

Unicode 所見香港特有漢字初探

		a [jím ¹]	a [gúŋ ²]
	343		342
劇春	仁起者空起也俗讀仁若拱說文仁大腹也讀若紅	肚仁起	
劇春者猶言割勢也俗讀劇若淹正法念經卷四十八	音義引通俗文以刀去陰曰劇集韻劇渠言切音度削	此或作剗非是男子勢何以謂之春蓋俗名勢為陽爾	
雅春為青陽呂覽孟春行冬令注春陽也此名勢曰春	之義也勢為男子陰此稱甚古周禮司刑宮罪五百注	宮者丈夫割勢韻會外腎為勢今讀勢音近漆韻補勢	
宮者式列切音設木華海賦羣山既略百川潛渫決滌滌	叶式切音設木華海賦羣山既略百川潛渫決滌滌	波赴勢此勢字之音讀如設者也以廣州音讀設	泞騰波赴勢此勢字之音讀如設者也以廣州音讀設

引自：詹鑑慈，《廣州語本字》卷七，頁 145-146。（影像作者提供）

袁國華

香港恒生管理學院中文系

古今論衡 第27期 2015.4

一、前言

香港土地面積有 1,098 平方公里。二〇〇〇年十月五日香港政府統計處，用一套新的推算人口數字方法，估計三十年後香港人口會達到 9,000,000 人；而至二〇〇九年底，香港人口估計為 7,026,400 人。^①而據二〇一一年普查所得結果，當年常住人口約為 6,860,000 人，加上流動居民約 210,000 人，居港人口實際已逾 7,070,000 人。^②

按國籍作分類，二〇一一年香港居民仍以中國人和華人為多數，約佔總人口的 91.8%；中國人和華人以外的種族，以印尼人、菲律賓人人數較多，共約 28 萬人，約佔總人口 3.8%；其次為南亞人、英國人、日本人和韓國人。

此外，還有港英時期到港工作的南亞裔人士，以及來自泰國等國家的家庭幫傭。香港中國籍人口除了漢人以外，也有如回族、滿族等其他少數民族。^③

香港居民半數以上出生於香港。過去的人口調查，都有記錄市民的祖籍。當中廣府人佔多數，其次是四邑人及潮汕人，再其次是來自上海及江浙地區的新移民。自從一九八〇年代初期中國大陸改革開放，來自各省市的移民數量增加，以至於來自各省市的居民分布更為廣泛。^④

54

香港居民當中有 89.5% 的人口，以廣府話作為日常用語，^⑤理所當然，漢字成為常用文字。^⑥再加上法律明文規定，漢字便成為香港法定語文之一。根據基本法第一章第九條規定，中、英文都是香港法定語文，然而規範漢字用字的條文當中，並沒有

① 資料見香港政府統計處網站二〇一〇年二月十一日：http://www.censtatd.gov.hk/home/index_tc.jsp。

② 資料見香港政府統計處網站二〇一一年二月二十一日，人口普查結果，簡要報告，表 8：

http://www.censtatd.gov.hk/hong_kong_statistics/statistics_by_subject/index_tc.jsp?subjectID=1&charsetID=2&displayMode=T。

③ 資料見香港政府統計處網站二〇一一年二月二十一日，人口普查結果，簡要報告，表 9：

http://www.censtatd.gov.hk/hong_kong_statistics/statistics_by_subject/index_tc.jsp?subjectID=1&charsetID=2&displayMode=T。又參維基百科：<http://zh.wikipedia.org/zh/%E9%A6%99%E6%B8%AF%E4%BA%BA%E5%8F%A3#.E7.A8.AE.E6.97.8F>。

④ 資料見香港政府統計處網站二〇一一年二月二十一日，人口普查結果，簡要報告，表 7：

http://www.censtatd.gov.hk/hong_kong_statistics/statistics_by_subject/index_tc.jsp?subjectID=1&charsetID=2&displayMode=T。又參維基百科：<http://zh.wikipedia.org/zh/%E9%A6%99%E6%B8%AF%E4%BA%BA%E5%8F%A3#.E7.A8.AE.E6.97.8F>。

⑤ 資料見香港政府統計處網站二〇一一年二月二十一日，人口普查結果，簡要報告，表 12：

http://www.censtatd.gov.hk/hong_kong_statistics/statistics_by_subject/index_tc.jsp?subjectID=1&charsetID=2&displayMode=T。

⑥ 香港法律文件及基本法條文都採用「中文」一詞代表「漢字」。事實上狹義的「中文」或「中文字」等同「漢字」；而廣義的「中文」或「中文字」可同時指稱「中國境內所有民族使用的文字」。顯然香港使用「中文」一詞，係專指「漢字」。

指定用正體字或簡化字，更何況官方所有法律文件，皆用正體字，已是不成文的規定，不必言喻，正體字就是香港法定漢字用字。^⑦

中文和英文是法定語文，政府語文政策為兩文三語——廣府話^⑧、普通話和英文。

二、廣府話與香港特有漢字

由於香港居民當中，廣府人佔多數，廣府話一直以來是香港的日常用語。因為粵方言有許多方言字可能是有音無字的，故此，就產生了許多特有的廣府話用字。以下一段文字出自維基百科廣府話條目：

廣州話，又叫省城話、白話，即係老人家講的廣府話。佢個名就講明佢係來自廣東省城廣州。因為佢係廣東省城（廣府），政治經濟都有特殊地位，所以啲粵語專家都會認為佢係粵語嘅標準音。廣府話流行廣州、佛山、中山、香港、澳門，呢啲珠三角地區，廣西、海外都有好多人係講廣府話。因為有佢響廣東有特殊地位，唔少人直情叫佢做粵語、廣東話。之不過，「廣東話」呢個叫法同學術上嘅「粵語」定義唔同，廣東重有人講潮州話、客家話。廣東話呢個詞學術上未有定義。粵語方言有好多種，好似臺山話、南寧白話咁，有啲可以同廣州話互通，有啲就好難互通。

不過唔係成個廣州地方講嘅話都係廣府話，以前淨係省城部份係講標準嘅廣府話，出咗城就有差異。後尾英國人響香港起城，香港同廣州有好多商貿來往，廣府話就慢慢代替咗當地粵語分支。而自鴉片戰爭之後，有好多珠江三角洲嘅人來回廣州、香港做嘢。因為咁，廣府話以佢嘅影響力，佢個範圍就擴大好多。另外，廣府地區（粵語粵海片）、韶關、清遠、粵語非廣府地區（粵西）有啲後生仔都識講廣州話，有啲甚至擺廣府話嚟代替自己原有嘅方言作為母語。香港電視，流行曲呢類廣府話流行文化亦加速廣府話擴散。而家講的廣州話同香港粵語其實係同一種語言，只不過表達上因地方唔同，所以會有啲分別，但完全可以理解。^⑨

^⑦ 資料見香港政府網站二〇〇八年三月十七日：http://www.basiclegal.gov.hk/text/tc/basiclegaltxt/chapter_1.html。

^⑧ 香港法律文件中三語為：廣州話、普通話及英語，法律文件所稱「廣州話」嚴格定義應為「廣府話」，至於「廣州話」、「廣府話」以及「廣東話」的細緻分別，下節「二、廣府話與香港特有漢字」引文有詳細說明可參。

^⑨ 資料見維基百科：<http://zh-yue.wikipedia.org/wiki/%E5%BB%A3%E5%BA%9C%E8%A9%B1>。

上述一段文字分析了廣府話、廣州話與廣東話在定義上的差異，並說明了廣州話與廣東話不等同廣府話的原因，其中就收錄了許多廣府話方言字，如：佢、啲、嘅、唔、咁、𠂇、𠂇、𠂇。對於熟識廣府話的人士，閱讀以上文章感受必然分外親切。但是以上香港人口語常用的「佢」等 8 字，從前都是電腦缺字，如果香港人打算將常用口語特有廣府話方言用字加入句子之中，卻是不能在電腦字庫中找到的，因此造成經常使用電腦的香港人的諸多不便。

在《香港增補字符集》制訂並公布之前，大部分特殊人名、地名、廣府話方言用字，以及罕用字，並未收錄於如大五碼及 ISO 10646 國際編碼標準等常用的中文電腦編碼標準裡面。結果導致香港電腦使用者，每當有需要使用特別中文字符之際，卻無法在電腦之上找到可用字符的窘境。雖然使用者可利用電腦造字區，自行編製，或利用「小畫家」等應用軟體繪製、組合個人急需的中文字符；但前者未必能在其他電腦上呈現正確的字符，甚或導致資料交換時，因編碼混亂，造成難以預期的錯誤；¹⁰而後者因屬圖形檔案，雖可正確傳送和呈現字符，卻失去文字檢索的功能，亦難免顧此而失彼。

三、電腦所見香港特有漢字

56

一九九九年香港特別行政區政府與中文界面諮詢委員會為了解決以上問題，於當年九月公布編製《香港增補字符集》(HKSCS)。《香港增補字符集》(HKSCS)由最初一九九九年香港通用特別字符 4,702 個，後歷經五次增補修訂，至二〇〇九年新版本《香港增補字符集－2008》字符數目已增加至 5,009 個。

期間於二〇〇一年公布《香港增補字符集－2001》，較一九九九年《香港增補字符集》新增了 116 個中文字，並在技術上配合國際標準化組織於二〇〇一年十一月公布的「ISO/IEC 10646-2:2001」。¹¹

又於二〇〇五年公布《香港增補字符集－2004》，增收了 123 個字符，在技術上配合國際標準化組織在二〇〇四年四月公布的「ISO/IEC 10646:2003」。「ISO/IEC 10646:2003」及其修訂版收納了《香港增補字符集－2004》的全部字符。

¹⁰ 參見香港特別行政區政府共通中文界面網站：http://www.ccli.gov.hk/textonly/tc_chi/itpro/itpro_ccli.html。

¹¹ ISO 乃「國際標準化組織」的英文簡稱，ISO 10646 是一個國際標準編號。ISO 頒布 ISO 10646（或稱 ISO/IEC 10646）標準所定義的標準字符集，亦即「通用字符集」(Universal Character Set，UCS)。通用字符集又稱 Universal Multiple-Octet Coded Character Set，中國大陸譯為「通用多八位編碼字符集」，臺灣譯為「廣用多八位元編碼字元集」。此標準替世界各種語言的字符（包括正體及簡體的中文字）及附加符號，編訂統一的內碼，以實現全球語言統一編碼為目標。

又於二〇〇九年政府公布《香港增補字符集－2008》，增收了68個字符，在技術上配合國際標準化組織在二〇〇五年十一月和二〇〇九年十月公布的「ISO/IEC 10646:2003」第一及第六修訂版。「ISO/IEC 10646:2003」及其第一至第六修訂版收納了《香港增補字符集－2008》全部的字符。¹²

香港居民及政府部門共同使用《香港增補字符集》後，電腦資料傳送的準確度無疑已大幅提升，而字符錯誤呈現或字符亂碼，以及無法檢索等問題，亦得以迎刃而解。¹³

四、《香港增補字符集》字符的分類

對於《香港增補字符集》字符的分類，香港特別行政區政府共通中文界面網站，有如下說明：

《香港增補字符集－2008》共收錄5,009個字符，當中絕大部分為中文字符（共4,568個，佔總數91%）。這些字符大多可在大型字典（包括《康熙字典》、《漢語大字典》、《漢語大詞典》、《中華字海》）找到。由於每個中文字符都可作不同用途，因此《香港增補字符集－2008》沒有為每個字符作詳細分類。不過，這四千五百多個字符大體上可歸入以下類別：

專有名詞用字，如人名、地名、公司名稱

粵方言用字

科學名詞用字

部首或部件

這些字符主要由政府部門（如律政司、稅務局、司法機構、地政總署、香港警務處、入境事務處、公司註冊處）、學術組織、教育機構，以及一般市民建議收錄。

至於餘下的441個字符，則屬各類符號，包括漢字部件、漢語拼音字母、國際音標符號、日本平假名和片假名等。¹⁴

¹² 參見香港特別行政區政府共通中文界面網站：http://www.ccli.gov.hk/textonly/tc_chi/itpro/itpro_ccli.html。

¹³ 參見香港特別行政區政府共通中文界面網站：http://www.ccli.gov.hk/textonly/tc_chi/itpro/itpro_ccli.html。

¹⁴ 稱引自香港特別行政區政府共通中文界面網站：http://www.ccli.gov.hk/textonly/tc_chi/itpro/itpro_ccli.html。

五、Unicode 所見香港特有漢字舉例說明

目前所見 Unicode 字集所收香港特有漢字字符共計 949 字。¹⁵

香港新增電腦漢字，除了前述專有名詞用字（如人名、地名、公司名稱）、粵方言用字、科學名詞用字，以及部首或部件以外，還有小篆隸定字形，如：

𠂇（粵音 zin2）U+382D ¹⁶

義項說明：「展」字異體。¹⁷

今案：「𠂇」字見許慎《說文解字》，意謂「極巧視也」。字的結構由四個「工」字組合而成。六書屬「會意」。《玉篇》曰：「𠂇今作展」。可見「𠂇」與「展」是一組異體字。

𠂔（粵音 cap1）U+35CA

義項說明：1. 衆口。2. 喧嘩。3. 同「雷」。4. 一種有機化合物名（porphin）。或稱「卟吩」。

今案：「𠂔」字見許慎《說文解字》，意謂「衆口也」。字的結構由四個「口」字組合而成。六書屬「會意」。此字雖為小篆隸定字形，但是收錄的原因，可能與其為「科學名詞用字」有關。

此外，從目前 Unicode 字集所收香港特有字符，得見以下現象，分別敘述於後。

（一）有本字可循的 Unicode 所見香港特有漢字

有學者認為部分廣府話是有本字可循的。過去，詹憲慈《廣州語本字·自序》說：

廣州之人操廣州語，莫以爲難也。而寫之於紙，則遲疑不能遽下筆矣。蓋廣州常語往往口所能道，而筆不能寫也？其本字未易一一知也。¹⁸

由此可見詹氏認為由於使用廣府話的人士，大多不知廣府話的本字，遂造成書寫時的困難。試舉「咁」、「嘅」兩字為例。

¹⁵ 參中央研究院資訊科學研究所，電腦漢字字形及詞彙整合知識庫：<http://chardb.iis.sinica.edu.tw/>。

¹⁶ 此為 Unicode 字集編碼，下同。

¹⁷ 義項未注明出處者，主要參考《漢語大字典》，以下例同。見漢語大字典編輯委員會編，徐中舒主編，《漢語大字典》（武漢：湖北辭書出版社；成都：四川辭書出版社，1986-1990）。

¹⁸ 詹憲慈，《廣州語本字》（香港：中文大學出版社，1995）。

咁（粵音 gam2）¹⁹ U+5481

義項說明及用例：1. 這麼、那麼。例如：枝竹竿要咁長先啲用（竹竿要這麼長才合用）。

2. 多麼。例如：天氣鬼咁熱（天氣多麼熱）。²⁰

詹憲慈《廣州語本字》「恁樣」條下云：

恁者，如此也。俗讀恁若紺，或寫作咁，誤也。《朱子語錄》：「聖人作易，教人去占，占得恁地便吉，恁地便凶。」所謂恁地即恁樣也。廣垌人以咁爲恁，不知恁之爲本字耳。²¹

詹氏認爲「恁」是「咁」的本字，實無必要另寫爲「咁」。今案：雖然詹氏所作推敲頗具學理，但是「咁」粵音讀 gam2，而「恁」粵音讀 jam6，發音部位不同，兼且一爲上聲，一爲去聲，故此爲反映實際口語語音，字庫新增「咁」字仍有其需要。

嘅²²（粵音 ge1）U+5497

義項說明及用例：1. 表示領屬關係。例如：我嘅書（我的書）。²³

2. 用於定語之後。例如：佢都幾聰明嘅（他是蠻聰明的）。²⁴

詹憲慈《廣州語本字》「是我忌」條下云：

忌者，助詞也。俗讀忌音近這。《詩·大叔于田》：「叔善射忌」，《箇云》：「忌讀如彼其之子之己」；〈候人〉：「彼其之子」，《箇云》：「其或作記，或作己，讀聲相似。」《書·微子》：「若之何其」，《鄭注》：「其，語助也。齊魯之間聲如姬。」《禮記》曰：「何居」，《檀弓》注「讀爲姬姓之姬。」忌也、己也、記也、其也、居也，皆助詞也。²⁵

詹氏認爲「忌」是「嘅」的本字，實無必要另寫爲「嘅」。今案：雖然詹氏所作推敲頗具學理，但是「嘅」粵音讀 ge1，而「忌」粵音讀 gei6，收音部位不同，兼且一爲平聲，一爲去聲；更何況「嘅」除了用作語尾助詞之外，另外亦用作結構助詞：置於名詞或代名詞後，表示所屬、所有的關係，故此爲反映實際口語語音，同時「嘅」字又擔負其他語法用途，字庫新增字「嘅」仍有其需要。

¹⁹ 粵音參香港中文大學「粵語審音配詞字庫」，下同。網站：<http://humanum.arts.cuhk.edu.hk/Lexis/lexi-can/>。

²⁰ 參詹伯慧主編，《廣州話正音字典》（廣州：廣東人民出版社，2002），頁242。

²¹ 參詹憲慈，《廣州語本字》卷一，頁5。

²² 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

²³ 參李榮主編，白宛如編纂，《廣州方言詞典·現代漢語方言大詞典·分卷》（南京：江蘇教育出版社，2003），頁27。結構助詞：置於名詞或代名詞後，表示所屬、所有的關係。

²⁴ 參李榮主編，白宛如編纂，《廣州方言詞典》，頁27。句尾助詞：置於句尾，表示肯定或加強的語氣。

²⁵ 參詹憲慈，《廣州語本字》卷一，頁10。

(二) 無本字可循的 Unicode 所見香港特有漢字

香港電腦字庫部分新增字，為日常生活普遍使用的，則未必都有本字可循。

咗 ㉖ (粵音 zo2) U+5497

義項說明及用例：廣府話用在動詞後面，表示動作已經完成，相當於普通話或國語的「了」。^㉗例如：做咗成朝嘅，想早啲食飯（工作了整個早上，想早點吃飯）。

吓 ㉘ (粵音 haa5) U+5413

義項說明及用例：1. 語氣詞。相當於「啊」、「呀」。
2. 嘆詞。
3. 方言。同「下」。一吓，即一下。
4. 「嚇」的簡化字。
5. 廣府話「～吓～吓」，疊用格式，由兩個相同的動詞連著，後隨相關的量詞，表示「一會兒～一會兒～」。例如：洗吓一件，洗吓一件，卒之（終於）將啲衫洗晒。洗吓一件，洗吓一件，卒之將啲衫洗晒（一會兒洗一件，一會兒洗一件，終於把衣服洗完了）。^㉙

以上所舉「咗」、「吓」兩字是廣府話經常用於口語的「語助詞」。

哋 ㉚ (粵音 dei6) U+54CB

義項說明及用例：1. 廣府話人稱代詞，複數。例如：我哋即我們；你哋即你們；佢哋即他們。^㉛
2. 助詞，用在疊音形容詞後面，表示輕微、稍微的意思。例如：
件衫仲係濕濕哋（這件衣服還是有點濕）。^㉜

「哋」字除了用作「語助詞」，最常用為「複數的人稱代詞」，但說「人哋」時，即是「我」的意思，是「第一人稱代詞」的用法。

㉖ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

㉗ 參饒秉才主編，《廣州音字典》（廣州：廣東人民出版社，1992），頁200。

㉘ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

㉙ 參鄭定歐編纂，《香港粵語詞典》（南京：江蘇教育出版社，1999），頁22。

㉚ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

㉛ 參饒秉才主編，《廣州音字典》，頁202-203。

㉜ 參鄭定歐編纂，《香港粵語詞典》，頁100。

𠵼 ③ (粵音 tam5) U+20C41 (p197)

義項說明及用例：字從口、𠵼聲④。𠵼意為小坑。澳門有一處地方叫做「𠵼仔」。

𠵼，是用他人喜歡聽的話使其順從。例如：𠵼細路哥（哄小孩兒）。⑤

嚙 ⑥ (粵音 dyut1) U+2106F

義項說明及用例：翹起嘴巴。例如：有乜嘢唔開心呀，嚙長條嘴（有甚麼不開心的事兒呀，瞧你翹起嘴巴）。⑦

㗎 ⑧ (粵音 loe1 loe2) U+20F31

義項說明及用例：1. 把東西從嘴裡吐出來。例如：㗎口水（吐口水 / 吐痰）。

2. 故意糾纏以達到某種目的。例如：先例不可開，你點㗎都無用㗎（先例不可開，你怎樣磨都不行）。⑨

「𠵼」、「嚙」、「㗎」三個動詞如果不懂得廣府話的人士，一般是無法從字形了解字的音義的。但只要稍加觀察，便不難發現以上三字，都是從「口」的形聲字。「口」是義符，許慎《說文解字》云：「口，人所以言食也」，就是說：口是人賴以說話、飲食的器官。因此，在創造廣府話口語字時，以「口」代表說話義，右邊的「𠵼」、「奪」、「累」就表示字的讀音。與許慎所言六書中形聲字造字——以事為名，取譬相成——的原理，十分融合。

(三) 為他字「異體」的 Unicode 所見香港特有漢字

香港電腦字庫部分新增字，有一些是純粹的「異體字」。「異體字」是指「音、義相同，而字形不同的字」。「異體字」又可分成「異構字」和「異寫字」。「異構字」是「組合部件有所改變造成的」，如「婉」字異構作「挽」，由「女」字改為「子」字。「異寫字」則只是「字形筆畫的改變或增減」，如「帆」、「承」、「爲」等字。

挽 ⑩ (粵音 min5) U+3743

「婉」字異構。意為「分娩」。

③ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

④ 𠵼，《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

⑤ 參鄭定歐編纂，《香港粵語詞典》，頁 197。

⑥ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

⑦ 參鄭定歐編纂，《香港粵語詞典》，頁 354。

⑧ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

⑨ 參鄭定歐編纂，《香港粵語詞典》，頁 60。

⑩ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

帆 ① (粵音 faan4) U+3836

「帆」字異寫。

烝 ② (粵音 sing4) U+2424B

疑為「丞」。「丞」字異寫。

爲 ③ (粵音 wai4 wai6) U+24521

「為」字異寫。

今案：以上四例中，「婉」、「爲」二字被視作「挽」、「為」異寫；事實上從「挽」字字形出現較早，以及「爲」字字形結構更符合古文字構造原意來看，似可商榷。

「婉」字異構作「挽」，由「女」字改為「子」字。從子之「挽」字為說文正字，④從女之「婉」字為唐以後常用字。中華民國標準字採用從女之字，而香港則保留從子之「挽」。「爲」字，甲骨文字形作𦥑，從「爪」從「象」；《說文》小篆作𦥑，「象」字字形訛變，而「爪」旁則保持原貌。由此可見，從子之「挽」與從「爪」的「爲」字，更適合視作正體本字。《香港增補字符集》(HKSCS)收錄了「挽」字，為「婉」字字形來源與演變的考究，保留了十分重要的媒介。

62

六、結語

綜上所述香港電腦特有漢字，除專有名詞用字（如人名、地名、公司名稱），粵方言用字，科學名詞用字，部首或部件外，至少還有小篆隸定字形和異體字等。其實粵方言用字當中，還可分為本有其字的粵方言字，以及本無其字的粵方言字。凡此種種都可以作為研究或發展廣府話電腦新增字時之參考。

但過去 Unicode 以及 ISO 等組織在收錄字符時，只著重字形，而忽略字符的語音及意義，往往將不同來源，甚至音義皆不同的字混為一談。這些「音義不同，而字形相同的字」，文字學界稱之為「同形字」。如上文所述之「𠃔」字，即是一例。作為「一種有機化合物名」的「𠃔」，與許慎《說文解字》所錄義作「衆口也」的「𠃔」，原本是兩個音義皆不同的字，只因字形相同，才混而為一。這種不求甚解的處理文字方式，令使用漢字的人士，難以正確認識漢字形音義三者的關係。

① 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

② 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

③ 《康熙字典》和香港中文大學「粵語審音配詞字庫」未收此字形。

④ 「挽」字字形可上溯戰國楚簡及甲骨文。參見季旭昇，《說文新證》(臺北：藝文印書館，2004)，下冊，頁283。

此外，Unicode 以及 ISO 等組織在收錄字符時，完全不講究字符的用例與出處所本，致令有志人士，無法據之作深入分析、研究，實有欠嚴謹。以上所述均為標準化組織有待改進之處。

最後，還有一個構想，贅述如後。在電腦漢字編碼中，Unicode 具有兼容多國文字的好處，可以預見的是，在全球化的數位世界裡，將會有更多的數位資訊使用 Unicode 作為基本字集。然而，Unicode 3.1 版收錄的七萬餘個漢字中有許多是音義相同、字形不同的異體字，或形體相近的日、韓、越、港等地專用漢字。由於這些字的音義相同，卻有不同的字碼，再加上使用者的用字習慣不同，未來可能造成電腦資訊處理時的諸多困難，如中文域名解析的不精確性，和中文資訊檢索的不完整性等等。故在發展異體字資訊處理技術前，需要整理一套資訊處理用的 Unicode 異體字表和辭典。

在「數位典藏與數位學習國家型科技計畫・數位技術研發與整合計畫」分項下的「系統整合與規範研發計畫・異體字研究」執行期間，工作團隊發現漢字的使用除了中、港、臺、星、馬等中國人和華人地區之外，漢字文化圈的其他國家如日本、韓國以及越南等，皆曾在不同時間使用或借用漢字，且發展出一套屬於自己的漢字系統。鑑於日本、韓國以及越南等地電腦漢字的整理與研究，在網路通訊無遠弗屆的今日，早已經是不可忽略的課題，本國學者如能與日本、韓國以及越南等國漢字專家，共同合作整理研究電腦漢字，必可達到事半功倍之效。

以「正體字」而非「簡化字」，作為建構「國際電腦漢字及異體字知識庫」的主要參照體系，應是目前最合理的作法。以「正體字」分別與「中國簡化字」、「日本漢字」、「韓國漢字」、「越南漢字」，在平等互惠的基礎上作學術交流。此不但符合漢字是以「正體字」在「漢字文化圈」內傳播的歷史事實，也將有利於由臺灣在電腦漢字的發展過程中不會被輕易忽略，進而取得主導地位，對「正體字」的保存可略盡綿薄之力。

除了結合國際漢字專家學者共同參與本計畫的研究工作，各國學者一方面各自在其國家內推動國家標準，一方面與計畫團隊合作與 IETF 和 Unicode Consortium 等國際標準組織聯繫。最終目標期望研發成果，能成為國際社會以及 Unicode 漢字系統後續開發的重要參考依據，是工作團隊所期盼。

後記：本文原題目為：〈Unicode 所見香港特有漢字初探舉隅〉，曾於二〇一〇年七月於臺北中央研究院歷史語言研究所主辦「國際漢字研究與網路技術工作坊」上宣讀，得以聽取與會學者專家的高見。另外，又得兩位學者匿名審查，並提供寶貴修改意見，特此一併致謝！

附表：《香港增補字符集—2008》特有漢字 919 個所見首 30 及最後 30 字⁴⁵
(據 Unicode 編碼順序排列)

國際電腦漢字及異體字知識庫
International Encoded Han Character and Variants Database

查詢結果：共 949 筆資料，從第 0 筆開始顯示

1	U+3440	俵	2	U+34C1	𠂇	3	U+3553	墨	4	U+3618	𦵃	5	U+363E	𡇧
6	U+372D	嫏	7	U+3D8A	瀨	8	U+3DFD	熿	9	U+3ED0	瓊	10	U+3EE9	璫
11	U+3EF4	墁	12	U+3F00	𡔉	13	U+41CA	竑	14	U+43ED	眡	15	U+44C0	苦
16	U+2003E	奩	17	U+2004E	彔	18	U+20068	𠂔	19	U+200EE	𠂔	20	U+20118	承
21	U+201AB	𠮩	22	U+201C1	𠂈	23	U+20214	金	24	U+20274	𠂈	25	U+20275	𠂈
26	U+20299	𠂈	27	U+202BF	𠂈	28	U+202C0	𠂈	29	U+2030A	𢚎	30	U+20325	𠂈

64

920	U+2959E	颯	921	U+29857	馥	922	U+298A4	駟	923	U+298EA	駟	924	U+29903	駟
925	U+29905	駘	926	U+2992F	驪	927	U+2995D	駘	928	U+2996A	駘	929	U+2999D	駘
930	U+29B05	鬱	931	U+29B0E	鬱	932	U+29BD5	鬱	933	U+29CAD	鬱	934	U+29D3E	鬱
935	U+29EAC	鯷	936	U+29EC3	鯷	937	U+29F23	鯷	938	U+29F30	鯷	939	U+2A0E1	鯷
940	U+2A0F8	鷗	941	U+2A150	鷗	942	U+2A192	鷗	943	U+2A193	鷗	944	U+2A1DF	鷗
945	U+2A220	鷗	946	U+2A29F	鷗	947	U+2A2DF	鷗	948	U+2A3A9	鷗	949	U+2A6A9	鷗

⁴⁵ 詳細資料請見網址：<http://chardb.iis.sinica.edu.tw/charlist/hk100.html>。