

電腦用漢字粵語拼音表

1. 主題內容

- 1.1 本文件旨在提供一套漢字的粵語拼音，為電腦的漢字處理提供粵語的語音信息，作為參考，以利粵語應用軟件的開發。
- 1.2 本文件標注的漢字包括國際標準《ISO/IEC 10646-1:2000》和《香港增補字符集 – 2001》(Hong Kong Supplementary Character Set – 2001) 內的全部漢字，共 29,145 個。
- 1.3 本文件內漢字的粵語注音採用香港語言學學會制定的羅馬化《粵語拼音方案》。

2. 參考標準與文獻

2.1 標準

- 2.1.1 ISO/IEC 10646-1:2000(E), International Standard: Information Technology-Universal Multiple-Octet Coded Character Set (UCS) – Part 1: Architecture and Basic Multilingual Plane, ISO/IEC, 2000.
- 2.1.2 GB 12200.1-90《漢語信息處理詞彙 01 部分：基本術語》，國家技術監督局，中國標準出版社，1990 年 2 月。
- 2.1.3 GB2312-80《信息交換用漢字編碼字符集 基本集》，國家標準總局，1981 年 5 月。

2.2 文獻

- 2.2.1 《粵語拼音字表》（第二版），香港語言學學會，2002 年 12 月。
- 2.2.2 〈第二次漢字簡化方案（草案）〉，《現行漢字規範問題》，高更生，商務印書館，2002 年 12 月。
- 2.2.3 《商務新詞典》縮印本，黃港生，商務印書館（香港）有限公司，2002 年 10 月。
- 2.2.4 《新華字典》1998 年修訂本，商務印書館，2002 年 7 月。
- 2.2.5 〈新舊字形對照表〉，《新華字典》1998 年修訂本，商務印書館，2002 年 7 月。
- 2.2.6 《廣州話正音字典》，詹伯慧，廣東人民出版社，2002 年 7 月。
- 2.2.7 《香港增補字符集 – 2001》，香港特別行政區政府，2001 年 12 月。
- 2.2.8 《粵音韻彙》重排本，黃錫凌，中華書局（香港）有限公司，2001 年 9 月。
- 2.2.9 《漢字標準字典》，許嘉璐，遼寧大學出版社，2001 年 8 月。
- 2.2.10 《粵音正讀字彙》，何文匯、朱國藩，香港教育圖書公司，2001 年。
- 2.2.11 《康熙字典》，中華書局，1997 年 2 月。
- 2.2.12 〈國共先後所頒佈的簡體字簡化字異同對照一覽表〉，《海峽兩岸簡體字研究》，路燈光、路燈照，臺灣學生書局，1992 年 11 月。
- 2.2.13 《常用字廣州話讀音表》（一九九二年修定本），香港教育署語文教育學院中文系，香港教育署，1990 年。
- 2.2.14 《新中日簡體字研究》，謝世涯，語文出版社，1989 年 3 月。
- 2.2.15 《李氏中文字典》，李卓敏，中文大學出版社，1989 年 1 月。

- 2.2.16 《廣州話標準音字彙》，周無忌、饒秉才，商務印書館香港分館，1988年2月。
- 2.2.17 《漢語大字典》，漢語大字典編輯委員會，湖北辭書出版社、四川辭書出版社，1986年10月。
- 2.2.18 《印刷通用漢字字形表》，文字改革出版社，新華書店北京發行所，1986年1月。
- 2.2.19 《簡化字總表》，中國文字改革委員會，語文出版社，1986年。
- 2.2.20 《第一批異體字整理表》，中華人民共和國文化部和中國文字改革委員會，語文出版社，1955年12月。

3. 術語

- 3.1 **統一的中日韓文字 CJK Unified Ideographs (ISO/IEC 10646 《信息技術通用多八位編碼字符集 (UCS) 》)**
ISO/IEC 10646 國際標準包括的中國、日本、韓國及其他國家和地區經認同規則整合後所採納的字符的統稱。
- 3.2 **漢字 Chinese Character / Hanzi (GB 12200.1-90 《漢語信息處理詞彙 01 部分：基本術語》)**
原指書寫漢語的符號；在本文件內特指統一的中日韓文字及《香港增補字符集 – 2001》內的字。
- 3.3 **華語漢字**
原指內地、台灣、香港和新加坡所使用的所有漢字；在本文件內特指國際標準《ISO/IEC 10646-1:2000》內接受的由內地、台灣、香港、澳門和新加坡提交的所有漢字或可在《康熙字典》或《漢語大字典》等可查證的漢字，以及《香港增補字符集 – 2001》內的所有漢字。
- 3.4 **非華語漢字**
本文件內指國際標準《ISO/IEC 10646-1:2000》編碼字符集內所有不屬於華語漢字的其他漢字。
- 3.5 **簡化字 Simplified Chinese Character (GB 12200.1-90 《漢語信息處理詞彙 01 部分：基本術語》)**
特指內地於 1986 年公佈的《簡化字總表》(見上文 2.2.19 段) 所列的字。
- 3.6 **類推簡化字**
指沒有在《簡化字總表》內列出，但根據《簡化字總表》的簡化偏旁部首類推出來的簡化字。
- 3.7 **異體字**
狹義：指與《第一批異體字整理表》(見上文 2.2.20 段) 和《新華字典》(見上文 2.2.4 段) 內所列的同音、同義而寫法不同的字。
廣義：包括狹義異體字和《商務新詞典》、《廣州話正音字典》、《廣州話標準音字彙》(見上文 2.2 段) 內所列的同音、同義而寫法不同的字。
- 3.8 **新舊字形**
指包含〈新舊字形對照表〉(見上文 2.2.5 段) 內部件新舊寫法不同的漢字。

- 3.9 **繁體字環境**
指在香港及台灣應用的漢字書寫環境。電腦內以大五碼 (Big5) 編碼的中文平台是典型的繁體字書寫環境。
- 3.10 **簡體字環境**
指內地規範的漢字書寫環境（包括簡化字）。電腦內以國標 (GB2312-80) 編碼的中文平台是典型的簡體字書寫環境。
- 3.11 **訓讀**
指借用同義或近義而讀音不同的字，取其字義而不取字音。
- 3.12 **破讀**
指同一個字形有兩個或以上讀音，而讀音有比較明確的分工。
例：好 (hou2) 壞、喜好 (hou3)
- 3.13 **羅馬化拼音**
指主要用羅馬字母來表示的拼音方案。
- 3.14 **UCS-2 編碼**
指 ISO/IEC 10646 國際標準中以雙八位定義的編碼系統。在使用 UCS-2 編碼表示 ISO/IEC 10646 國際標準中的字符時，以符號 U+XXXX 表示。其中 XXXX 是 UCS-2 編碼的十六進制值，如：
‘一’ 字以 UCS-2 編碼表示為 U+4E00；
‘漢’ 字以 UCS-2 編碼表示為 U+6F22。
- 3.15 **原字集 CJK Ideograph Source (ISO/IEC 10646 Annex S 《信息技術 通用多八位編碼字符集 (UCS) 》附件 S)**
ISO/IEC 10646 國際標準基本多文種平面內的中日韓文字是在整合中國、日本、韓國及其他國家和地區原有的電腦文字標準後得出的。這些原有的電腦文字標準或規範稱為原字集。原字集的頒布國家和地區均有代號：中國內地(G)、香港(H)、日本(J)、韓國(K)、新加坡(S)、臺灣(T)和越南(V)。
- 3.16 **認同規則 Unification Procedure (ISO/IEC 10646 Annex S 《信息技術 通用多八位編碼字符集 (UCS) 》附件 S)**
指 ISO/IEC 10646 國際標準中使用的一套漢字統一整合規則¹。根據認同規則被認同的漢字不會分別編碼，而在不違反認同規則的原則下，ISO/IEC 10646 國際標準不限制某一碼位漢字的具體字形。
- 3.17 **原字集分別編碼原則 Source Separation Rule (ISO/IEC 10646 Annex S 《信息技術 通用多八位編碼字符集 (UCS) 》附件 S)**
凡在原字集中具有各自代碼的異體字不受認同規則限制，在 ISO/IEC 10646 國際標準中予以保留並分別編碼，以保證 ISO/IEC 10646 與原字集之間的對映關係不受影響。
4. **標音原則**
- 4.1 字音分類
- 4.1.1 單音字不予分類。
- 4.1.2 多音字分純異讀字和破讀字兩類。

¹ 認同規則原文見 ISO/IEC 10646 Annex S (Informative) Procedure for the Unification and Arrangement of CJK Ideographs.

- 4.1.2.1 純異讀字的讀音簡單分爲兩類：
無標記 (unmarked) 音：最可接受的讀音
又音：其他讀音
- 4.1.2.2 破讀字的讀音分工可分別標注不同的無標記音，其餘的爲又音。
- 4.2 本表以《廣州話正音字典》、《廣州話標準音字彙》和《商務新詞典》(見上文 2.2 段) 爲主要參考書。凡此三本詞典已有粵音的漢字，本表會收錄或參考該注音；一字多音時，會先將多義多音區分；同義多音時，則以該字的常用性爲大致的排序。
- 4.3 凡收錄在《粵語拼音字表》(見上文 2.2.1 段) 中附有 ‘*’ 號和 ‘#’ 號的粵音，本表一概收錄；另外，本表亦會收錄《粵語拼音字表》所載香港增補字符集內漢字的粵音。
- 4.4 本表標音依循字與字之間的正異、繁簡及新舊關係：
以《第一批異體字整理表》、《新華字典》及三本詞典(見上文 4.2 段) 所列，作爲正異關係的依據，並以《康熙字典》、《漢語大字典》及《漢字標準字典》核實三本詞典的資料；以《簡化字總表》及據以類推出來的漢字作爲繁簡關係的依據；以《印刷通用漢字字形表》、《新華字典》的〈新舊字形對照表〉及據以類推出來的漢字，作爲新舊關係的依據。
- 4.5 凡收錄在《現行漢字規範問題》的〈第二次漢字簡化方案(草案)〉和《海峽兩岸簡體字研究》的〈國共先後所頒佈的簡體字簡化字異同對照一覽表〉(見上文 2.2 段) 及據以類推出來的漢字，會依其對應的有音漢字注音。
例： ‘灑’ (U+3CD5) 作爲 ‘灑’ (U+703E) 的異體字處理；
‘桜’ (U+685C) 作爲 ‘櫻’ (U+6AFB) 的異體字處理。
- 4.6 尙未能標上粵音的漢字，會以《漢語大字典》作參考，以書中標注的第一義第一音作參考，按以下原則標上粵音：
- 4.6.1 等同字注音：
凡有“同某字”關係者，依其等同字之音。
例： 把‘裹’ (U+8901) 標上其等同字‘裹’ (U+88F9) 的粵音；
把‘僕’ (U+5040) 標上其等同字‘英’ (U+82F1) 的粵音。
- 4.6.2 直音注音：
凡古已有之的字有“音某字”關係者，依其同音字之音。
例： ‘𠂔’ (U+4EC8) 字依其直音字‘八’ (U+516B) 標注；
‘𠂔’ (U+5498) 字依其直音字‘布’ (U+5E03) 標注。
- 4.6.3 反切注音：
凡《漢語大字典》有反切的漢字，會以《粵音正讀字彙》的反切及粵音作參考，按以下次序標上粵音：

- 4.6.3.1 如該反切在《粵音正讀字彙》是同切單音，會以此粵音標注。
 例：‘棹’ (U+6932) 是‘于鬼切’，切音是‘wai5’，以此音標注；
 ‘儉’ (U+50CB) 是‘他紺切’，切音是‘taam3’，以此音標注。
- 4.6.3.2 如該反切在《粵音正讀字彙》是同切多音，按以下原則注音：
- 4.6.3.2.1 如該漢字可以判斷為部首加聲旁結構，而其聲旁有音與切音相同，則標上聲旁之音或聲旁之首音。
 例：‘筴’ (U+41EE)、‘桀’ (U+6820) 是‘如甚切’，有切音‘jam5’和‘jam6’，按聲旁‘任’ (U+4EFB) 的音‘jam6’和‘jam4’，標音為‘jam6’；
 ‘蔭’ (U+84A2) 是‘直魚切’，有切音‘ceoi4’和‘cyu4’，按聲旁‘除’ (U+9664) 的首音‘ceoi4’標注。
- 4.6.3.2.2 如該漢字可以判斷為部首加聲旁的結構，其聲旁之音與切音不符，會參考有相同聲旁的漢字或字形相似的字作類比注音。
 例：‘祛’ (U+7D36)、‘祛’ (U+3B55)、‘芸’ (U+44A7) 是‘去魚切’，有切音‘heoi1’和‘keoi1’，與聲旁‘去’ (U+53BB) 的粵音‘heoi3’、‘heoi2’不符。因此，以‘佉’ (U+4F49)、祛 (U+547F)、祛 (U+62BE)、祛 (U+80E0)、祛 (U+795B) 作類比，標音為‘keoi1’；
 ‘庖’ (U+5E89) 是‘徒損切’，有切音‘deon6’和‘teon5’，與聲旁‘屯’ (U+5C6F) 的粵音‘tyun4’和‘zeon1’不符。因此，以‘囤’ (U+56E4)、沌 (U+6C8C)、炖 (U+7096)、鈍 (U+920D) 作類比，標音作‘deon6’。
- 4.6.3.2.3 如該漢字可以判斷為部首加聲旁的結構，但不能以聲旁或類比注音，會參考《康熙字典》‘同某字’或‘音某字’關係確定其反切注音。
 例：‘攄’ (U+63D3) 是‘施智切’，有切音‘si3’和‘ci3’，與聲旁‘施’ (U+65BD) 的粵音‘si1’和‘ji6’不符；‘攄’ (U+6938)、箴’ (U+7BB7) 的粵音為‘ji4’，不能作類比。
 《康熙字典》注‘音翅’，故以‘翅’ (U+7FC5) 粵音‘ci3’為‘攄’標注。
- 4.6.3.2.4 如該漢字無確定的聲旁，會參考《古今漢字粵音》²稿本的標音。
- 4.6.3.3 如該反切沒有收錄在《粵音正讀字彙》中，會參考《古今漢字粵音》稿本的標音。
- 4.7 未能參考《漢語大字典》標音的漢字，會以《漢字標準字典》的等同字關係注音。

² 《古今漢字粵音》由張日昇、張群顯、麥耘編著，未發表。

例： 把‘數’(U+6573)標上其等同字‘鼓’(U+6571)的粵音；
把‘數’(U+656F)標上其等同字‘警’(U+668B)的粵音。

4.8 無任何注音可參考者，按以下之方法注音：

4.8.1 以全字形比較判斷注音

4.8.1.1 參考繁簡關係注音

4.8.1.1.1 凡有簡化部件之簡體字，若其對應之繁體字不在表中因而未能按繁簡關係處理，會以《漢語大字典》查證其繁體字之音而注音（見上文 4.6 段）。

例： ‘乾’(U+4880)是簡體字，但‘車乞’不在表中。故以《漢語大字典》查證‘車乞’字之音而注音。

4.8.1.1.2 繁體字若含兩個或以上可簡化之部件，而另一字只簡化其中部分部件，此字作為其對應之繁體字的俗簡字處理，並依該繁體字之音而注音。

例： ‘鯪’(U+9BF5)的右旁‘參’是‘參’的簡化字，而左旁‘魚’可作簡化但沒有簡化作‘魚’，故把‘鯪’作‘鯪’(U+9C3A)的俗簡字處理，依‘鯪’注音。

4.8.1.1.3 凡漢字有部件為不可作簡化偏旁的簡化字，會以該部件類推其對應之繁體字，作俗簡字處理，並依該繁體字之音而注音。

例： ‘妘’(U+E0FD)的右旁‘么’是‘麼’的簡化字。但‘么’是不可用作簡化偏旁的簡化字。故把‘妘’作‘嫪’(U+5B24)的俗簡字處理，依‘嫪’注音。

4.8.1.1.4 凡漢字含懷疑錯誤類推的簡化部件，會以該部件推論其對應之繁體字，作俗簡字處理，並依該繁體字之音而注音。

例： ‘鈺’(U+9229)的右旁‘戶’，是簡化字‘戶、芦、炉、驴’中所簡化的部件。與相應的未簡化字‘廬、蘆、爐、驢’比較，‘戶’或會錯誤類推為部件‘廬’的簡化。故把‘鈺’作‘鑪’(U+946A)的俗簡字處理，依‘鑪’注音。

4.8.1.2 依類似的字形類推注音

4.8.1.2.1 凡漢字有部分變形，作其正形的古俗異體或類推正體關係處理，依其正形的字注音。

例： ‘僦’(U+3DDB)按其正形的字‘僦’(U+7172)注音；
‘蕪’(U+3DFC)按其正形的字‘蕪’(U+71D5)注音。

4.8.1.2.2 凡部首相同而聲旁部分為正異關係的漢字，可類推作正異體關係處理，依其正體字注音。

例： ‘涵’(U+3D20)與‘函’(U+6DB5)近似，而《康熙字典》135頁註‘函別作函’，故‘涵’可作為‘函’的異體處理。

4.8.1.2.3 依有相同部件移位組合的漢字注音。

如‘峰’(U+5CF0)，部件‘山’由左方移到上方成‘峯’(U+5CEF)，意思、讀音一樣；另外如‘棋’(U+68CB)，部件‘木’由左方移到下方成‘碁’(U+68CA)，意思、讀音一樣。

例：‘堪’(U+364B)字中的部件‘土’從‘其’的左方移至下方，作為‘基’(U+57FA)的俗異處理；
‘點’(U+3E03)為‘黠’(U+9EDE)的下部四點的移展，故作為‘黠’的俗異處理；
‘𪔐’(U+3CA9)為‘髡’(U+6BFE)的左右部件移位，故作為‘髡’的俗異處理；
‘鴛’(U+4CB6)為《康熙字典》1483頁中‘元鳥’的部件移位，故作為‘元鳥’的俗異處理（標音作‘元’）。

4.8.1.2.4 依字形最相似的漢字注音。

例：‘熊’(U+3DF1)與‘熊’(U+718A)字形相似，唯右邊有異，故作‘熊’之異體字處理；
‘寔’(U+3684)與‘寔’(U+7590)字形相似，故作‘寔’之異體字處理；
‘尋’(U+3777)與‘尋’(U+5C0B)字形相似，故作‘尋’之異體字處理；
‘外’(U+3688)與‘外’(U+5916)字形相似，故作‘外’之異體字處理。

4.8.2 以部分字形比較判斷注音

4.8.2.1 凡可被判斷為部首加聲旁的漢字，會依其判斷聲旁注音。

例：‘楞’(U+3B76)、‘楞’(U+3B77)、‘榧’(U+3BE7)、‘榧’(U+3BBC)都有常見部首‘木’，因此判斷‘木’為此類字形的部首，而‘芳’(U+82B3)、‘男’(U+7537)、‘超’(U+8D85)、‘扇’(U+6247)為其聲旁，故以此聲旁注音；
‘媯’(U+36E2)、‘媯’(U+3700)、‘媯’(U+3701)、‘媯’(U+371F)都有常見部首‘女’，因此判斷‘女’為此類字形的部首，而‘秀’(U+79C0)、‘相’(U+76F8)、‘查’(U+67E5)、‘敢’(U+6562)為其聲旁，故以此聲旁注音。

4.8.2.2 凡判斷聲旁為簡化部件或簡化字的漢字，而表中沒有其繁體之字，會依其判斷聲旁之繁體字注音。

例：‘埔’(U+364C)在表中並沒有相對的繁體字‘土肅’，故會以其判斷聲旁‘肅’的繁體字‘肅’(U+8085)注音；
‘姪’(U+36C7)在表中並沒有相對的繁體字‘女莊’，故會以其判斷聲旁‘庄’的繁體字‘莊’(U+838A)注音。

- 4.8.2.3 凡判斷聲旁為簡化部件或簡化字的漢字，而表中沒有其或其聲旁之繁體字，會依其聲旁的近似字注音。
 例： ‘嘖’ (U+35F7) 的右旁與 ‘賽’ (U+8D5B) 近，故從 ‘賽’ 音；
 ‘獮’ (U+3E99) 的右旁與 ‘實’ (U+5BE6) 近，故從 ‘實’ 音。
- 4.8.2.4 從其近似部件注音
 例： ‘才’ (U+3427) 從 ‘才’ (U+624D)；
 ‘忒’ (U+3773) 從 ‘忒’ (U+38FA)。
- 4.8.3 無既定理據注音：需專家小組討論通過。
 例： ‘蠱’ (U+3F53)、‘犇’ (U+E1A4)、‘𦉳’ (U+E657)
- 4.9 一些特殊漢字可以多音節字標注類別如下
- 4.9.1 表示度量單位的雙音節字，如：‘𠵼’ (U+74E9) (千瓦) 及 ‘𠵼’ (U+55E7) (加侖)。如某字兼具單音，以該單音作為無標記 (unmarked) 音。
- 4.9.2 ‘卅’ (U+5345) 及 ‘卌’ (U+534C) 兩字由於涉及粵語口語讀法，所以本表沒有依從辭書所注的單音節音，而賦予實際的雙音節發音。
- 4.10 訓讀處理：凡參考辭書有載的訓讀，本表均會收錄。
- 4.11 部件的注音處理：
- 4.11.1 在工作範圍內的字符如為漢字的基本部首、部件、或非獨立漢字（統稱部件）³，會予以獨立分類。
- 4.11.2 字形既是獨立漢字又是部件（含部首），除已注漢字音外，會附加標注為部件類，並以該漢字的首選音作為部件的名稱。
 例： ‘金’ (U+91D1) 既是獨立漢字（音 ‘gam1’），又是部件，部件名稱為 ‘金’（音 ‘gam1’）；
 ‘大’ (U+5927) 既是獨立漢字（音 ‘daai6’ 和 ‘taai3’），又是部件，部件名稱為 ‘大’（取首選音 ‘daai6’）。
- 4.11.3 字形如非獨立漢字或有香港常用叫法，會另外賦予名稱並以名稱注音。
 例： ‘𠵼’ (U+8FB6) 使用香港常用叫法，稱作 ‘撐艇’ (caang1 teng5)；
 ‘𠵼’ (U+91D2) 使用香港常用叫法，稱作 ‘金字邊’ (gam1 zi6 bin1)；
 ‘丶’ (U+4E36) 的漢字音是 ‘zyu2’，部件名稱是 ‘點’ (dim2)；

³部首以《康熙字典》內部索引表為依據。獨立漢字：字元獨立成字，ideograph proper

‘勹’ (U+52F9) 的漢字音是 ‘baau1’，部件名稱是 ‘包字頭’ (baau1 zi6 tau4)。

4.11.4 如果該部件不是獨立漢字而且沒有既定讀音，會由專家根據部件的形而賦予名稱。

例： ‘𠂔’ (U+EEB7) 的名稱是 ‘開口皿’ (hoi1 hau2 ming5)；
‘𠂔’ (U+F30A) 的名稱是 ‘三節豎’ (saam1 zit3 syu6)。

4.12 非華語漢字處理：

4.12.1 非華語漢字與華語漢字字形有異，但有華語漢字可考，依該華語漢字注音。

例： ‘塩’ (U+5869) (‘鹽’ (U+9E7D))，標粵音 ‘jim4’；
‘壹’ (U+58F1) (‘壹’ (U+58F9))，標粵音 ‘jat1’。

4.12.2 以形聲注音（見 4.3.3.2 段）。

例： ‘樁’ (U+697F) (‘香’ (U+9999))、
‘靄’ (U+825D) (‘雪’ (U+96EA))、
‘𠂔’ (U+3407) (‘九’ (U+4E5D))、
‘𠂔’ (U+3409) (‘丁’ (U+4E01))

4.12.3 無既定理據注音：需專家小組討論通過（見 4.8.3 段）。

例： ‘𠂔’ (U+3402)、‘𠂔’ (U+3403)

4.13 首選音（default 音）：

4.13.1 為配合電腦的文字語音轉換系統，必須每一個字給予一個首選音。每一個漢字只有一個首選音。

4.13.2 就純異讀字來說，首選音就是無標記音。

4.13.3 破讀字的首選音：

4.13.3.1 破讀字以該字最常用的讀音分工（一般為辭書排列在首的讀音分工）的無標記音為該字的首選音。

4.13.3.2 若發現辭書有錯或可疑之處，將提出討論。

4.14 音序：

4.14.1 多音字注音的排序原則上以音的使用率為據。

4.14.2 多音字破讀可按與意義無關和與意義有關兩個層面區分。

4.14.3 破讀與意義無關（例：希臘），則常用度高的定為無標記音，其他為又音。

4.14.4 破讀與意義有關時（例：快樂、音樂），可分別標注為不同的無標記音。

4.15 繁簡環境處理：

有些漢字在簡體字環境下的發音與繁體字環境下發音會有不同。例如：‘乾’ (U+4E7E) 字，在簡體字環境只可讀作 ‘kin4’（乾坤），但在繁體字環境可讀作 ‘gon1’（乾燥）及 ‘kin4’。因此本表會標示這類字發音的簡繁體環境，並採用《新華字典》及《簡化字總表》作為繁簡字體的依據。類推簡化字則以《簡化字總表》內的類推規則作依據。

4.16 異讀處理細則

4.16.1 傳統系統性異讀

- 4.16.1.1 n- → l- : 以 n- 為無標記音
- 4.16.1.2 ng- → 零- : 以 ng- 為無標記音 (涉變調⁴另議)
- 4.16.1.3 零- → ng- : 以零- 為無標記音
- 4.16.1.4 gw/kw → g/k / [ɔ:]⁵ : 以 gw/kw- 為無標記音
- 4.16.1.5 gw/kw → g/k / [u:]⁶ : 以 g/k- 為無標記音
- 4.16.1.6 ng → m : 以 ng 為無標記音

4.16.2 半系統性異讀

4.16.2.1 文白

4.16.2.1.1 文白分工型：處理同 4.16.2.1.2 或 4.16.2.1.3

例：坐 (U+5750) zo6~co5

4.16.2.1.2 常文型（常規）：以文讀為無標記音

例：零 (U+96F6) ling4~leng4

4.16.2.1.3 常白型（例外）：以白讀為無標記音

例：井 (U+4E95) zing2~zeng2

4.16.2.2 形態 (morphological) 變調

4.16.2.2.1 T3-6 → T2

4.16.2.2.1.1 一般：以 T3-6 為無標記音

例：情 (U+60C5) cing4~2

4.16.2.2.1.2 久變失繫：以 T2 為無標記音

例：橙⁷ (U+6A59) caang4~2

4.16.2.2.2 T2-6 → T1 : 以 T2-6 為無標記音

例：尾 (U+5C3E) mei5~1

4.16.2.2.3 Tx → T4

4.16.2.2.3.1 疊字親屬稱謂：以 Tx 為無標記音

例：弟 (U+5F1F) dai6~4

4.16.2.2.3.2 其他⁸：以 Tx 為無標記音

例：芝 (U+829D) zi1~4

4.16.2.3 文白異讀涉 T2 形態變調

例：名 (U+540D) ming4~meng2 : 以文讀 T3-6 為無標記音

4.16.2.4 T3→T5 : 以原調 (可 T3 或 T5) 為無標記音

⁴ 例：額 ngaak6→2

⁵ 即 o, ong, ok 這幾個韻，不包括 ou 韻。

⁶ 即 u, ui, un, ut 這幾個韻，而不包括 ung, uk, um, up 這些韻。

⁷ 同類例子有“絨、院、柚”等。

⁸ 只限人名的一種親切化叫法。

- 例： 試⁹ (U+8A66) si3~5
- 4.16.2.5 陽調變陰：須逐字考慮
例： 演¹⁰ (U+6F14) jin5~2
- 4.16.2.6 不送氣變送氣：須逐字考慮
例： 昆 (U+6606) g~k(wan1)
- 4.16.2.7 辨義破讀：須逐字考慮¹¹
例： 樂 (U+6A02) lok6~ngok6
- 4.16.3 零散異讀：須逐字考慮
- 4.16.3.1 純異讀
例： 縷 (U+7E37) lei5~lau5
- 4.16.3.2 涉其他因素異讀
- 4.16.3.2.1 涉“配詞”破讀
例： (希) 臘 (U+81D8) laap6~lip6
- 4.16.3.2.2 涉文白異讀
例： 在 (U+5728) zoi6~coi5
- 4.16.3.2.3 涉形態變調
例： 調 (U+8ABF) diu6~2

5. 《粵語拼音表》

本表列載合共 29,145 個漢字，包括國際標準《ISO/IEC 10646-1:2000》和《香港增補字符集 – 2001》(Hong Kong Supplementary Character Set – 2001) 內的全部漢字。

- 5.1 漢字排列方式
- 5.1.1 漢字按 UCS-2 編碼從小到大順序排列。
- 5.2 欄目說明
- 5.2.1 **ucs2**：UCS-2 編碼的十六進制值。
- 5.2.2 **ch**：字符。
- 5.2.3 **typ**：漢字類型。如多於一類，以分號 ‘;’ 分隔。
‘<’：《簡化字總表》內列出的簡化字。
‘#’：不在《簡化字總表》內的類推簡體字。
‘>’：在《簡化字總表》內列出的被簡化的字，即與 ‘<’ 類相對應的繁體字。

⁹ 同類例子有“舍、舅、臼”等。

¹⁰ 同類例子有“忍、繞、擾、阮、戀”，另去聲有“撰、捷”。

¹¹ 本小組判斷為最常用的音為 default/首選(以別於 unmarked/無標記)。

- ‘S’：不在《簡化字總表》內的可被類推簡化的繁體字，且表內有相對應的簡體字，即與‘#’類相對應的繁體字。
- ‘\$’：不在《簡化字總表》內的可被類推簡化的繁體字，表內沒有相對應的簡體字。
- ‘*’：《第一批異體字整理表》列出的異體字。
- ‘**’：據《新華字典》，不在《第一批異體字整理表》的異體字。
- ‘***’：不在《第一批異體字整理表》及《新華字典》的異體字。
- ‘****’：在《漢語大字典》中列明是某字的訛字。
- 5.2.4 **rg**：該字所屬區域的編號。
- ‘1’：區域 1，範圍包括以原則 4.2 至 4.4，或 4.9 所標注粵音的所有漢字。
- ‘2’：區域 2，範圍包括以原則 4.5 至 4.7 所標注粵音的所有漢字。
- ‘3’：區域 3，範圍包括以原則 4.8 或 4.12 所標注粵音的所有漢字。
- 5.2.5 **rad**：該字是否部件（包括部首）；如無任何標示，則不是部首或字形部件。
- ‘部’：是部件。
- 5.2.6 **ref**：列出之字音是出自那些參照字（附 UCS-2 編碼的十六進制值）。如參照字不在表內，則以‘@’符號表示。
- 5.2.7 **en**：字音用於那個環境；如無任何標示，則繁簡環境相同。
- ‘t’：繁環境。
- ‘s’：簡環境。
- 5.2.8 **pn**：字音編號：由 1 開始。比較常用的，用小編號。首音為首選音。
- 5.2.9 **jyutping**：粵語拼音。
- 5.2.10 **cl**：粵音類別；如無任何標示，則為無標記音。
- ‘又’：又讀。
- 5.2.11 **sc**：異讀類別，以原則 4.16 內的編碼標示。