇇	一 丨 一 丿	乃

筆畫圖示來源與表 4 相同。

筆畫反映韻律結構最明顯的方式是「特徵突出」(prominence) 上，也就是筆畫或筆畫群在強勢位置中放大。例 (12) 中的例子連同相關的韻律模板一起呈現。

- (12) a. 底部： 昌多哥 二土工手車里羊 $\begin{bmatrix} W \\ S \end{bmatrix}$
 b. 右邊： 珙林比炊林 川州 $\begin{bmatrix} W \\ S \end{bmatrix}$

延展法主要是縮小左邊的語義部首，同時延展右下方的筆畫。而韻律模板也有助於解釋延展法的過程，如例 (13) 的分析。

(13) 走~起 [W {S}]

即使當橫向筆畫的突出特徵異常置於底部以外的位置，它也符合另一個韻律規則：避免「衝突」(clash)，也就是當一個較弱的突出特徵與較強的突出特徵相接時，較弱的一方會不知去向而丟失（如 *refórm*、*rèformátion* 這兩個字，重音移位是為了避免衝突，這樣的情況和 *elícit*、*elicitátion* 相反）。因此在例 (14a) 中，我們可以看到異常置於突出特徵位置的筆畫會落在最底部的筆畫，且維持短的長度。但在 (14b) 中，最底部的筆畫較長是因為異常較長的筆畫並沒有落在最底部的筆畫。

- (14) a. 士 末 壬 華
b. 聿 幸 重 垂 事

漢字韻律在筆畫字形的角色之進一步證據來自縱向「筆畫彎曲」(curving) 的分布狀態。這類筆畫只會出現在左邊的漢字組成成分，意即處於一個弱勢位置，如 (15a) 左邊彎曲筆畫和 (15b) 非左邊之直線縱向筆畫的對比。

- (15) a. 川介升非拜弗爪月周角用 片爿 大九刀力虎戶底右看
b. 十中木不平下車年市 耳斗

正如 Wang (1983, 第 203 頁到第 206 頁) 首次觀察到但未解釋的現象，彎曲在寬距的漢字組成成分中是不受青睞的。根據此現象於表 7 中加以定量化，資料呈現出直線筆畫的變體「冂」更有可能出現在寬距的漢字組成成分中（如粗體字所標註之漢字）。這些觀察支持例 (16) 的韻律分析方式，其結果顯示彎曲筆畫在強勢位置上受阻的。

表 7. 彎曲屬性與各漢字組成成分維度的關係

		主軸線		
筆畫字形	橫向	縱向	無	
彎曲		月 甩 周 有 舟 角	丹 用	
直線	冊 冂 岡 巾 內 向 兩 肉 市 冫	再 甬 高 商 喬	同 冂 冉 束	

- (16) a. 高、細的組成成分： 月 [W S]
b. 分開的筆畫群： 門 [S][S]
c. 寬距的組成成分： 冊 [S][S]

然而，並非所有筆畫規則都有韻律的動機因子。事實上，比起音韻學，豎提在本質上看起來更像是例子驅動的類比結果，因為它通常只出現在令人出奇的特定環境，也就是直接在交叉筆畫的左邊，如例 (17) 所示。

- (17) 氏 民 長 艮 良 衣 喪 辰 派

豎鉤更顯得較少產出，如例 (18) 中類似最小對立體的例子。

- (18) a. 于 丁 事 乎
b. 干 下 聿 平

然而，如同 Wang (1983, 第 206 頁到第 210 頁) 首次非正式觀察到的現象，它仍然符合一組成對的統計通則：豎鉤在非對稱的漢字組成中更為常見，在不對稱的組成成分中，目標筆畫會接觸頂部位置。根據此現象於表 8 中加以定量化，表格中的資料將具有中心縱向筆畫的組成成分進行交叉分類（請注意粗體字）。

表 8. 不對稱與頂部接觸在豎鉤筆畫的角色

		不對稱	對稱
上鉤	頂部接觸	17 (了...)	5 (丁...)
	非頂部接觸	9 (事...)	1 (小)
非上鉤	頂部接觸	20 (下...)	18 (平...)
	非頂部接觸	26 (牛...)	28 (中...)

完全生產力的缺乏在詞彙語法中是可以被預期的，而且事實上漢字音韻學呈現出字構音韻學 (lexical phonology) 的典型特性 (Hargus 與 Kaisse, 1993)。如我們所見，大多數的規則不僅具有詞彙例外，也無法創造新穎結構 (即結構保存 [structure-preserving] 的原則)，如例 (19) 所示。

- (19) a. 對角線法：子~孩 子 vs. 子
 b. 特徵突出：日~昌 日 vs. 日

漢字音韻學對漢字構詞學也相當敏感，因為它大多數的通則都可以應用在內部詞素 (在組成成分當中)，如例 (20) 所示。

- (20) a. 在組成成分邊線的特徵突出：圭 (參見 土) 順 (參見 川)
 b. 在組成成分邊線的彎曲部分：所 淵 (參見 片月)

其他的音韻模式對相互作用的構詞機制類型同樣敏感。正如我們前面所看到的分析，語素變體特異化實際上只以加綴法 (語義部首) 為目標，而非組成成分的類型。語素變體特異化也不能應用在重複法上，如例 (21) 所示。相比之下，規律的簡化法會影響組成成分的各種類型，包含重複的元素，如例 (22) 所示。

- (21) 火：炎 (參見 灬：熱)
 水：淼 (參見 氵：江)
 手：龠 (參見 扌：拾)

- (22) a. 對角線法：珏 竝 銓 孛 鑫 弄
 b. 點散法：林 炊 林 𣎵 焱 森

漢字音韻學的詞彙制約證明漢字音韻學無法完全簡化為漢字語音學。更廣泛來說，漢字音韻學不僅是生理上的：漢字音韻學的範式還保留了機械化式所產出的字型，且大多數是強制性的。即使是左撇子，漢字音韻學的規則也可以規律性地應用 (例如，筆畫在左邊精確彎曲)。我們可以說，至少就歷時性的角度而言，漢字語音學「啟發」了漢字音韻學。特別的是，位在左邊的漢字組成成分會寫在右邊成分之前，而每一個漢字組成成分本身的書寫方向也都是從左到右、從上到下 (大致而言是如此；請參見下文)。在書寫時，需要將書寫工具從第一個組成成分的右下方移動到下一個組成成分的左上方，

因此點散法和對角線法對於簡化書寫距離上具有一定的影響。然而，雖然例（23a）的最後一個筆畫是縱向筆畫，但同樣是最後一劃，例（23b）的筆畫卻是從右上到左下的對角線法完成。換句話說，正如我們預期的一樣，語音與音韻之間具有共時關係，且筆畫順序同步化取決於對角線法，而非恰好相反。

- (23) a. 牛
b. 物

歷時的語音動機因子也可以在特徵突出化、彎曲、上鉤和韻律模版等部分被觀察到，但這當中沒有一個能夠就共時漢字音韻學上提出令人滿意的解釋，漢字音韻學保留了自身抽象的本質。然而，漢字語音學還是很值得單獨研究的，因為漢字語音學代表了筆畫方向和筆畫順序的領域，這也是在漢字研究文獻中被廣泛討論的議題。Wang (1983) 甚至將筆順放在他的分析系統的中心。然而，因為構成這些驅動力的基礎是源自於心理與生理，所以筆順的分析應該置於語法之外。

對於筆畫方向與筆順的主要影響是書寫過程，尤其是以右手持書寫工具寫字的實際活動。拉動書寫工具比推動書寫工具容易，這也是為什麼漢字「丁」和羅馬字「T」在書寫方向和方式完全一樣：亦即從左到右、從上到下。約定俗成的漢字筆順也會簡化漢字整體的字距（林桂生，2014），如圖 4 所示。



圖 4. 「正」字的筆畫慣性方向與筆順

當可替代的順序具有相似的書寫方便程度時，筆順變體就會隨之產生。依照中文的書寫慣例，通常會先書寫位在交叉位置的橫向筆畫，但是依照西方的書寫慣例，包含美國，通常都會先寫縱向筆畫。如圖 5 所示，兩種選擇都產出一樣短的書寫路徑。Goodnow 與 Levine (1973) 追蹤美國跨年齡書寫習慣的發展歷程，他們發現最年輕的書寫者（四歲五個月）幾乎沒有一個筆順凌駕於另一個筆畫之上的書寫偏好。

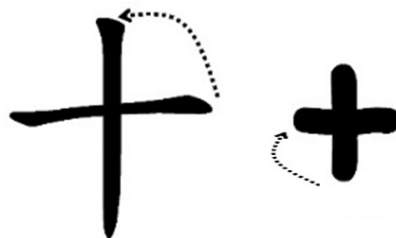


圖 5. 兩種書寫十字的同等效益選擇

在漢字書寫中也可以看到一樣的變化。例（24）的漢字在繁體字與簡體字系統中都是相同的，但是根據張小衡與張煥淇（2013）的研究發現，在台灣和中國學習中文的孩童與成人學習者所學習到的筆順卻不相同：在台灣，通常在寫十字的時候，學到的順序是先寫橫向筆畫、再寫縱向筆畫，但是在中國學到的筆順則剛好相反。

(24) 里 黑 冉 重

筆順變體也可以從書寫方式和感知限制間的競合狀態而產生。根據張小衡與張煥淇（2013），在例（25）中兩種組成成分的筆順都遵守中國的筆順對稱原則（點筆畫最後寫）。但是依照台灣的筆畫順序，例（25b）組成成分的筆順（此為語義部首「心」的語素變體）則是遵循書寫方式誘發的從左到右的順序。

- (25) a. 小
b. 忄

由於語音學主要在於實現而非塑造詞彙的表達形式，筆順也可能隨著音韻或形態結構的字符不同而有所改變。Wang（1983，第139頁）對於例（26a-b）中組成成分順序的分析方法格外令人注目：在例（26a）的漢字中，位於左邊和底部的語義部首是最後寫的，而例（26b）的漢字，則是先寫左邊和底部的語義部首。這是由於只有第二組的部首是經由延展法的方式，從例（26c）中的預設字形由左邊和底部加以延伸而來。因此例（26a）的組成成分順序是透過視覺佈局（空間填充元素優先）的因素來激發的，而例（26b）則是因書寫方式的實際活動而激發的。

- (26) a. 建 道
b. 趕 麵 颱 魅
c. 走 麥 風 鬼

4. 以語料庫為本的漢字語法證據

語料庫的量化與質化分析提供漢字語法能創性的額外證據。例如，圖（6）呈現兩種根據中研院語料庫（Chen, K. J.等，1996）統計後的能創性模型，左方模型比較形聲字、會意字和其他漢字類型的關係，右方模型則比較單邊語義部首在形聲字的不同位置。每一張圖都呈現出已揭示的新漢字（類型）的增加狀況，如同更多漢字形符在語料庫中的採樣一樣，產生成長曲線。圖中的黑線代表觀察性資料，其後延伸出去的虛線和短劃線則採用齊夫定律（Zipf's law）詞頻分佈的假設（Evert 與 Baroni，2007）。成長曲線顯示形聲字和語義部首在左邊位置的漢字不僅最普遍，而且最有能產力，如預估的斜率所示將遠超過實際的語料庫規模。

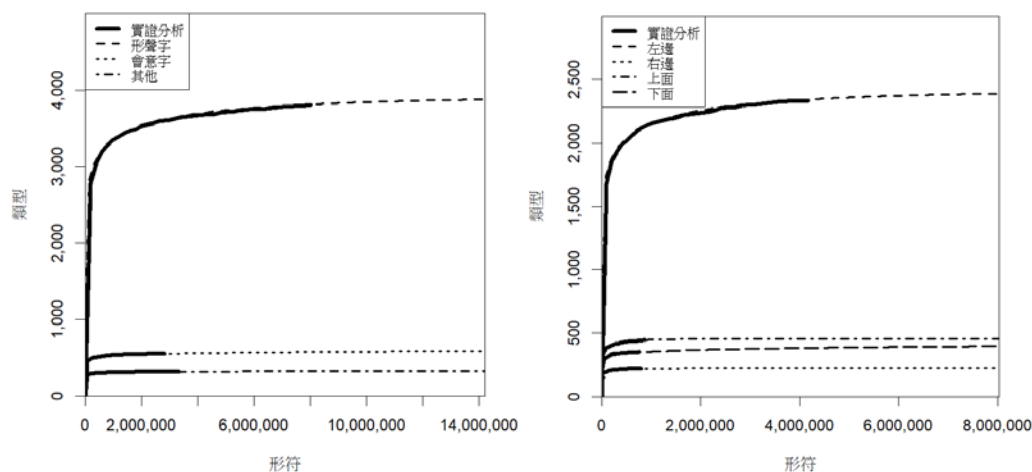


圖 6. 多種漢字類型的相對能創性

我們可以從遙遠的小篆（全盛時期約於西元前 200 年）開始，進一步透過擴展當代繁體字以外的語料庫來探討能創性。漢字構詞學從那時候開始有了些微變化，如例(27a)所示，雖然當時小篆還沒出現語素變體特異化的現象（X > Y 表示 Y 從 X 經由歷史衍生而來）。三橫畫重複的現象並未因此而不合語法，如例(27b)所示，儘管這樣的現象已經很稀少，而且一些其他在更古老的中文文字中已證實的重複字形也早已消失了。

- (27) a. 𠄎 ~ 悟 > 心 ~ 悟 𠄎 ~ 泥 > 水 ~ 泥
b. 𠄎 > 眾

在小篆的漢字音韻學中，除了一些頂部邊緣有收縮的現象外，並沒有規律簡化的現象，如例(28a)所示。在這些語料中，也沒有底部邊緣或右側邊緣突出化的現象，如例(28b)所示，而在這兩種邊緣的彎曲或者不存在、或者對稱出現在兩邊，並非侷限於左側邊緣，如例(28c)所示。韻律音韻學在現代漢字中是構成部首簡化和筆畫字形的基礎，不過綜觀所有觀察結果，韻律音韻學似乎尚未發展。

- (28) a. 𠄎 ~ 根 > 木 ~ 根 土 ~ 地 > 土 ~ 地 走 ~ 起 竹 ~ 筆
b. 𠄎 > 未 未 > 末 土 > 土 土 > 士 I > 工 川 > 川
c. 𠄎 > 并 井 > 井 周 > 周 同 > 同

我們也可以擴展我們的語料庫，讓語料庫的資料包含中國的現代簡體字系統（部分字體是藉由恢復可替代與年代久遠的字形而創造的，有些則是書法傳統）。一般而言，簡體字繼續往規模更大的構詞系統性的趨勢而發展，最為明顯的就是替代許多具有預設形聲字結構的漢字，如例(29)（此處 X < Y 代表 Y 被 X 替代）。具有位置依賴性的語素變體特異化已經延伸出另外的部首，如例(30)所示。字形簡化的策略同樣應用到重複法上，例(31)顯示這種構詞機制依舊保持漢字的能創力。

- (29) 听 < 聽 响 < 響 体 < 體 肤 < 膚 惊 < 驚

- (30) 词 ~ 警 < 詞 ~ 警 铅 ~ 鑒 < 鉛 ~ 鑒 红 ~ 紫 < 紅 ~ 紫

- (31) a. 保留原本的構字模板/基礎或僅基礎：𧈧 < 羸 虫 < 蟲 齿 < 齒
b. 大略保留原本的構字模板：𧈧 < 棗 双 < 雙 聶 < 聶 轰 < 轟
c. 應用到新例的模板：众 < 眾 宫 < 宮 网 < 網

然而，如同小篆一樣，簡體字的漢字音韻學與繁體字的漢字音韻系統略有不同。例(32a)的突出化延展更為明顯，而彎曲/筆直的對比度則忽略組成成分的寬度，如例(32b)所示。如例(32c)所示，向右上鉤不再依賴向右的交叉筆畫，雖然這與 Wang (1983) 的歸納一致，但當組成成分呈現非對稱的狀態時，有時候會促發向左上鉤，如例(32d)所示。然而，漢字語音系統仍然依賴漢字音韻系統且保持同步，當最低的筆畫被對角線化後，在例(32e)的簡體字組成成分會反轉筆畫順序。

- (32) a. 来 < 來 金 < 僉
b. 贝 < 貝 门 < 門 风 < 風
c. 长 < 長

- d. 东 < 東 (參見 书 < 書 车 < 車)
 e. 车~较 < 車~較

漢字語法也有助於理解不同字型之間的差異。如例(33)所見，比起黑體字，如毛筆書寫的楷書更容易應用在語素變體特異化的字體。

- (33) a. 糸~絲 示~神~禮
 b. 糸~絲 示~神~禮

字型也會在構詞分解上有所不同。如例(34a)顯示在常規字體中沒有出現突出部分接連而相互抵觸或不協調的狀況，但是在例(34b)的宋體字型卻出現不協調的情況，彷彿後方所接連的突出筆畫是在互不相干的組成成分裡。

- (34) a. 美 ≠ 羊 + 大
 b. 美 = 羊 + 大

實際手寫筆跡的變化也顯示漢字構詞學的知識起了作用。正如 Wang (1983, 第 129 頁到第 134 頁) 的觀察，書寫者有時候會將部首「提高」，以便讓整體字形看起來更符合部首位於左側的規範結構，如(35a)所示。在 Moser (1991) 收集的筆誤資料中，其中有不少是因為寫錯位置而產生異常的字，如例(35b)(於原文中第 32 頁的第 85 個例子)。在例(35e)中，異常迴避了突出部分相互抵觸或衝突的情況，暗示書寫者認為此漢字是單字素構字(源自 Ch'en 等, 1989, 第 73 頁)。

- (35) a. 哲 → 摺 塗 → 塗
 b. 真没意思 → 真没童
 c. 黑 → 黑

最後，當書寫者因後設語言的目的而創造新穎漢字時，他們通常遵循規範的形聲字結構，雖然整體漢字的詮釋方式更類似會意字。相關例子包含例(36a)從電視廣告即興創作的玩笑(Mair, 2015)，例(36b)避免遭到誹謗指控的失敗嘗試(顏宏駿, 2018)，以及例(36c)是從香港雨傘革命中具有視覺/聽覺的雙關語「撐傘」而來的(Ho, 2014)。

- (36) a. 甓 *duāng* (當下的感嘆聲)(成龍 *Chéng Lóng*)
 b. 蕩苒 (妓女)
 c. 揸 (扌[手], 傘 *saan*³, 撐 *caang*¹)

5. 漢字語法的實驗證據

閱讀與書寫的各項實驗對於漢字語法各方面的確立至關重要，包含漢字語法具有心理上的真實性(Liu 與 Wu, 1997)、與漢字的「產出式知識(productive knowledge)」密切相關(請參閱下文所述的假字相關研究)、屬於非象似性的抽象概念(例如 Xiao 與 Treiman, 2012 的研究發現在 213 個基本漢字中，母語非中文的閱讀者只能猜對 15 個)、獨立於知覺系統之外(amodal)(如 Tan 等(2005)是表明漢字的輸入與產出能力與漢字學習具有緊密關係的研究之一)，並且是「模塊化(modular)」的(視覺字形處理區對於中文閱讀也是至關重要的；Liu 等, 2008)。

許多表面上專注在漢字閱讀與書寫即時處理的實驗，實際上也提供了構詞知識上的證據。例如，Chen, Y. P.等（1996）請中文簡體字閱讀者依照類似表 9 的語料進行成對漢字的異同判斷，實驗結果發現，漢字組成成分越多，閱讀者的反應越慢，意味著閱讀漢字存有自動分解機制。Chen 與 Cherg (2013) 發現漢字在書寫時也會自動分解，因為如果相同的組成成分越早呈現在一組漢字的開頭（書寫形式），繁體字的書寫形式會越快被啟動。在這兩項研究中，漢字組成成分（不論是意義上或發音上）的可解讀性無關宏旨。

表 9. 漢字差異檢測任務的語料樣本

漢字類型	不同漢字字組	相同漢字字組
兩個組成成分	裘 杂	雪 雪
三個組成成分	喝 唱	读 读

然而，其他實驗證明漢字組成成分是可以被解讀的。Feldman 與 Siok (1999) 將真字和結構完整的假字展示給簡體字閱讀者，接著請閱讀者決定何者為真、何者為假。每一個主要刺激字 (target) 出現之前會先出現觸發刺激字 (prime)，觸發刺激字有時共享同樣的語義部首或相關的漢字字義，有時則與語義部首或漢字字義無關，如同表 10 的例子所示。此實驗的關鍵結果是如果觸發刺激字與主要刺激字之間的部首意義和整體字義不相稱，詞彙判斷速度就會變慢（如表 10 中的粗體字，其包含部首「讠」，但是字義和言談的意思無關）。

表 10. 觸發詞彙判斷實驗的語料樣本

主要刺激字: 论		語義部首	
		相同	不同
漢字字義	相關	评	述
	不相關	诸	竿

語音成分也可以藉由發音來解釋（例如 Lee 等，2006）。不過，和類詞綴的語義部首有所不同，語音成分更類似詞幹的角色，閱讀者通常會將這些語音成分視為開放詞類，像是在詞彙判斷作業被排除的假字無論它們的語音成分是不是真的用在真字上，這些假字被排除的速度都相同（Mattingly 與 Hsiao，1999）。

在實驗研究的相關文獻中，漢字音韻學鮮少被研究，Myers (2016) 的研究是少有的例外。在他的研究中，中文繁體字的閱讀者需要對一組假字進行接受度判斷，這些假字包含交叉詞彙狀態（意即無論重複形式是否確實出現在真字中）與語法狀態的重複結構（意即無論重複字形是否適合韻律模版），如表 11 所示。研究主要發現詞彙狀態與語法狀態這兩個因素各自獨立影響判斷的結果。此外，該研究發表後再重新分析的結果（相關細節請參閱 Myers，2019）顯示，這兩個因素幾乎同時影響判斷結果，如圖 7 中的圖表所示。圖中顯示出受試者隨著時間過去給出「是」的回應的累積發生率（Scheike 與 Zhang，2011），而 95% 信賴帶首次高於基線則在圖中以直線標示之（根據 Baayen 與 Blanche，2017 的分析）。

表 11. 假字接受度判斷作業的語料樣本

字形	合語法		非合語法	
	詞彙	非詞彙	詞彙	非詞彙
橫向	棚	襁	棚	襁
縱向	窠	窠	窠	窠
三角	窠	窠	窠	窠

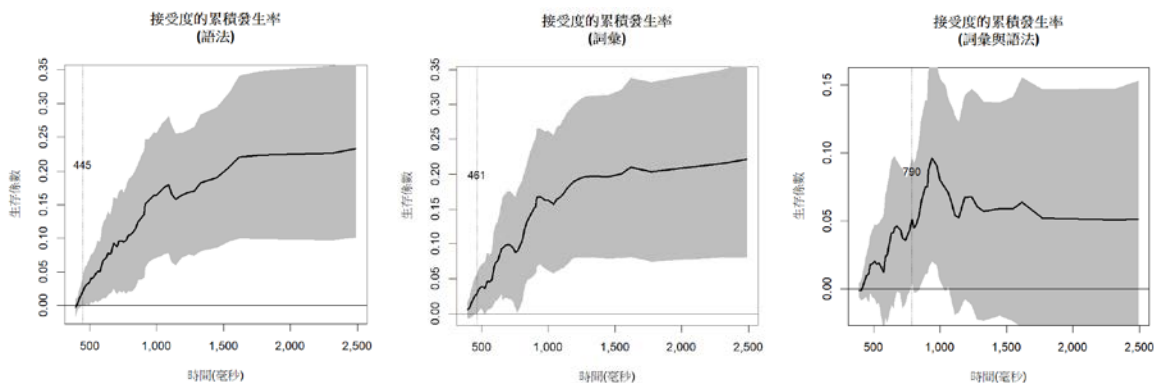


圖 7. 語法和詞彙對漢字接受度在不同時間階段的影響

另一個實驗（在 Myers (2019) 中報告）請中文繁體字的閱讀者評估圖 8 左邊的三筆畫組合體，圖 8 右邊僅呈現八種上鉤組合體的評估結果。評估結果顯示，閱讀者不僅知道向左上鉤屬於右側邊緣，同時也與 Wang (1983) 的觀察一致，意即他們也偏好具有頂部物件的類型。

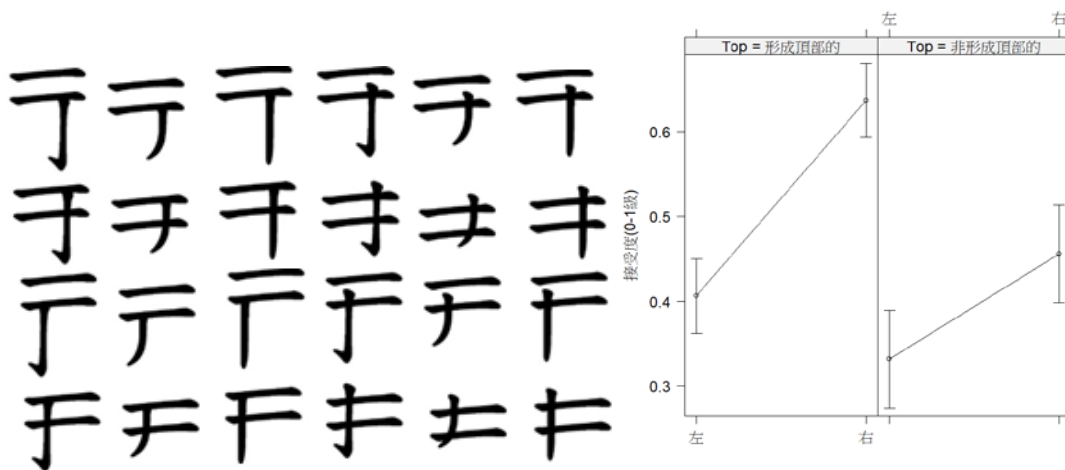


圖 8. 假字接受度作業的刺激（左圖）與部分結果（右圖）

6. 研究意涵與應用

在最後這一節，我們將探討上述語料證據的研究意涵，以及如何有效應用。關於第一節所提到的，漢字語法似乎有四種常規類型的解釋。第一種解釋留意到漢字語法系統在其中發展的特殊文化和語言環境。非語素文字系統（non-logographic system）也有視覺語法（例如 Evertz, 2018），但差別在於非語素文字系統的字素更少，比起二維結構較偏好線性結構，且相對於口語的直接詮釋，非語素文字系統的語法豐富度更顯貧乏。此外，雖然其他語素文字系統也具有語法，包含類詞綴的語義標記（例如埃及象形文字

中的限定符號：Ritner, 1996)，不過漢字書寫文字是這類拼字法的唯一倖存者之事實，可能是起因於一種獨特歷史因素的匯流，特別是口語中以音節為基礎的音韻系統與以詞根為基礎的構詞系統，以及本身文化背景的集中性和孤立性。

如同我們在前幾節所暗示的，第二類解釋觸發出非常普遍的動作性、感知性與認知性的限制約束。例如，漢字「韻律」會與細粒度與粗粒度訊息的平行處理路徑有關（例如，關於視覺：Yamaguchi 等，2000）。如圖 9 所示，粗粒度模式強調重複的結構部分，而細粒度模式則讓基本要素更為顯著。

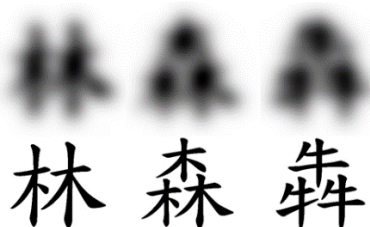


圖 9. 重複結構在粗粒度與細粒度下的視域狀態

第三類解釋認為，人腦已經進化到可以處理比任何特定情態（modality）還更抽象的語言，這個概念和自然手語的存在一致。支持這個觀點的一種說法是人們容易在靜態的二維視覺顯示中學習語法（例如，Pothos 與 Bailey, 2000）。但反對這個觀點的說法則是認為，和拼寫（視覺字形處理區）最為明顯連結的腦區與視覺皮質相鄰，但與布洛卡區（Broca's area）有一段距離。

最後，漢字語法可能從構成所有複雜系統的數學定律（在喬姆斯基 Chomsky, 2015 和其他研究中推測的分類）衍生而來，Abler (2005) 和 Zhang (2006) 單獨將這樣的概念應用在漢字研究上。然而，無論因果理論的解釋或說明是否似乎仰賴某一人的人格傾向 (Myers, 2012)，類人語法明顯地不是不可避免的，而此事實暗示著一個完整的解釋應該需要不只少數幾個簡潔的數學方程式。

無論我們用什麼方式解釋，漢字語法仍然可以在現實世界中應用。例如，幼兒教育者應該認知到，如同口語的（和手語的）語言，孩童要掌握一套拼寫系統是需要大量仰賴於他們自己的（可能是無意識的）各種發現。事實上，當孩童在自己的拼寫探索過程中受到鼓勵，有些孩童的確會從中獲得助益（如 Chan 等，2008）。因為漢字語法，就如同所有語法一樣，是獨立於知覺系統之外的（amodal），基礎閱讀者也像寫作者一樣需要獲得經驗（如 Tan 等，2005）。更具體而言，孩童需要學習漢字解構（如 Anderson 等，2013）、語義部首的位置偏好（如 Tsai 與 Nunes, 2003）、語義部首的意思（如洪國鈞等，2010）、語音部件的發音（如 Li 等，2018），以及結構完整的創新筆畫組合體之形成原因（如劉敏，2013）。然而，有鑒於筆順比語法本身更歸屬於語音層面，筆順這種需要高度反覆訓練的字詞學習方式，應該能夠值得重新思索。例如，Law 等人（1998）研究在孩童習寫繁體字所公認的筆順錯誤，他們討論數個例子，這些例子實際上都是漢字解構（漢字構詞）中的變體。

針對成人外國學習者的字詞教學也能從語法角度獲得助益。如同學習新的口語或手語，在學習新的拼寫系統時，成人不僅需要抑制既有的先前知識（Bassetti, 2013），還需要比孩童更強烈的後設語言覺察（meta-linguistic awareness），使他們得以從外顯的語法分析中有所收穫（就如同 Li, 2015 表示華語教學更為普遍）。成人也不需要嚴格地依照低頻到高頻的順序來學習漢字，這樣的方式可以幫助他們避免過度接觸到傾向於較高形符頻率（如同英文中的 *ate* 與 *ingested*）的不規則漢字。在針對外國成人日文學習者的日文漢字教學中，Tollini (1995) 明確提倡這種教學策略。

在臨床語言學中，語法觀點有助於解釋孩童的發展性（先天性）閱讀障礙（developmental dyslexia）與成人的後天性閱讀障礙（acquired dyslexia）和書寫障礙（agraphia）之間本質上的差異。以孩童而言，主要是學習漢字語法受到損害，像是對於結構完整但實為非詞彙漢字的異常反應所呈現的樣子（例如 Tzeng 等，2018）。然而以成人而言，漢字語法是多餘的，而且甚至彌補詞彙觸接的不足，像是過度依賴發音中的語音成分（例如 Yin 等，2005）和書寫文字的組成成分替代錯誤（例如 Han 等，2007）。

最後，計算語言學也應該被提醒語法的價值。語法的價值曾經位於閱讀模型的核心位置（例如 Dai 等，2007），但是因類推能力持續成長的「深度學習」（deep learning）網絡而日漸被忽視（Schmidhuber，2015）。然而，即便進行了超人般的訓練，這樣的模型與人類相比，仍然傾向異常地沒有彈性（Waldrop，2019）。例如，為了學習 1,000 個漢字，Cireşan 等（2012）訓練一組具有成千上萬節點的網路在 14 小時內處理超過 25 萬的形符（tokens）。雖然訓練過程在經由 Elman（1993）提出的「從小處著手」（starting small）策略之後速度加快，但是這套參與其中的簡化初始訓練設置，在某種程度上未見自然語言習得的過程。然而，作為我們實際大腦功能中現今最好的模式，人工神經網路也在語法研究上具有極高的潛力（例如 Silfverberg 與 Hulden，2018），並在其他部分幫助提取那些對於（有意識的）人眼來說不太明顯的樣態。

因此，我相信，漢字語法不僅具有實證上的合理性和理論上的重要性，而且也能從許多可行的研究方向中有所收穫，例如：分析的、歷史的、比較的、語料庫的、實驗的、教學的、臨床的和計算機的等。本研究（Myers（2019））所言絕非定論，而是期待拋磚引玉，鼓勵大家繼續探索這塊富饒的研究領域。

參考文獻

- 林桂生（2014）。汉字笔顺规范的理论是路径法。「黄冈师范学院学报」，34（1），91-96。
- 洪國鈞，李姝慧，陳修元，周泰立（2010）。語意部件與關聯強度對成人與國小五年級孩童漢字語意處理效應的差異。「中華心理學刊」，52（3），327-344。
- 張小衡，張煥淇（2013）。兩岸漢字規範筆順比較。「中國語文通訊」，92（1），17-26。
- 黃德寬（2003）。汉字构形方式的动态分析。「安徽大学学报（哲学社会科学版）」，27（4），1-8。
- 維基詞典（2019）。「維基詞典」。<https://www.wiktionary.org/>。
- 劉敏（2013）。「中文兒童組字規則覺識發展」。臺北市立教育大學幼兒教育學系碩士班碩士論文。
- 顏宏駿（2018）。自創文字罵人「妓女」判拘 30 天。「自由時報」，三月二十日，<http://news.ltn.com.tw/news/society/paper/1185367>。
- Abler, W. L. (2005). *Structure of matter, structure of mind*. Sofia: Pensoft.
- Anderson, R. C., Ku, Y. M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to see the patterns in Chinese characters. *Scientific Studies of Reading*, 17(1), 41-56.
- Baayen, H., & Blanche, P. (2017). Dynamic survival analysis. Talk presented at *The Workshop on New Methods in Statistics*. Tübingen, Germany, 19-20 January 2017.
- Bassetti, B. (2013). Bilingualism and writing systems. In T. K. Bhatia & W. C. Ritchie (Eds.), *The handbook of bilingualism and multilingualism, 2nd ed* (pp. 649-670). Hoboken, NJ: Wiley-Blackwell.
- Behr, W. (2006). Homosomatic juxtaposition and the problem of 'syssymantic' (huìyì) characters. In F. Bottéro & R. Djamouri (Eds.), *Écriture chinoise: données, usages et*

- représentations* (pp. 75-114). Paris: École des hautes études en sciences sociales, Centre de recherches linguistiques sur l'Asie orientale.
- Berent, I., & Dupuis, A. (2018). The unbounded productivity of (sign) language. *The Mental Lexicon*, 12(3), 309-341.
- Bloomfield, L. (1933). *Language*. New York: Holt.
- Brentari, D. (2011). Sign language phonology. In J. Goldsmith, J. Riggle, J., & A. C. L. Yu (Eds.), *The handbook of phonological theory, second edition* (pp. 691-721). Hoboken, NJ: Wiley-Blackwell.
- Ch'en, T. T., Link, P., Tai, Y. J., & Tang, H. T. (1989). *Chinese primer: Character workbook*. Cambridge, MA: Harvard University Press.
- Chan, L., Cheng, Z. J., & Chan, L. F. (2008). Chinese preschool children's literacy development: From emergent to conventional writing. *Early Years*, 28(2), 135-148.
- Chen, J. Y., & Cheng, R. J. (2013). The proximate unit in Chinese handwritten character production. *Frontiers in Psychology*, 4, Article 517.
- Chen, K. J., Huang, C. R., Chang, L. P., & Hsu, H. L. (1996). Sinica Corpus: Design methodology for balanced corpora. *Proceedings of the 11th Pacific Asia Conference on Language, Information and Computation (PACLIC 11)*, 167-176.
- Chen, Y. P., Allport, D. A., & Marshall, J. C. (1996). What are the functional orthographic units in Chinese word recognition: The stroke or the stroke pattern? *The Quarterly Journal of Experimental Psychology*, 49(4), 1024-1043.
- Chomsky, N. (2005). Three factors in language design. *Linguistic Inquiry*, 36(1), 1-22.
- Chomsky, N., & Halle, M. (1968). *The sound pattern of English*. New York: Harper and Row.
- Cireşan, D. C., Meier, U., & Schmidhuber, J. (2012). Transfer learning for Latin and Chinese characters with deep neural networks. *The 2012 International Joint Conference on Neural Networks(IJCNN)*, IEEE. 1-6.
- Dai, R., Liu, C., & Xiao, B. (2007). Chinese character recognition: history, status and prospects. *Frontiers of Computer Science in China*, 1(2), 126-136.
- Dehaene, S., & Cohen, L. (2011). The unique role of the visual word form area in reading. *Trends in Cognitive Sciences*, 15(6), 254-262.
- Elman, J. L. (1993). Learning and development in neural networks: The importance of starting small. *Cognition*, 48(1), 71-99.
- Emmorey, K. (2001). *Language, cognition, and the brain: Insights from sign language research*. London: Psychology Press.
- Evert, S., & Baroni, M. (2007). zipfR: Word frequency distributions in R. *Proceedings of the 45th Annual Meeting of the Association for Computational Linguistics, Posters and Demonstrations Sessions*, 29-32.
- Evertz, M. (2018). *Visual prosody: The graphematic foot in English and German*. Berlin: Walter de Gruyter.
- Feldman, L. B., & Siok, W. W. (1999). Semantic radicals contribute to the visual identification of Chinese characters. *Journal of Memory and Language*, 40(4), 559-576.
- Frishberg, N. (1975). Arbitrariness and iconicity: Historical change in American Sign Language. *Language*, 51(3), 696-719.
- Goodnow, J. J., & Levine, R. A. (1973). 'The grammar of action': Sequence and syntax in children's copying. *Cognitive Psychology*, 4(1), 82-98.
- Han, Z., Zhang, Y., Shu, H., & Bi, Y. (2007). The orthographic buffer in writing Chinese characters: Evidence from a dysgraphic patient. *Cognitive Neuropsychology*, 24(4), 431-450.
- Handel, Z. (2018). The cognitive role of semantic classifiers in modern Chinese writing as reflected in neogram creation, University of Washington ms.

- Handel, Z. (2019). *Sinography: A cross-linguistic study of the borrowing and adaptation of the Chinese script*. Leiden: Brill.
- Hargus, S., & Kaisse, E. M. (Eds.) (1993). *Studies in lexical phonology*. Cambridge, MA: Academic Press.
- Ho, T. L. M. (2014). Letters from Hong Kong: Characters under the Cantonese umbrella. *Asian Review of Books*. <http://www.asianreviewofbooks.com/pages/?ID=2053>.
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203(3), 88-96.
- Householder, F. W. (1971). The primacy of writing. In F. W. Householder (Ed.), *Linguistic speculations* (pp. 244-264). Cambridge, UK: Cambridge University Press.
- Jackendoff, R. (2010). The ecology of English noun-noun compounds. In R. Jackendoff (Ed.), *Meaning and the lexicon* (pp. 413-451). Oxford: Oxford University Press.
- Kordek, N. (2013). *On some quantitative aspects of the componential structure of Chinese characters*. Poznań: Wydawnictwo Rys.
- Ladd, D. R. (2014). *Simultaneous structure in phonology*. Oxford: Oxford University Press.
- Law, N., Ki, W. W., Chung, A. L. S., Ko, P. Y., & Lam, H. C. (1998). Children's stroke sequence errors in writing Chinese characters. *Reading and Writing: An Interdisciplinary Journal*, 10, 267-292.
- Lee, C. Y., Tsai, J. L., Chiu, Y. C., Tzeng, O. J. L., & Hung, D. L. (2006). The early extraction of sublexical phonology in reading Chinese pseudocharacters: An event-related potentials study. *Language and Linguistics*, 7(3), 619-636.
- Li, A. Y. H. (2015). Facilitating language learning: A generative perspective. In H. Tao, Y. H. Lee, D. Su, K. Tsurumi, W. Wang & Y. Yang (Eds.), *Proceedings of the 27th North American Conference on Chinese Linguistics, vol. 1*, 1-21, UCLA, Los Angeles.
- Li, L., Wang, H. C., Castles, A., Hsieh, M. L., & Marinus, E. (2018). Phonetic radicals, not phonological coding systems, support orthographic learning via self-teaching in Chinese. *Cognition*, 176, 184-194.
- Liu, C., Zhang, W. T., Tang, Y. Y., Mai, X. Q., Chen, H. C., Tardif, T., & Luo, Y. J. (2008). The visual word form area: evidence from an fMRI study of implicit processing of Chinese characters. *Neuroimage*, 40(3), 1350-1361.
- Liu, I. M., & Wu, J. T. (2017). Reading characters and words, behavioral studies. In R. Sybesma, W. Behr, Y. Gu, Z. Handel, C. T. J. Huang & J. Myers (Eds.), *Encyclopedia of Chinese language and linguistics*, vol. III (pp. 532-536). Leiden: Brill.
- Lui, H. M., Leung, M. T., Law, S. P., & Fung, R. S. Y. (2010). A database for investigating the logographeme as a basic unit of writing Chinese. *International Journal of Speech-Language Pathology*, 12(1), 8-18.
- Mair, V. (2015). Duang. *Language Log*, 三月一日, <http://languagelog.ldc.upenn.edu/nll/?p=17913>.
- Matthews, P. H. (1991). *Morphology, 2nd ed.* Cambridge, UK: Cambridge University Press.
- Mattingly, I. G., & Hsiao, P. L. (1999). Are phonetic elements in Chinese characters drawn from a syllabary? *Psychologia*, 42(4), 281-289.
- McCarthy, J. J., & Prince, A. (1998). Prosodic morphology. In A. Spencer & A. M. Zwicky (Eds.), *The handbook of morphology* (pp. 283-305). Oxford: Blackwell.
- McCawley, J. D. (1994). Some graphotactic constraints. In W. C. Watt (Ed.), *Writing systems and cognition: Perspectives from psychology, physiology, linguistics, and semiotics* (pp. 115-127). Berlin: Springer.
- Morioka, T. (2008). CHISE: Character processing based on Character ontology. In T. Tokunaga & A. Ortega (Eds.), *Large-scale knowledge resources. Construction and application* (pp. 148-162). Berlin: Springer.
- Moser, D. (1991). Slips of the tongue and pen in Chinese. *Sino-Platonic Papers*, 22, 1-45.

- Myers, J. (2012). Cognitive styles in two cognitive sciences. In N. Miyake, D. Peebles & R. P. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society*, 2067-2072. Austin, TX: Cognitive Science Society.
- Myers, J. (2016). Knowing Chinese character grammar, *Cognition*, 147, 127-132.
- Myers, J. (2019). *The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system*. London: Routledge.
- Naka, M. (1998). Repeated writing facilitates children's memory for pseudocharacters and foreign letters. *Memory & Cognition*, 26(4), 804-809.
- Pacton, S., Perruchet, P., Fayol, M., & Cleeremans, A. (2001). Implicit learning out of the lab: The case of orthographic regularities. *Journal of Experimental Psychology: General*, 130(3), 401-426.
- Padden, C. A., & Perlmutter, D. M. (1987). American Sign Language and the architecture of phonological theory. *Natural Language & Linguistic Theory*, 5(3), 335-375.
- Peng, X. (2017). Stroke systems in Chinese characters: A systemic functional perspective on simplified regular script. *Semiotica*, 2017(218), 1-19.
- Pothos, E. M., & Bailey, T. M. (2000). The role of similarity in artificial grammar learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26(4), 847-862.
- Ritner, R. K. (1996). Egyptian writing. In P. T. Daniels & W. Bright (Eds.), *The world's writing systems* (pp. 73-84). Oxford: Oxford University Press.
- Sandler, W., & Lillo-Martin, D. (2006). *Sign language and linguistic universals*. Cambridge, UK: Cambridge University Press.
- Scheike, T. H., & Zhang, M. J. (2011). Analyzing competing risk data using the R timereg package. *Journal of Statistical Software*, 38(2), 1-15.
- Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural Networks*, 61, 85-117.
- Silfverberg, M., & Hulden, M. (2018). Initial experiments in data-driven morphological analysis for Finnish. In T. A. Pirinen, M. Rießler, J. Rueter & T. Trosterud (Eds.), *Proceedings of the Fourth International Workshop on Computational Linguistics of Uralic Languages*, 100-107, Association for Computational Linguistics, Stroudsburg, PA.
- Sproat, R. (2000). *A computational theory of writing systems*. Cambridge, UK: Cambridge University Press.
- Stalph, J. (1989). *Grundlagen einer Grammatik der sinojapanischen Schrift*. Wiesbaden: Harrasowitz Verlag.
- Sweetser, E. E. (1988). Grammaticalization and semantic bleaching. In K. Lambrecht, S. Axmaker, A. Jaisser & H. Singmaster (Eds.), *Proceedings of the Fourteenth Annual Meeting of the Berkeley Linguistics Society*, 389-405, Berkeley Linguistics Society, Berkeley.
- Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *Proceedings of the National Academy of Sciences*, 102(24), 8781-8785.
- Tollini, A. (1994). The importance of form in the teaching of kanji. 「世界の日本語教育」, 4, 107-116.
- Tsai, K. C., & Nunes, T. (2003). The role of character schema in learning novel Chinese characters. In C. McBride-Chang & H.-C. Chen (Eds.), *Reading development in Chinese children* (pp. 109-125). London: Praeger.
- Tzeng, Y. L., Hsu, C. H., Lin, W. H., & Lee, C. Y. (2018). Impaired orthographic processing in Chinese dyslexic children: Evidence from the lexicality effect on N400. *Scientific Studies of Reading*, 22(1), 85-100.
- Waldrop, M. M. (2019). What are the limits of deep learning? *Proceedings of the National Academy of Sciences*, 116(4), 1074-1077.

- Wang, J. C. S. (1983). *Toward a generative grammar of Chinese character structure and stroke order*. Ph.D. thesis, University of Wisconsin-Madison.
- Wikimedia Commons. (2017). *Chinese character decomposition*. https://commons.wikimedia.org/wiki/Commons:Chinese_characters_decomposition.
- Xiao, W., & Treiman, R. (2012). Iconicity of simple Chinese characters. *Behavior Research Methods*, 44(4), 954-960.
- Yamaguchi, S., Yamagata, S., & Kobayashi, S. (2000). Cerebral asymmetry of the 'top-down' allocation of attention to global and local features. *Journal of Neuroscience*, 20(9), Article RC72.
- Yin, W., He, S., & Weekes, B. S. (2005). Acquired dyslexia and dysgraphia in Chinese. *Behavioural Neurology*, 16(2-3), 159-167.
- Yip, P. C. (2000). *The Chinese lexicon: A comprehensive survey*. Abingdon: Routledge.
- Zhang, H. Y. (2006). The evolution of genomes and language. *EMBO Reports*, 7(8), 748-749.