

香港中文字形問題整理

劉擇明 陳輝恒

摘要

香港沒有一套通行的規範中文字形標準，多年磨合下雖然有不少寫法分歧已定於一尊，但仍存有不少個人差異。近年，小學中文教育常以《常用字字形表》的標準字形作為參考，排除其他異體字。然而該表未有適時修訂，對筆畫及部件的要求時有背離慣常用法，影響中文的識字教學。本研究旨在審視字形判斷的基本原則，在字形表的基礎上有系統地提取各種字形問題，討論縮小標準字形與民間習慣差距的方向。本文把現行標準字形及異體字之間的問題分類為異字、異構、異寫、部件混同四大類共 180 項，並按照字頻、轄字量及網上用字分佈，選出具有爭議性的 66 項作為手寫體的調查點，為日後制定新版字形標準提供基礎。

關鍵詞：常用字字形表 標準字形 手寫體 異體字

一、引言

漢字數目眾多，同一漢字經歷各種字體變化，不時在部首、部件、筆畫和配置方式上出現大小差異，形成各種變體。漢字文化圈中，不少地方進行

劉擇明博士，香港教育大學語言學及現代語言系助理教授，近年的研究範疇為數碼人文及語言資源建設與評估，尤其關注低資源語言的編碼及數據建模問題，項目包括粵語詞典資源建設、為香港少數族裔開發的 TypeDuck 粵拼輸入法等。

陳輝恒，擁有工商管理學（資訊系統）學士學位，現職軟件工程師。2016 年起以個人身份參與國際標準化組織轄下的國際表意文字小組，負責整理及審批 ISO/IEC 10646 內的新中日韓統一表意文字擴展區的編碼工作。2023 年起通過青年委員自薦計劃，獲委任成為香港特別行政區中文界面諮詢委員會委員。

了文字標準化，依據當地書寫或印刷習慣，為每個字元選定一個寫法作為教學或工業標準，並適時修正。大灣區當中，香港和澳門情況相似，都使用繁體字，香港最主要的字形整理工作成果是《常用字字形表》，由教育工作者和文字專家帶領編纂，但只供教師參考，並非官方標準；澳門雖然有少數本地造字，但字形方面與香港沒有明顯區別；其他大灣區省市的中文運用雖受到港澳影響，但沒有發展出獨特的字形習慣。因此，本文只集中討論香港狀況。

香港民間使用的字形比較寬鬆，保留了不少與該表不同的字形。社會大眾的書寫習慣隨時發生變化，加上各種社會及技術原因，字形表及其後續的版本在內容和教育應用上已出現各種問題（陳曙光，2021）。本文首先探討字形問題，然後從語言學角度為字形問題分類，以此為基礎整理異體字的主要爭議點，最後探討日後修訂字形表的啟示。

二、香港的字形問題

本文探討的「字形」參考了以下兩個定義：

1. 字形 (character form)

特指構成每個方塊漢字的**二維圖形**。構成漢字字形的要素是筆畫、筆數及漢字部件的位置關係等。（中華人民共和國教育部，2005）

2. 字形 (character glyph)

特指構成每個方塊漢字的**可識別抽象圖形符號**，為**實際圖形的抽象表示**。構成漢字字形的要素是筆畫、筆數、漢字部首和部件，以及其相對位置等。（政府資訊科技總監辦公室與公務員事務局法定語文事務部，2016，頁4）

對一般中文使用者而言，上面第二個定義比較符合大眾對字形的理解：字形不似圖形般可以用像素比對，而是「可識別的抽象圖形符號」。每個漢字的構成要素應該有一定準則，不能任意改動，例如「早」不可以改成左右結構「叶」，「香」不能改變部件成為「杳」，「千」不能以橫代改撇改成「干」等等；但準則也有灰色地帶，例如以香港習慣而言，「竹」字的最後一畫可以

寫作「丨」或「丨」，「舌」的上半可以寫作「千」或「干」。這種抽象性也是字形問題的癥結所在：標準不單要指出怎樣寫可以接受，同時也要容許無關痛癢的細微風格差異，如筆畫的起點終點、元件的大小比例等。本文的重點在於探討現行標準及日常使用者對漢字的構成要素的分歧，集中整理字形的具體要求、容許範圍、標準更新所引起的問題。

電腦普及前，輸入字元不受軟件限制，手寫選字、字形相當自由。在這個背景下，香港的教育工作者於 1986 年制定《常用字字形表》（香港教育署語文教育學院中文系，1986），並已列出 4,721 字的建議寫法（後增至 4,762 字），包括部件及筆畫的要求，並於建議寫法後列舉常用異體字。其後不論是教學還是工程上的整理都以該表作為主要參考。教學上的延伸整理包括該表多個修訂版，其中《常用字字形表》（2000 年修訂本）（李學銘，2000）為該表最後一個獨立印刷的版本。另外還有以 1993 版為基礎編制的《香港小學學習字詞表》（2007）中以附錄形式出現的重排版，及獨立印行的《常用字字形表（2007 年重排本）附粵普字音及英文解釋》（2012）。教育局也提供不斷更新的網頁查詢系統，增加字元並對字形作微調。另一方面，業界也需要工業規格標準文件，以製作本港日常使用及教學的電腦字體。資訊科技辦公室轄下的中文界面諮詢委員會為業界出版了《香港電腦漢字楷體字形參考指引》（2002）及《香港電腦漢字宋體（印刷體）字形參考指引》（2002）、《香港電腦漢字參考字形》（2016）等參考資料。後續討論以《字形表》指《常用字字形表》的六個印刷版本，並以「網站版」指網頁查詢系統。

《常用字字形表》說明（2000，頁 XIX）明確指出選字原則有三項——普遍性、學術性及規律性，並列出優先考慮順序，明確指出「只符合『普遍性』」的優先順序高於「符合任何兩項，不包括『普遍性』」。這代表一旦有了約定俗成的寫法，符合學術性和規律性也應當讓路。重視普遍性代表該表有持續修訂及維護的必要，但現時卻沒有任何公開機制去達成這個目標。資料辦的工業字表修訂沒有與網站版同步，也沒法反映在教學上。結果而言，雖然《字形表》已提供大部分香港常用字的標準字形，但在缺乏公開維護的流程下，其在日常教學及應用上的推廣遇到以下一些困難：

(續表)

調查項	描述	影響字元	調查點
白白	標準字形作「白」；另有區分「白、白」的寫法	叟嫂庾插搜瘦叟艘莢颯餽 (11)	瘦
𠂇四	標準字形於 2007 年曾作修改，原本右側是「賣」，中間從「四」，現時改為「𠂇」	犢瀆犢犢寶續讀賣讀黷 (10)	讀
𠂇凡丸	隸定從「𠂇」的部件中，「恐」字頂標準字形為「豎點」（點穿），另有「撇點」、從「凡」寫法；「執、孰」等標準字形為「撇點」（點穿），右半與「丸」標準字形「點在撇內側」的寫法不同	恐鞏築 (3)；執塾塾孰擊熟 (6)	執、恐
不調查部分：類近「另」的偏旁（拐拐別）；「囟、囟」（囟窗）			

上面列出異字、異構、異寫、部件混用問題等項，共有 66 項調查點應作廣泛調查。

四、餘論

本文審視現有字形標準，探討其問題並提出更新方案，以改善中文教育及應用。

回應文首提出的研究問題，《字形表》有異體寫法的條目最少有 1192 項，是難以逐一調查的數量。就香港情況而言，社會一般以繁體字為預設字體，而在日常電腦使用上也受到台灣的規範與用法影響。整理標準字形時，除了按照異體字和語言層級之間的對應，區分異字、異構、異寫和部件混同的例子之外，也可以從網上語料和現行台灣標準出發，比對差異。上面列表儘量窮盡潛在爭議——異體字表中大部分標準字形已經是主流用法，但仍有與日常使用習慣不同之處。上面提出以調查點的方式，觀察各種類別的異體狀況。

事實上《字形表》開宗明義是以「普遍性」為首要原則的，其次才是「學術性」和「規律性」。有些字形由始至終都採取從俗方向，如「溫」「說」等；2007 版的參考字形「潛」由「先」改作「无」，取消「賣」與「讀」的區分也是同樣的改變方向。標準字形違反普遍習慣的，理應都要有強力的字源以支持，基於學術性、規律性原則予以保留。陳曙光（2021）分析《字形表》的字形取捨。他指出雖然該文指出編訂多數是符合字理的，例如按字理區分「冂、曰」「𠂇、𠂇」「壬、王」，收俗字（戶、溫）同時接受符合字源寫法（戶、溫），

但《字形表》也有值得商榷之處，如「田」「用」橫畫不相連、「𠂇」「齊」帶鈎、採用後起的俗字「粧」而非「妝」，又合併「飢、饑」「歎、嘆」等問題。今次有較完整的列表，就可以窺探問題規模。

本文歸入「異寫」的字形當中，部分是出於《字形表》的筆畫要求，但這些要求不一定有充分的字源理據，例如「致」採用符合字源的「攴」，而非坊間大部分電腦字體以及國家標準採用的「攴」。一些手寫體的筆勢問題，《字形表》也嚴格指定，卻難以保持一致，例如「屯」在舊版無勾，2007年起加勾，「在、左、亦」都用「豎勾」避讓，但「竹」「『選』右上『巳』」「『潛』右上『死』」都沒勾。撇—左點（「小、少」採撇，「業」採左點），捺—右點（「尖」採捺，「矣」改成點）等。「陷」改成不符字源的帶勾寫法，純粹按照部件位置來區分是否帶勾。另外註明「首筆本作橫畫」的字，例如「舌」「吞」等，依從台灣「國字標準」但和一直以來的寫法不同。這些細微差異是否值得執着，應回歸到普遍性、學術性、規律性原則討論。

而本文指出的「部件混同」和「異構字」問題，不少是近代至現代的字形規範下，試圖逆轉合流的產物。

其一為「月、肉」的分立。兩者是自隸變以來已經合流，清《康熙字典》(1716)嘗試以兩橫不觸右方為「月」，兩橫相觸右方為「肉」，後者因手寫筆勢而寫成「點提」（即提肉旁）而被台灣「國字標準」採用，不論部件位置都寫作提肉旁。香港似乎為了兼顧美觀，「胃」「胡」等字「肉」在下方或右側時改從「兩橫」之形（目）。香港刻意區分「月肉」但採用獨有的「兩橫」之形，這個特別寫法與「月」相似，也犧牲了標準字形的規律性。

其二為形近「匕」的幾個部件，「叱」作「豎彎勾、撇」，「化」作「撇、豎彎勾」，「老」作「撇、豎彎」，而來源同為「匕」的「匕、比」作「橫、豎彎勾」，「此、北」作「撇、豎彎勾」但接受「橫、豎彎勾」，等。這些要求沒有詳細說明，也和國家標準不一致，不便教學。

其三是「慈、滋、磁」等字，不依《說文解字》從「兹」（1963，頁218），而根據段玉裁（1981，頁504）《說文解字註》用從「艸」的「茲」，而「孳」卻繼續從「兹」。「兹」「茲」的區別尚有爭議，而從「艸」的寫法只有香港採用為標準字形。

其四是「執、巩」（本從「𠂇」）與「丸」「凡」的區別，指「執」點不穿

和「丸」不同，指「恐、鞏、築」與「凡」不同，這些要求或缺普遍性支持，《字形表》也沒有說明原因。

標準字形的獨特寫法要求辨義的部件有理據支持，而其規律性理應有助於中文教學，但小部分字形推廣近四十年，依然沒有影響主流用法，或許都反映這些學術性原則曲高和寡，因而一般人難以理解。為追溯字源而改變文字固有面貌的做法並不可取，也難免會犧牲美感或提高學習難度。

本文的工作希望有助於明確區分學者堅持的寫法、一般使用者的偏好以及因習慣而形成的無關緊要的差異，以助建立更合理的字形標準及容許範圍，避免因死記字形令學生進行大量無謂的抄寫更正。如果標準字形反映香港用法，即使與電腦字體有異，也應保留並推廣（兇、衛）。如果確定標準字形與普遍習慣不同，就應該向中文教師傳遞清晰訊息。長遠而言，本文的資料表可成為系統化的資料庫，追蹤記錄標準字形的決定依據和使用狀況，並公開相關數據供教師和社會大眾參考。最理想的情況是透過以上行動，推動往後的《字形表》的重訂、修正與推廣。

參考文獻

1. 陳曙光 (2021)：《香港小學學習字詞表》所收字形研究——兼論香港寫字教學之困境，《國文學報》，(70)，269-300。
2. 東網港澳版 (2016.01.12)：《規範字：小學習字表字型怪 學生無所適從》，檢自 https://hk.on.cc/hk/bkn/cnt/news/20160112/bkn-20160112070104524-0112_00822_001.html，檢索日期：2024.7.31。
3. 鄺媛萍 (2017)：淺談香港幼稚園教師的漢字字形教學：由「𠂇」和「𠂈」說起，《語文建設通訊》，(114)，53-57。
4. 李國英 (2007)：異體字的定義和類型，《北京師範大學學報（社會科學版）》，(03)，46-50。
5. 李學銘主編 (2000)：《常用字字形表》(2000年修訂本)，香港，香港教育學院。
6. 香港教育局課程發展處中國語文教育組編 (2007)：《香港小學學習字詞表》，香港，香港特別行政區政府教育局。
7. 香港教育局課程發展處中國語文教育組編 (2012)：《常用字字形表》(2007年重排本) 附粵普字音及英文解釋，香港，香港特別行政區政府教育局。
8. 香港教育署語文教育學院中文系 (1986)：《常用字字形表》，香港：香港政府印務局。
9. 香港特別行政區政府教育局課程發展處中國語文教育組 (2024)：《香港小學學習字詞表》網上版，檢自 https://www.edbchinese.hk/lexlist_ch/，檢索日期：2024.7.31
10. 熊丹、陸勤 (2019)：ISO/IEC 10646 國際編碼標準下的香港電腦漢字編碼及字形原則，《中文信息學報》，(07)，136-142。

11. 〔漢〕許慎（撰）（1963）：《說文解字》，北京，中華書局。
12. 〔漢〕許慎（撰）、〔清〕段玉裁（註）（1981）：《說文解字註》，上海，上海古籍出版社。
13. 尤盛（2017）：香港漢字書寫系統中異體字部件的簡化——以三組形似部件為例，《中國語文通訊》，96（1），19-29。
14. 〔清〕張玉書等（編）（1716）：《康熙字典》。
15. 政府資訊科技總監辦公室與公務員事務局法定語文事務部合編（2016）：《香港電腦漢字參考字形》，檢自 https://www.ccli.gov.hk/tc/download/reference_glyphs.html，檢索日期：2024.7.31。
16. 政府資訊科技總監辦公室與公務員事務局法定語文事務部合編（2017）：《香港增補字符集—2016》，檢自 <https://www.ccli.gov.hk/tc/download/terms01.html>，檢索日期：2024.7.31。
17. 中華人民共和國教育部（2005）：《漢語信息處理詞彙 02 部分：漢語和漢字》，檢自 http://www.moe.gov.cn/jyb_sjzl/ziliao/A19/201001/t20100115_75613.html，檢索日期：2024.7.31。
18. 資訊科技署與法定語文事務署合編（2002）：《香港電腦漢字楷體字形參考指引》及《香港電腦漢字宋體（印刷體）字形參考指引》，檢自 https://www.ccli.gov.hk/tc/download/terms_glyph.html，檢索日期：2024.7.31。
19. Chan, H. F. H. (2023). IDS. Retrieved from <https://github.com/hfhchan/ids/blob/main/release/ids-20230408.txt>，Accessed July 31, 2024.
20. Lam, H. C. (2020). Linguistic variation in Chinese characters: Knowledge essential for teachers. *Electronic Journal of Foreign Language Teaching*, 17(2), 489-504.
21. Lu, Q., Chan, S. T., Li, Y., & Li, N. L. (2002). Decomposition for ISO/IEC 10646 ideographic characters. In *COLING-02: The 3rd Workshop on Asian Language Resources and International Standardization*.

A Compilation of Chinese Character Glyph Issues in Hong Kong

LAU, Chaak Ming CHAN, Fai Hang Henry

Abstract

Hong Kong lacks an official set of standardized Chinese character glyphs. Despite the convergence of multiple handwritten variants to one agreed form for many characters over the past years, individual variation still exists in many cases. There is a trend in recent years in primary school Chinese education to use the standard glyphs found in *List of Graphemes of Commonly-used Chinese Characters* as the correct form, and rejecting other character variants. This is however problematic, due to this list not being timely revised, with many of its requirements for strokes and components deviating from common usage, which negatively impacts the teaching of Chinese literacy. This study aims to review the basic principles of judgment in the practice of glyph selection, systematically extract various glyph issues based on existing lists and tables and discuss directions for narrowing the gap between the designated standard glyphs and common practices. The chapter categorizes the issues based on the relation between the standard glyph and the variant(s) for each character, which fall into four major types: difference in character selection, difference in structural/component composition, difference in stroke/writing,

and component disambiguation, with a total of 180 identified issues. These issues were then further sorted, based on character frequency, the number of affected characters, and distribution in online datasets, into 66 items as investigation points for handwriting preference. This list will provide a foundation for future construction of a new character glyph standard.

Keywords: *List of Graphemes of Commonly-used Chinese Characters*, standard glyphs, handwritten form, character variants

