

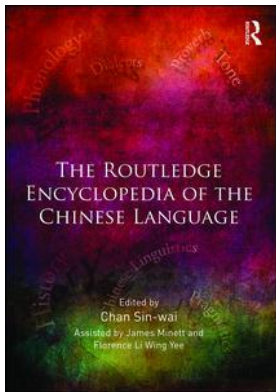
This article was downloaded by: 10.3.97.143

On: 13 Aug 2023

Access details: *subscription number*

Publisher: *Routledge*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: 5 Howick Place, London SW1P 1WG, UK



The Routledge Encyclopedia of the Chinese Language

Sin-wai Chan, James Minett, Florence Li Wing Yee

Cantonese Romanization

Publication details

<https://www.routledgehandbooks.com/doi/10.4324/9781315675541.ch3>

Siu-Pong Cheng, Sze-Wing Tang

Published online on: 15 Feb 2016

How to cite :- Siu-Pong Cheng, Sze-Wing Tang. 15 Feb 2016, *Cantonese Romanization from: The Routledge Encyclopedia of the Chinese Language* Routledge

Accessed on: 13 Aug 2023

<https://www.routledgehandbooks.com/doi/10.4324/9781315675541.ch3>

PLEASE SCROLL DOWN FOR DOCUMENT

Full terms and conditions of use: <https://www.routledgehandbooks.com/legal-notices/terms>

This Document PDF may be used for research, teaching and private study purposes. Any substantial or systematic reproductions, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The publisher shall not be liable for an loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

3

CANTONESE ROMANIZATION

Cheng Siu-Pong and Tang Sze-Wing

THE CHINESE UNIVERSITY OF HONG KONG, HONG KONG, CHINA

The diversity of Cantonese Romanization is a reflection of different backgrounds and purposes in rendering Cantonese into Roman letters. Some forms of Romanization are marked for their historical significance and influence on later proposals; others are noteworthy for their current popularity, as they are often found in books and other materials. Individual frameworks are discussed in detail, and the concluding tables summarize these discussions. It follows that such diversity compensates for the lack of a predominant form of Cantonese Romanization.

1. Introduction

Cantonese Romanization methods are as varied as they are numerous. Up to the end of the last century, over 30 varieties have existed (Wu 1997; Lam 2009). None of them, however, is in a predominant position as *pinyin* is to Mandarin Chinese. Lack of government-sponsored standardization programs is part of the reason. The general public in Cantonese-speaking regions also cannot reach a consensus on how a Cantonese word should be Romanized. Unlike Mandarin or Vietnamese, the use of Chinese characters to write Cantonese has not been subjected to challenge by the proponents of the Roman alphabet, and no serious attempt has ever been made to change the written form of Cantonese into an alphabetical system. The main function of Cantonese Romanization is to transcribe Cantonese for educational and research purposes. Another major function is to transliterate Cantonese terms into English or other European languages. Most Romanization systems aim at the former function, and different schemes have been introduced to cater for different target recipients. The following section first describes the sound system of Cantonese. After that, individual Romanization frameworks will be discussed. If looking for a shortcut, you can move to the Romanization tables in the concluding section.

2. Cantonese phonology

There are 19 consonants in modern-day standard Hong Kong Cantonese. These include plosives /p, p^h, t, t^h, k, k^h, k^w, k^{wh}/, nasals /m, n, ŋ/, affricates /ts, ts^h/, fricatives /f, s, h/, and ‘semi-vowels’ as well as other approximants /j, w, l/. These sounds are considered *phonemes*

because they change the meaning of a word when one consonant is replaced by another. Take the (un)aspirated bilabial stops as examples. Unlike English, aspirated /p^h/ (like *p* in *pie*) and unaspirated /p/ (like *p* in *spy*) are two phonemes, and a minimal pair between them can be found, e.g. /p^ha:55/ ‘crouch’ vs. /pa:55/ ‘daddy’. Other plosives are the same and as a result there are two sets of plosives in Cantonese, one aspirated and another unaspirated. /ts, ts^h/ may be realized as [tʃ, tʃ^h], but they are either *allophones* (when the following vowel is rounded, especially when it is /y:/) or simply free variants of /ts, ts^h/, depending on the speaker’s accent. All of these consonants are syllable-initial. Nasals are also syllable-final, and syllable-final plosives /p, t, k/ are unreleased, expressed as [p̚, t̚, k̚]. The 19 consonants reflect the *standard* pronunciation. Many native speakers do not articulate /k^w, k^{wh}, n, ŋ/ syllable-initially and /k, ŋ/ syllable-finally (except in rhymes /ek, ok, eŋ, oŋ/; see Bauer and Benedict 1997). This is of particular interest as far as Romanization is concerned, because many Cantonese dictionaries and textbooks consider the drop or replacement of these sounds to be *non-standard* or even ‘sloppy’ and retain the standard way of pronunciation.¹

Vowels in Cantonese include both front and back, and high and low monothongs /i:, e, y:, u:, o, ɛ:, ɔ:, œ:, ə, ɐ, a:/. Depending on which analysis is adopted, the use of the vowel length mark may be optional or conditional,² and even the total number of vowels can be reduced as some of them occur in *complementary* phonetic contexts. Whether or not they are deemed phonemes depends on how these sounds are perceived. For example, /e/ may lose its phonemic status and be argued as an allophone of /i:/, since they are not only similar but also complementary: /e/ before /k, ŋ/ and /i:/ elsewhere. This is the same for the sound pair of /o, u:/. In the standard, conservative accent, the sound pair of /œ:, ə/ follows a similar distribution, although /œ:/ can stand alone without a final consonant and monothong /ə/ is always followed by /t/ or /n/. This part being introductory, the academic debate associated is set aside. What concerns us is the fact that many Romanization schemes use a single letter for both sounds of the pair. In addition to these monothongs, there are also ten gliding vowels in Cantonese: /iw, ɔy, uɣ, ej, ɔj, ow, ɛj, ɐw, a:j, a:w/. In some analyses, the final glides are rendered as a vowel symbol. Notice that in the same syllable, a gliding vowel (or diphthong) is not followed by a consonant. Only a monothong precedes a consonant coda. For this reason, the final glide in a gliding vowel is often considered as a coda as well. The vowels above are all phonemic, with the apparent ‘exceptions’ of /ɔy, uɣ/. Minimal pairs can be found only when they are preceded by /k/, and also /k^h/ if tones are disregarded. For the sound pairs of /ɛj, a:j/ and /ɐw, a:w/, they are clearly distinctive, and it is easy to find a minimal pair to differentiate the two members within. But given that they are similar in sound quality, these distinctive sounds, together with /ɔy, uɣ/, are sometimes represented by a single *grapheme* when they are rendered in Romanization. This will be further addressed in the following sections. In Cantonese, a syllable always contains a vowel, with the exception of two *syllabic* nasals /m, ŋ/. They can stand alone without a vowel to follow, in which case they are rendered as [m̩, ŋ̩].

Cantonese is a tonal language. There are six or nine tones in Cantonese, depending on one’s approach to categorize them. Traditionally, tones in Chinese are classified into Flat, Rising, Departing, and Entering. In Cantonese, the first three groups contain two tones (one Dark and one Light), and together with three Entering tones (Upper Dark, Lower Dark, Light), they amount to a total of nine tones. For the non-Entering tones, Cantonese has Dark Flat tone /55/, Light Flat tone /21/, Dark Rising tone /25/, Light Rising tone /23/, Dark Departing tone /33/, and Light Departing tone /22/.³ For the Dark Flat tone, there is also a variant which is high-falling /53/. The two Dark Flat tones are either not distinctive (Cheung 1986), or in fact, the falling variant might have been fading out altogether (Matthews and

Yip 1994). Therefore, the maximum number of tones in Cantonese is still said to be nine. For the Entering tones, they specifically refer to the tones of those syllables ended with an unreleased stop [p̚, t̚, k̚], and the Upper Dark, Lower Dark, and Light Entering tones are given as /5/, /3/, and /2/, which respectively correspond to the Dark Flat, Dark Departing, and Light Departing tones. Given this, these Entering tones are often not viewed as separate tones, and as such, Cantonese is said to have six tones only. The Entering tones are only viewed as the short or extra-short variants of the other tones (Zee 1999). Different Romanization schemes variably adhere to the six-tone and nine-tone categorization, and this gives rise to a noticeable difference in how a tone is transcribed. It has also been observed that some native speakers merge some of these tones, especially the two Rising tones /25/ and /23/ (Bauer et al. 2003). This is again seldom reflected in dictionaries or textbooks with words transcribed in Cantonese Romanization. The term *Cantonese* itself, when used in the Hong Kong variety, is literally translated as ‘Yue language’ or ‘Guangdong speech’. Regardless of tone merger or any pronunciation deemed ‘non-standard’, the phonetic notation for the two words is given as /jyt2 jy:23/ and /kʷɔŋ25 tɔŋ55 wa:25/ respectively.

3. Various Romanization systems

3.1. Earlier Romanizations

The publication of Cantonese dictionaries in the early nineteenth century, intended for the Western readers, marked the beginning of a time when substantial work was done to Romanize Cantonese. Robert Morrison’s *A Vocabulary of the Canton Dialect* in 1828 was one of these earliest attempts to offer transcription for Cantonese words in Roman letters. Like many other dictionaries and glossaries that followed, every effort was made to ensure that the Romanization given resembled the native language of the target readers. For example, /i:/ and /u:/ are transcribed as *ee* and *oo*, written in a way in conformity with the English orthography. Letter *u* is used for /ɤ/, which only occurs in closed syllables. For instance, /sɛm/ ‘heart’ is rendered as *sum*. At this stage, there was no one-to-one mapping between letters and sounds. For the same letter *u*, when it is used in open syllables, it refers to /y:/. For example, /tɕy:/ ‘pig’ is given as *chu*. But when /y:/ occurs in a closed syllable, it is rendered as *ue*, as in *uet* /jyt/ ‘moon’. Besides, Morrison’s Romanization omits the indication of tones. Therefore, even though /tʰin55/ ‘sky’ and /tʰin21/ ‘field’ are in different tones, they are all Romanized in the same way as *teen*. Aspiration is not addressed either. For instance, unaspirated /t/ is indistinguishable from aspirated /tʰ/, so *teen* may be used for /tʰin/ as well as /tin/, as in /tin22/ ‘lightning’. Readers cannot know whether it refers to sky or field or lightning simply through this Romanization.

Several Cantonese dictionaries and textbooks emerged after Morrison’s work. These include E. C. Bridgman’s *Chinese Chrestomathy in the Canton Dialect* (1841), S. W. Williams’ *A Tonic Dictionary of the Chinese Language in the Canton Dialect* (1856), J. Chalmers’ *An English and Cantonese Pocket-dictionary* (1859), E. J. Eitel’s *A Chinese Dictionary in the Cantonese Dialect* (1877), and J. D. Ball’s *Cantonese Made Easy* (1883). Compared with Morrison’s Romanization, correspondence between letters and phonemes was improved in later schemes. In Bridgman’s dictionary, for example, distinction was made between aspirated and unaspirated stops and affricates by using inverted apostrophes. Therefore, /t/ and /tʰ/ are represented by *t* and *t'*, and the Romanized forms for /tʰin/ ‘sky’ and /tin/ ‘lightning’ came to be distinctive. Another feature is the use of semicircular marks to represent tones. This is not an invention by Western lexicographers, but adoption of the Chinese traditions. The Flat,

Rising, Departing, and Entering tones are indicated by a semicircular mark placed at four different corners: lower left for Flat, upper left for Rising, upper right for Departing, and lower right for Entering. The Dark and Light tones are differentiated by marking the Light tones with a line underneath the semicircle. For instance, /t^hin55/ ‘sky’ is rendered as *t’in*, whereas for /t^hin21/ ‘field’, the tone mark is further underlined. Ball’s *Cantonese Made Easy* followed the same methods to mark aspiration and tones, and by the same token also achieved a higher degree of phonemic distinction. B. F. Meyer and T. F. Wempe’s *The Student’s Cantonese–English Dictionary* (1934) also has a Romanization system similar to Eitel’s and Ball’s. Comparatively speaking, Y. R. Chao’s *Cantonese Primer* (1947) was a more recent work, yet his Romanization scheme is worth mentioning for historical purposes. Chao’s system is fully alphabetical in that tones are also represented by letters. Different sets of initial consonant letters are used for Dark and Light tones and, for each set, the representation of the same rhyme changes with different tones. To illustrate, the Dark /pa:55/, /pa:25/ and /pa:33/ are represented as *pa*, *pax*, and *pah*, and their Light counterparts are written as *ba*, *bax*, and *bah*. Under Chao’s system, /jyt2 jy:23/ and /k^wɔŋ25 tɔŋ55 wa:25/ are expressed as *Yutyux* and *Kwoagtong-wah** respectively. Notice that Chao’s system is particular about whether the tone is the base tone or changed from another one. Since /wa:/ here is originally in the Light Departing tone /22/, it is so expressed in its original tone as *wah*, rather than as *uax*, with an asterisk added to mark its change to a rising tone. Chao’s Romanization is historically significant as it is arguably the most complex scheme ever proposed for Cantonese Romanization. It has fallen into oblivion though.

3.2. Government Romanization

In Hong Kong, the Government Romanization is the most recognizable means to Romanize Cantonese. It is mainly used for local place names and personal names. To many Cantonese speakers in Hong Kong, this is the only Cantonese Romanization they can understand, and perhaps provide if they are asked to. As previously suggested, it would be dubious to call it a ‘system’ (Kataoka and Lee 2008); it is rather a hodgepodge of various ways of Romanization (Kataoka 2014). For obvious reasons, this is not a Romanization scheme – there is a low degree of correspondence between letters and sounds. The same letter is often used for different sounds. This is likened to the English orthography, in which letter *c* represents /s/ in *cede*, but /k/ in *candy*. In this form of Romanization, the four pairs of Cantonese plosives /p, p^h/, /t, t^h/, /k, k^h/, /k^w, k^{wh}/ are rendered the same way as *p*, *t*, *k*, and *kw*, regardless of whether they are aspirated. Given that these sounds are phonemes, distinctive words such as /tin55/ ‘insane’ and /t^hin55/ ‘sky’ would become indistinguishable when they are Romanized as *tin*. The same also applies to affricates /ts, ts^h/, but there is one more complication: while the two consonants are Romanized the same way, they are variably rendered as *ts* or *ch*. Thus, although the first character for the place names of *Chai Wan* /ts^haj21 wan55/ and *Tsuen Wan* /ts^hyn21 wan55/ share the same initial consonant, they are written in different initial letters.⁴ Nasals /m, n, ŋ/, fricatives /f, s, h/, and approximants /w, l/ are straightforwardly Romanized as *m*, *n*, *ng*, *f*, *s*, *h*, *w*, *l*. As in English, *y* is used for /j/.

The Hong Kong Government Romanization is a *defective* script, and in fact it is defective in both directions – devoid of one-to-one correspondence from grapheme to phoneme, as well as from phoneme to grapheme. This resembles the English orthography more than, for example, the French orthography. The one-to-many mapping from letters to sounds is illustrated in the case of (un)aspirated consonants. And phoneme-to-grapheme inconsistency

is shown in the affricates and more often in the case of some vowels. Take /i:/ as an example. It can be Romanized differently in the place names, as *i* for *yi* in *Tsing Yi*, as *ee* for *see* in *Sai See Street*, as *z* for *tsz* in *Tsz Wan Shan*, and as *ze* for *sze* in *Sze Tei Shan*. Personal names transliterated from Cantonese show the same inconsistency. /tsy:55/, a common surname, is rendered as *Chu*, but /tsyn/ would likely become *Chuen* or *Tsuen* when transliterated. On the other hand, *Chun* (as in the name of Chief Executive *Leung Chun-Ying*) is reserved for /ts^hɛn/. Added to this is the fact that not all letter combinations are legitimate, even though they may represent a legitimate syllable. This is again likened to the English orthography, in which both letters *s* and *c* can refer to /s/, but only *see*, rather than *cee*, is accepted as a valid written unit. Therefore, /tsɛn/ is expressed as *Chun*, but *Yun* for /jɛn/ is unlikely. The adding up of these instances of inconsistency implies that many who know this Romanization have had to learn it by rote, no matter whether they are native Cantonese speakers. Only generalizations can be made here: (i) correspondence between monothongs and single vowel letters: /aɪ, ɐ/ as *a*; /ɛɪ/ as *e*; /iɪ, e/ as *i*; /ɔɪ/ as *o*; /yɪ, uɪ, o, ø/ as *u*; and (ii) correspondence between gliding vowels and double vowel letters: /iw/ as *iu*; /øɥ, uɥ/ as *ui*; /ej/ as *ei*; /ɔj/ as *oi*; /ow/ as *ou*; /ɛj, aɪ/ as *ai*; /ɛw, aɪw/ as *au*. These generalizations are far from comprehensive since a monothong can also be (and is often) expressed in more than one letter, and that /ow/ is more commonly expressed in a single letter, i.e. *o*, not to mention the use of consonant letters (plus *y*) and other irregularities in vowel representation (see Cheng 2014).

The Government Romanization is not meant to be a teaching tool for Cantonese learners. Its purpose is to translate Cantonese names in the form of transliteration. Therefore, it almost never goes with the tonal indication. Unlike other ways of Romanization, tone is out of concern under most circumstances. Besides, not all place and personal names in Hong Kong are Romanized from Cantonese. Some place names in English are literally translated, e.g. *Green Island* for /ts^hɛŋ55 tsɛw55/, or they are named with no regard to how they are called in Cantonese, e.g. *Aberdeen* for /hɔŋ55 kɔŋ25 tsɛj25/. Furthermore, not all place names are transliterated from standard Hong Kong Cantonese. The most notable example is the name *Hong Kong* itself. *Hong* is based on the pronunciation of the boat people, or Tankas, during the time when the British arrived at the territory in the nineteenth century. If transliterated from the standard variety, it will most likely be rendered as *Heung* or *Heong*. Personal names are not always transliterated from Cantonese either. This is especially true for those not of Cantonese descent. Surnames vary even greater, since a native Cantonese speaker might have inherited a non-Cantonese family name. Language change is another complicating factor. A place name was Romanized and subsequent sound change made it deviant from the current pronunciation. All these mean that it is easy to draw erroneous generalizations when one attempts to learn this Romanization through a list of local place names or personal names. Though unruly, the Government Romanization is very often the only Romanization known to the general populace, not only used in the administration to transliterate proper names, but also used by the general public for various purposes. For instance, the Cantonese ‘chat alphabet’ used in informal communication such as instant messaging is loosely based on the government’s way of Romanizing Cantonese. When /jyt2 jy:23/ and /k^wɔŋ25 tɔŋ55 wɔ:25/ ‘Cantonese’ are Romanized, if needed, they are given as *Yuet Yu(e)* and *Kwong Tung Wa(h)*.⁵

3.3. Yale Romanization

The origin of Yale Romanization for Cantonese is commonly attributed to Parker Po-Fei Huang and Gerald Kok for their collaborated textbook *Speak Cantonese*. First published in

1958, it has been reprinted several times over a period of four decades. In its sequel *Book II*, Huang states that the Romanization system was developed by Kok. Based on the year of its first publication, it is reasonable to assert that the scheme came into being no later than the late fifties. Huang's later work *Cantonese Dictionary: Cantonese-English, English-Cantonese* also adopted this system. Cantonese is one of the languages Romanized under a system created at Yale University. Other Yale Romanizations involve East Asian languages such as Mandarin, Korean, and Japanese. Since its introduction, the Yale system has been extensively used in English-language textbooks, dictionaries, and glossaries made for Cantonese learners. These include some of the most popular titles currently in circulation. Comparatively speaking, the use of the Yale system is less prevalent in Chinese-language publications. While the native speakers sometimes do need to know how to pronounce a particular Chinese character, they are often referred to another homophonous character. Even when Romanization is consulted (not very often), the Yale system is not the one needed. So far, it is chiefly used to study Cantonese as a *foreign* language, rather than offer a pronunciation guide for the native speakers. Thus, Cantonese learners using this system might encounter a situation where they know this Romanization much better than the native speakers.⁶

Under the Yale system, the 19 consonants are expressed in a way that makes a distinction between aspirated and unaspirated consonants. The plosives /p, p^h, t, t^h, k, k^h, k^w, k^{wh}/ are respectively rendered as *b, p, d, t, g, k, gw, and kw*. This contrasts with the Government Romanization. For example, in the Government version, *p* represents both /p/ and /p^h/, but Yale Romanization uses two separate letters *b* and *p*. In this case, the Yale system has one-to-one mapping between letters and phonemes. To achieve this goal, the Yale system uses letters traditionally reserved for voiced consonants to represent the voiceless ones. Affricates /ts, ts^h/ are given as *j* and *ch*. Similarly this distinction is not observed in the Government Romanization, and they are both expressed as *ch* or *ts*. Nasals /m, n, ŋ/, fricatives /f, s, h/, and /j, w, l/ are straightforwardly rendered as *m, n, ng, f, s, h, y, w, l*. With /j/ Romanized as *y*, the Yale system resembles the English orthography.

Phoneme-to-grapheme consistency is also achieved in the vowel representation, although a single letter may in separate situations be used to represent distinct sounds. Letter *a*, for example, is used for both /a:/ and /æ/. Yet /a:/ is expressed in single letter *a* only when it stands alone in the rhyme. Gliding vowels with /a:/ are given in double letters *aa*, and *aa* is also used when /a/ precedes a consonant coda. Since /æ/ (always given as *a*) never occurs syllable-finally, a distinction can still be made between the two vowels. Letter *i* is used for /i:/, e/, which is common to most Romanization schemes as they are in complementary distribution. The same applies to *u*, which represents both monothongs /u:/ and /o/. It should be noted, however, that *u* is also used in *yu* as a grapheme representing /y:/. The *y* in *yu* does not represent the semi-vowel (approximant), except in the case where the initial consonant is /j/. To illustrate, /sy:/ is expressed as *syu*, whereas /jy:/ is given as *yu*, instead of the incorrect *yyu*. The 'shorthand' rule takes advantage of the fact that /ju:/ does not exist as an indigenous syllable in Cantonese, but problems may arise when loanwords are involved. /ju:/, for example, can mean 'university' in daily speech. Apparently, the Yale system does not assign a separate Romanized form for this syllable. Letter *u* is also used for other vowels when it is associated with *e*. Letter *e* as a single letter represents /ɛ:/, but when combined with *u* to form *eu*, it refers to both /œ:/ and /ø/: /lœk/ is given as *leuk* and /lœt/ is given as *leut*. Letter *o* is used for monothong /ɔ:/. The same principles are generally extended to the gliding vowels, so /iw, ej, ɔj, ɛj, əw, a:j, a:w/ are expressed as *iu, ei, oi, ai, au, aai, aau*, with *u* representing the rounded glide and *i* for the unrounded one. However, /œɥ, ɥɥ/

are given as *eui* and *ui* even though /ɥ/ is considered rounded. /ow/ is given as *ou*, although monothong /o/ is Romanized as letter *u*. Thus, /sow/ is expressed as *sou*, but /sok/ is expressed as *suk*.

The tonal representation is what gives the Yale system a distinctive hue, as it makes use of both diacritics and letter *h*. It adheres to the six-tone categorization. In its original version, a distinction was made between the high and high-falling tones of the Dark Flat tones, so that /a:55, a:53/ are expressed as \bar{a} and \grave{a} , with the vowel differently marked by macron and grave accents. The Dark Rising and Departing vowels are respectively marked by acute and no accents. Thus, /a:25, a:33/ are given as \acute{a} and *a*. With considerations of the sound change factor (discussed above), subsequent variants may use a macron accent for all Dark Flat tones, whereas a few publications nowadays still follow the original design and use a grave accent. The Light counterparts are marked with the same diacritic mark (with Light Flat tone corresponding to the high-falling Dark Flat tone), with an *h* additionally attached to the vowel. Therefore, /a:21, a:23, a:22/ are given as $\grave{a}h$, $\acute{a}h$ and *ah*. As it follows the six-tone classification, a separate representation for Entering tones is unnecessary, and the three Entering tones are treated in the same way as the non-Entering ones: /at5, at3, at2/ are rendered as $\grave{a}at$, $\acute{a}at$ and *aat*. The tone letter *h* follows the vowel but not the whole rhyme; while /səŋ21/ is expressed as *səuih*, /sən21/ is expressed as *səuhn*. Taken all together, when /jyt2 jy:23/ and /kʷŋ25 tŋ55 wa:25/ ‘Cantonese’ are Romanized in the Yale system, they become *yuht yúh* and *gwóng dūng wá*. Regarding the tone letter *h*, an additional rule was stipulated in the original system, which states that when a syllable begins with *m*, *n* or *ng*, the *h* will not be used as an indicator. This is due to the fact that these syllables are seldom expressed in the Dark tones. /ŋɔ:23/ ‘me’, for example, should be rendered as *ngó* rather than as *ngóh*. Most later applications disregard this rule, with the possible exceptions of the two syllabic nasals. The peculiarity as a result of using tone letter *h* causes some later versions to discard it altogether,⁷ and replace the diacritics with tone number: /55/ (or /53/, /5/), /25/, /33/ (or /3/), /21/, /23/, and /22/ (or /2/) are numbered from 1 to 6. In this ‘simplified’ version, *yuht yúh* and *gwóng dūng wá* are given as *yut6 yu5* and *gwong2 dung1 wa2*.

3.4. LSHK Romanization

The LSHK Romanization Scheme, also known as *Jyutping*, was proposed by the Linguistic Society of Hong Kong (LSHK) in 1993. The LSHK is a non-governmental organization comprising linguistic academics and professionals based in Hong Kong. This Romanization scheme was collectively designed by scholars specializing in various subfields of language studies. A conference had been held earlier in 1992 to gather thoughts on the future scheme before the finalized version was released at the Fourth International Conference on the Yue Dialects in December 1993. The scheme aims to be multipurpose, systematic, inclusive, and user-friendly and takes care of both the earlier traditions and word-processing technology of the time (Cheung 1994). Since its introduction, the LSHK Romanization has gained a strong foothold in the academia. Up to this day, a majority of English and Chinese-language academic papers published in Hong Kong and also in the West adopt this system for Romanization, apart from the direct use of phonetic symbols. The Government of Hong Kong does not have a fixed position on which scheme should be the default option. Having said that, the post-handover Education Bureau, in collaboration with the tertiary scholars, has published materials that use this system as the pronunciation guide. They abandoned the earlier scheme (‘ILE Romanization’; to be discussed) used for another guidebook published by the Education Department in the early nineties. As the scheme was invented by tertiary

scholars, it is taught in local university courses that touch upon Cantonese pronunciation. Besides, its promotion has occasionally been extended to primary and secondary schools. But still most Hong Kong people, including university students, have never been required to learn Cantonese sound notation at school. Also due to its fairly recent date of emergence, the LSHK Romanization is still on its way to reaching a wider spectrum of people outside the academic community.

The representation of consonants has given this system some noticeably distinctive features. Same as the Yale Romanization, plosives /p, p^h, t, t^h, k, k^h, k^w, k^{wh}/, nasals /m, n, ŋ/, fricatives /f, s, h/, and approximants /w, l/ are conveniently given as *b, p, d, t, g, k, gw, kw, m, n, ng, f, s, h, w, l*. What makes the LSHK system spectacular is how it renders affricates /ts, ts^h/, which are given as *z* and *c*. It conforms to Mandarin *pinyin*, but deviates from the Government traditions in which single graphemes *z* and *c* are never used to represent consonants. In the Government Romanization, *c* is only used in *ch* for /ts, ts^h/, and *z* is never used for consonants and only used in few syllables like *tsz* /ts^hi:/ or *sze* /si:/. Another deviation from the commonplace Romanization is on how it expresses /j/. Letter *j* is used for /j/ in the LSHK system, but in Hong Kong's common practice, it is instead used for /ts/, as is often found in the name of local companies. Note that these peculiarities were made 'on purpose', so that it takes care of both the *pinyin* (for *z* and *c*) and IPA traditions (for *j*), and in so doing 'enhances user-friendliness' (Cheung 1994).

Regarding the Romanization of vowels, the LSHK system is arguably the most precise one in representing the distinctive sounds. One-to-one mapping from phoneme to grapheme is achieved in most vowels. And each rhyme corresponds to exactly one written unit. Monothongs /i:, e, y:, u:, o, e:, ɔ:, œ:, ø, ɐ, a:/ are respectively rendered as *i, i, yu, u, u, e, o, oe, eo, a*, and *aa*. /i:, e/ and /u:, o/ are expressed in one letter only (*i* and *u*) given that these sounds are in complementary distribution. Except them, each monothong matches one grapheme. To achieve phonemic distinction, the LSHK scheme resorts to the use of four digraphs *yu, oe, eo, and aa* to represent monothongs, in which *yu* is somewhat conspicuous as letter *y* is more often understood in English as a consonant letter. Unlike the Yale system, however, no 'shorthand' rule is added (*yyu* becomes *yu* for /jy:/); /jy:/ is directly expressed as *jyu*. Digraphs are specially designed so that only the basic Roman alphabet (26 letters) is used. This is of particular significance during the time when keying in a diacritic was a hassle. As for /ɐ, a:/, the distinction made by using *aa* and *a* is an inheritance of the earlier traditions. Another important feature is that it distinguishes /œ:, ø/. They have similar sound quality and have traditionally been expressed in only one grapheme (in the Yale system, for example, they are Romanized as *eu*). In so doing, it manages to render some 'non-indigenous' rhymes and syllables arising as a result of the recent sound change. In standard, conservative accent, /œ:, ø/ are complementary, but recent merger of codas /t, k/ and /n, ŋ/ into /t/ and /n/ has created new rhymes such as /œn, œt/. Thus, in this new variety, /œ:, ø/ are contrastive and can be differentiated with minimal pairs. Assigning two graphemes *oe* and *eo* separately for /œ:/ and /ø/ allows the system to become 'inclusive' by capturing this recent contrastive distinction. The same conversion rules are extended to gliding vowels: /iw, øɥ, uɥ, ɔj, ɛj, ɐw, a:j, a:w/ are given as *iu, eoi, ui, oi, ai, au, aai, aau*, with letter *u* representing /w/ and *i* representing /j/ or /ɥ/. Since the latter two glides are not contrastive, they are expressed in one single letter – a common practice in Cantonese Romanization. /ej, ow/ are treated as if the vowel were less high: monothongs /e, o/ are given as *i* and *u*, but as an element in the gliding vowels, they are rendered as *e* and *o*, and /ej, ow/ are given as *ei* and *ou*. The paradigm of the LSHK system allows the inclusion of new rhymes or syllables traditionally non-existent but recently created due to foreign-language influences. These include letter

pronunciation such as /ɛw55/ ‘L’ (*eu1*), /ɛm55/ ‘M’ (*em1*) and /ɛn55/ ‘N’ (*en1*), as well as transliterated loanwords, such as /ɛp5/ ‘app’ (*ep1*). Many other systems, on the other hand, do not specify how, or if at all, these syllables should be Romanized. For example, the use of *eu* for /œ:, ø/ in the Yale system implies that /ɛw/ does not exist; otherwise, the two vowels will share the same grapheme *eu*, which is implausible. Under the LSHK scheme, all these rhymes are regarded as existent so that it achieves ‘phonological inclusiveness’ (Cheung 1994). The scheme is systematic and highly inclusive, but it also risks producing some Romanized forms unintelligible to people who are only familiar with the Government Romanization. Examples of this include *jeon* (/jøn/) and *zyu* (/tʃy:/). They are commonly Romanized as *yun* and *chu* (*tsu*), quite dissimilar from what they are given in the LSHK format.

The LSHK system adheres to the six-tone categorization, meaning that there is no separate notation required for the Entering tones. Tone number is used to mark the tones: number 1 for /55/ (or /5/), 2 for /25/, 3 for /33/ (or /3/), 4 for /21/, 5 for /23/, and 6 for /22/ (or /2/). The LSHK working group decided that tone number 1 does not necessarily refer to /55/ (Cheung 1994). If tone /53/ exists in anyone’s speech, it is still marked by the same number. The use of numbering is desirable because numerals are in the basic ASCII character set. Since in computing Chinese characters are coded with at least one ‘non-basic’ ASCII character, like those with diacritic marks, inputting diacritics for tones might mistakenly produce Chinese characters in the earlier versions of Chinese word-processing software. The ordering also reflects a more common practice at the time. An alternative ordering puts the Flat-Rising-Departing-Entering sequence in the first order and the Dark–Light distinction in the second. To write the tone number, some earlier schemes stipulate that it be written in superscript style. This is not required in the LSHK scheme. Taken all together, when /jyt2 jy:23/ and /kʷɔŋ25 toŋ55 wa:25/ ‘Cantonese’ are Romanized in the LSHK system, they become *jyut6 jyu5* and *gwong2 dung1 waa2*. Often, tonal indication is omitted in translated terms, such as the term *Jyutping* itself.⁸ If tone number is omitted, *jyut6 jyu5* and *gwong2 dung1 waa2* will be expressed as *Jyutjyu* and *Gwongdungwaa*. It is assumed that the same capitalization rules of the target language apply to those terms transliterated in LSHK Romanization.

3.5. Rao’s Romanization

It refers to the Romanization scheme used in the dictionaries of which Rao Bingcai is one of the editors. The official name of the scheme is called the Scheme of the Cantonese Phonetic Alphabet. The system attributed to Rao is a modified version of the said scheme, which was officially announced in 1960 by the executive branch for education of Guangdong province. The original scheme was part of a larger project encompassing four different Romanization schemes for Guangzhou dialect (Cantonese), Kejia dialect (Hakka), Hainan dialect, and Chaozhou dialect (Swatow). In theory, this is a government standard for Cantonese Romanization. However, the official scheme has yet to gain widespread adoption. It should be best understood as a government invention without much standardization effort to follow it up. Rao’s modified version, on the other hand, has become predominant in mainland China, with its use extended from his own publications to a variety of Cantonese textbooks and dictionaries written by others. It has largely replaced the original scheme, which has since then been only occasionally referred to in the footnotes of the pronunciation guide. The changes made in the modified scheme include the replacement of some vowel graphemes with another one and also the addition of two syllabic nasals.⁹ They will be highlighted below.

Under this scheme, a phoneme may be represented by more than one grapheme, depending on the phonetic and orthographic conditions. When used syllable-initially as an onset, plosives /p, p^h, t, t^h, k, k^h/ are given as *b, p, d, t, g, k*. This agrees with their representation in most other systems. However, when /p, t, k/ are used syllable-finally as a coda, expressed as unreleased [p[̚], t[̚], k[̚]], they are given as *b, d, g*. This differs from all major forms of Romanization, including Government, Yale, and LSHK, in which they are rendered as *p, t, k*. Thus, /p^hit/ is expressed as *pid* in this scheme, but is expressed as *pit* in others. As for the other two plosives /k^w, k^{wh}/, they are given as *gu* and *ku*. Letter *u*, instead of *w* as in some other Romanizations, is used to represent the labialization of velar stops. The original scheme clarifies this by stating that *gu* and *ku* are syllables when they stand alone, i.e. in the case of /ku/, k^hu:/, whereas they are initials when they are not used independently in the phonetic notation; as in the case of /k^wa:/, k^{wh}a:/, they are given as *gua* and *kua*. Nasals /m, n, ŋ/, fricatives /f, h/, and approximants /w, l/ are as usual Romanized as *m, n, ng, f, h, w, l*, and /j/ is given as *y*. The most complex way of conversion occurs to /ts, ts^h, s/. They are variably expressed as *z, c, s* or *j, q, x* depending on the following vowel letter: as *j, q, x* when they precede vowel letters *i* or *ü* (shorthand as *u*; to be discussed), and as *z, c, s* elsewhere. To illustrate, /si:/ is given as *xi*, while /sa:/ is expressed as *sa*. Some previous literature suggests that it indicates a distinction between /ts, ts^h, s/ (as *z, c, s*) and /tʃ, tʃ^h, ʃ/ (as *j, q, x*). This is more apparent than real. First, whether the two sets of affricates are in complementary distribution is debatable. Even if they are so, the post-alveolar /tʃ, tʃ^h, ʃ/ usually precede round vowels, but *j, q, x* apply to both rounded (/y:/) and unrounded (/i:, e/) monothongs, so do *z, c, s*. Besides, monothong /e/ and gliding vowel /ej/ are differently expressed as *i* and *éi*, and different sets of letters are used for them. More importantly, the scheme explicates that there is no distinction between these two sets of consonants letters. This is totally out of orthographic concerns and is irrelevant to any ‘natural classes’ that show distinctive phonetic features. The orthographic considerations are in turn taken to resemble the Mandarin *pinyin* system. Such an arrangement is more pedagogically relevant than phonetically necessary. As far as the consonants are concerned, Rao’s modification is equivalent to the original scheme, except it additionally specifies that the two nasals can also be syllabic.

In Rao’s Romanization, monothongs /i:, e, y:, u:, o, ɛ:, ɔ:, œ:, ø, v, a:/ are respectively rendered as *i, i, ü, u, u, é, o, ê, ê, e, a*. Complementary /i:, e/ and /u:, o/ are expressed in one letter, i.e. *i* and *u*. /œ:, ø/ are both expressed as *ê*, a plausible arrangement for the standard accent (see the description above). Note that diacritics, instead of digraphs, are employed to render these sounds. Therefore, a monothong is always expressed by a single letter. Besides, the use of *ü* for /y:/ allows it to conform to the *pinyin* system. Meanwhile, the *pinyin* orthographic rules are transferred to the Cantonese system: *ü* is written as *yu* for *yu /jy:/, yun /jyn/, and yud /jyt/*; the two dots are also removed after *j, q, x*, e.g. *ju /tʃy:/, qu /tʃ^hy:/, xu /sy:/*. These are essentially the *pinyin* rules despite that in Mandarin /j/ is deemed non-existent, and letter *y* is added but represents no consonant. The ‘shorthand’ rule is possible only because /u:/ does not follow /ts, ts^h, s, j/ in the indigenous Cantonese phonological system. Borrowed syllable /sut/ (for ‘suit’) might be expressed as *sud*, making use of two sets of homophonous letters, *z, c, s* vs. *j, q, x*, but /ju:/ (for letter ‘U’ or ‘university’) cannot be represented without conflicting with the indigenous syllable /jy:/. Conformity with the *pinyin* system is also exhibited in the case of /v/. In other systems, the /v, a:/ distinction is often made by using single letter *a* and double letters *aa*. Here, *e* is used for /v/ and *a* is reserved for /a:/. The more basic *e* is reserved for /v/, whereas *é* and *ê* are for /ɛ:/ and /œ:, ø/. This is possibly because rhymes like *en /vɛn/* and *eng /vɛŋ/* resemble Mandarin *en /ən/* and *eng /əŋ/*, and syllables such as *zé, cé, sé /tsɛ:, ts^hɛ:, sɛ:/* do sound quite differently

from their near-homographs in Mandarin, i.e. *ze*, *ce*, *se* /tsɿ, tsʰɿ, sɿ/. Gliding vowels basically follow the same conversion pattern: the vowel element in /iw, øʏ, ʊʏ, əj, ɛj, ɐw, aɨ, a:w/ follows the same rules and so they are converted into *iu*, *êu*, *ui*, *oi*, *ei*, *eo*, *ai*, *ao* respectively. Apart from this, /ej, ow/ are given as *éi* and *ou*, treated as if they were the gliding variants of /ɛɿ, ɔ:/. In this regard, the Yale and LSHK systems also share the same treatment. Note that the glides in these vowels are variably represented. /w/ is given as *u* for *iu* /iw/ and *ou* /ow/, but as *o* for *eo* /ɐw/ and *ao* /a:w/; /ɥ/ in *êu* /øʏ/ is expressed as *u*, but is expressed as *i* in *ui* /ʊʏ/. The description above is for Rao's modified scheme. In the original scheme, the four rhymes /ɐw, ow, əj, ɔj/ are given as *ou*, *ô*, *ông*, *ong*, rather than as *eo*, *ou*, *ong*, *ung*. Phonetically speaking, these changes improve phoneme-to-grapheme consistency. Besides the obvious problem of using *e* for /ɐ/ but *ou* for /ɐw/, the use of *ô* also undermines the coherence of the scheme. Letter *ô* is the only non-digraph for gliding vowels. Even if the sound it represents is said to be monothongized as /o/, *ông* should better represent /ɔj/ than /əj/, not the other way around as it does in the original scheme. It appears that the original assignment of *ong* for /ɔj/ is deliberate because this rhyme is closer to Mandarin *ong* /ɔŋ/ than /əj/ is.

Both the original and Rao's modified schemes adhere to the six-tone categorization. As such, Dark Flat (Upper Dark Entering), Dark Rising, Dark Departing (Lower Dark Entering), Light Flat, Light Rising, and Light Departing (Light Entering) tones are numbered from 1 to 6 respectively. Cheung (2003) notes that the original 1960 scheme has established the now-common tonal notation traditions, in which Entering tones are no longer independent and the tonal order is first determined by the Dark–Light distinction rather than by the traditional tonal categories. Later systems like the LSHK scheme follow this ordering. Again, Tone 1 refers to either /55/ or /53/, a distinction prevalent in the older-day Guangzhou Cantonese for which the scheme was invented. Taken all together, /jyt2 jy:23/ and /kʷɔŋ25 tɔŋ55 wa:25/ 'Cantonese' are Romanized in Rao's system as *yud⁶ yu¹* and *guong² dung¹ wa²*. Unlike the LSHK scheme, a superscript style for tone number is required.

3.6. Other systems

Other influential systems include those proposed by S. L. Wong, Sidney Lau, and the Institute of Language in Education. In some areas, S. L. Wong's scheme is perhaps more influential than the Yale and LSHK ones. It originated from his book *A Chinese Syllabary Pronounced according to the Dialect of Canton* (1941) and forms the basis of the phonetic symbols used in a majority of the Standard Chinese dictionaries published in Hong Kong. Yet Wong's system and its variants may not fit the strict definition of 'Romanization' because phonetic symbols, rather than ordinary Roman letters, are used for transcription. Most dictionaries simply term it 'International Phonetic Alphabet'. It is often seen as a 'broad' phonetic transcription for Cantonese, and as a matter of fact, some conventionalized symbols should be understood in order to interpret them correctly. Most notably, it uses voiced consonant symbols for unaspirated plosives and affricates. Thus, unaspirated /p, t, k, kʷ, ts/ are rendered as /b, d, g, gw, dz/. Beginning learners with prior knowledge of IPA might risk viewing them as voiced consonants.¹⁰ Other features that make it a 'broad' transcription include the use of /œ/ for both /œ:/ and /œ/, and /i, u, y/ in gliding vowels, so that /iw, øʏ, ʊʏ, ej, əj, ow, ɛj, ɐw, aɨ, a:w/ are expressed as /iu, œy, ui, ei, ɔi, ou, ɛi, ɐu, ai, au/. Note that the vowel length mark is not used and /ɥ/ is variably expressed as /y/ or /i/, as in /œy, ui/. This system is sufficient for native speakers to look up the pronunciation of a Chinese character, but non-native learners using it might enunciate wrongly if not knowing the associated phonetic details.

The Institute of Language in Education (ILE) Scheme, on the other hand, is an exemplar of Cantonese Romanization. It was proposed by the said institute, which had then been a subsidiary of the colonial Education Department and was later incorporated into the Hong Kong Institute of Education (HKIEd). This scheme is a modification of the one proposed by P. C. Yu in his book *Rhyming Glossary* (1971). Same as the LSHK scheme, only 26 letters and Arabic numerals are used for Romanization. This system once acquired some degree of authority through the publication of *List of Cantonese Pronunciation of Commonly-used Chinese Characters*, printed by the Government Printer in 1990. Later publications by the administration, however, have changed to the LSHK system. A slightly modified version of the ILE scheme is used in *The Dictionary of Standard Cantonese Pronunciation* edited by Zhan Bohui (2002). In vowel representation, the ILE scheme renders /a:/ as *a* in the open syllables and as *aa* in the closed ones (same as the Yale Romanization). Yu's original scheme has the same treatment, except that the letter *a* for /ə/ is always italicized, whereas the one for /a:/, as well as all other letters, is not. Zhan's modification uses *aa* for /a:/ in all phonetic contexts.

Finally, Sidney Lau's Romanization is associated with a series of Cantonese textbooks written by him and published by the Government Printer starting from 1965. Lau has spent years in the Hong Kong Government teaching Cantonese and was once the Principal of the Government Language School. One of the main characteristics of this system is that it bears some resemblance to the English orthography. In addition to the use of *oo* for /u:/, it also resorts to the use of final *h* to close the syllable with /ɔ:/, ə:/ when they are not followed by a final consonant or glide, i.e. *goh* for /kɔ:/ but *gong* for /kɔŋ/. Lau stated that his scheme was an adaptation of the Meyer-Wempe system, which already included these features. The former two systems, Wong and ILE, adhere to the nine-tone categorization. The Upper Dark, Lower Dark, and Light Entering tones are numbered from 7 to 9, rather than being merged into the non-Entering ones for tonal notation. Sidney Lau's Romanization, on the other hand, adheres to the six-tone categorization.

4. Conclusion

In the practice of Cantonese Romanization, it is not uncommon for writers and academics to develop a new scheme or improve on existing ones – a reason for a substantial and growing number of systems. The Romanization forms discussed above are those that have historical significance and/or stimulated adoption by others who are *not* the inventors or definite proponents, or their associates. Tables 3.1–3.4 summarize the most popular schemes over the past half century.¹¹

Together with the Government Romanization, these Romanization forms are employed in various occasions and reach different groups of target audience. For instance, to transliterate a Cantonese term for a Hong Kong-based English bulletin, without much regard to phonemic precision, the Government Romanization allows the largest number of locals to understand what the transliterated term means. English-speaking Cantonese learners are often exposed to the Yale Romanization and might prefer a transcription in this system. Meanwhile, Mandarin-speaking learners of Cantonese who have prior knowledge of *pinyin* might find Rao's scheme most accommodating. In Hong Kong's academia, the LSHK Romanization has been becoming the *de facto* standard and is expected to be used by the faculty and students alike in academic papers. With such a diversity of uses and users of many backgrounds, the apparent chaos of Cantonese Romanization may well be thought of as a 'blessing in disguise'.

Table 3.1 Initial consonants

<i>IPA</i>	<i>Wong</i>	<i>Yale</i>	<i>Rao</i>	<i>ILE</i>	<i>LSHK</i>	<i>Lau</i>
p	b	b	b	b	b	b
p ^h	p	p	p	p	p	p
m	m	m	m	m	m	m
f	f	f	f	f	f	f
t	d	d	d	d	d	d
t ^h	t	t	t	t	t	t
n	n	n	n	n	n	n
l	l	l	l	l	l	l
ts	dz	j	z, j	dz	z	j
ts ^h	ts	ch	c, q	ts	c	ch
s	s	s	s, x	s	s	s
j	j	y	y	j	j	y
k	g	g	g	g	g	g
k ^h	k	k	k	k	k	k
ŋ	ŋ	ng	ng	ng	ng	ng
h	h	h	h	h	h	h
k ^w	gw	gw	gu	gw	gw	gw
k ^{wh}	kw	kw	ku	kw	kw	kw
w	w	w	w	w	w	w

Note

Wong: Variant of S.L. Wong's Scheme Used in Hong Kong's Dictionaries.

Yale: The General Scheme of Yale Romanization.

Rao: Rao Bingcai's Modification of the Scheme of the Cantonese Phonetic Alphabet.

ILE: Institute of Language in Education Romanization.

LSHK: Linguistic Society of Hong Kong Romanization.

Lau: The General Scheme of Sidney Lau's Romanization.

Table 3.2 Final consonants

<i>IPA</i>	<i>Wong</i>	<i>Yale</i>	<i>Rao</i>	<i>ILE</i>	<i>LSHK</i>	<i>Lau</i>
m	m	m	m	m	m	m
n	n	n	n	n	n	n
ŋ	ŋ	ng	ng	ng	ng	ng
p	p	p	b	p	p	p
t	t	t	d	t	t	t
k	k	k	g	k	k	k

Table 3.3 Vowels

IPA	Wong	Yale	Rao	ILE	LSHK	Lau
i:	i	i	i	i	i	i
e	i	i	i	i	i	i
y:	y	yu	ü	y	yu	ue
u:	u	u	u	u	u	oo
o	u	u	u	u	u	u
ɛ:	ɛ	e	é	e	e	e
ɔ:	ɔ	o	o	o	o	o(h)
œ:	œ	eu	ê	oe	oe	eu(h)
ə	œ	eu	ê	oe	eo	u
ɐ	ɐ	a	e	a	a	a
a:	a	a(a)	a	a(a)	aa	a(a)
iw	iu	iu	iu	iu	iu	iu
əɥ	œy	eui	êu	oey	eoï	ui
uɥ	ui	ui	ui	ui	ui	ooï
ej	ei	ei	éi	ei	ei	ei
ɔj	ɔi	oi	oi	oi	oi	oi
ow	ou	ou	ou	ou	ou	o
ɐj	ɐi	ai	ei	ai	ai	ai
ɐw	ɐu	au	eo	au	au	au
a:j	ai	aaï	ai	aaï	aaï	aaï
a:w	au	aau	ao	aau	aau	aau

Table 3.4 Tones

IPA	Wong	Yale	Rao	ILE	LSHK	Lau
55 / 53 (5)	a ¹ (a ⁷)	ā / à	a ¹	a ¹ (a ⁷)	a1	a ¹⁰ / a ¹ (a ¹⁰)
25	a ²	á	a ²	a ²	a2	a ²
33 (3)	a ³ (a ⁸)	a	a ³	a ³ (a ⁸)	a3	a ³
21	a ⁴	àh	a ⁴	a ⁴	a4	a ⁴
23	a ⁵	áh	a ⁵	a ⁵	a5	a ⁵
22 (2)	a ⁶ (a ⁹)	ah	a ⁶	a ⁶ (a ⁹)	a6	a ⁶

Notes

- 1 Since consonants are not obligatory in the syllable-initial position, a *zero* consonant is often suggested as a phonological unit (as onset).
- 2 Zee (1999) indicates that /i:, y:, u:, ɛ:, ɔ:, œ:, a:/ are long in open syllables and a third shorter in syllables closed by a plosive or nasal. The vowel length mark is not used here for the latter syllables.
- 3 Tone numerals are used. Dark Rising, Light Flat and Light Rising tones might also be given as /35/, /11/ and /13/. For simplicity, the description above follows Bauer and Benedict (1997).
- 4 According to Kataoka and Lee (2008), the use of *ch* and *ts* makes historical sense since Cantonese used to show distinction between /ts^(h)/ and /tʃ^(h)/ . The merger into /ts^(h)/ has now become the norm.
- 5 Neighbouring Macau, a former Portuguese colony, also has its own way to Romanize Cantonese. Precisely, this is not Romanization to English, but to Portuguese, although its Romanization is transferred directly into English contexts. Certain features differentiate it from the Hong Kong

- version, possibly designed to adapt for the Portuguese orthography. Major differences include its using *v* for /w/ (though *w* is also used) and *i* (or no letter at all) for /j/. /jyt2 jy:23/ and /k^wɨŋ25 tɔŋ55 wa:25/ may possibly be expressed as *Ut U* (or *Iut Iu*) and *Kuong Tong Va* in the Macau variety (see Cheng 2014 for a comparative study of Government Romanizations in Hong Kong and Macau).
- 6 For some Cantonese speakers, the Yale Romanization might be noticed through its association with the Google Cantonese Input Method; although it accepts a wide range of Romanization variants, the ‘suggestion’ given in the input box is based on the Yale system (in a ‘simplified’ version; to be discussed below).
 - 7 The use of letter *h* to mark the Light tones also creates a peculiar sorting order – in alphabetizing the dictionaries entries, for example, characters with the same phonetic segment but different tones are not put in consecutive order. /sɔn25/ shares the same rhyme with /sɔn23/ but not with /sɔŋ23/. But when the order is determined by the alphabetic sequence of their Yale Romanization, the three syllables will be in the sequence of ‘/sɔn23/-/sɔŋ23/-/sɔn25/’ as they are Romanized as *séuhn*, *séuih*, and *séun*. This differs from the normal practice in which characters having the same segment but different tones are sequenced consecutively.
 - 8 This is reminiscent of how Mandarin *pinyin* is employed in foreign-language publications and scholarly works. In the *pinyin* system, an apostrophe is used to reduce ambiguity by marking the end of one syllable and the beginning of another, when the second one starts with a vowel letter. Scholars using the LSHK scheme often follow the same practice and put an apostrophe before syllables that begin with a vowel as well as syllabic nasals, i.e., *co'ng* for *co3 ng6* ‘error’ and *cong* for *cong1* ‘tumor’.
 - 9 The changes are said to have been made ‘in response to public opinion’. It is not sure whether this may violate the specification of the original scheme which states that ‘in principle no more new letters will be invented and the letter pronunciation will not be casually changed’ (Guangdongsheng sizhong fangyan pinyin fang’an 1960).
 - 10 Wong did mention that /b, d, g, dz/ are in fact devoiced [b̥, d̥, g̥, dz̥] (/gw/ is treated as /g/ plus /w/). These phonetic details are normally omitted in dictionaries adopting Wong’s system.
 - 11 For easy reference, the details regarding the orthographic conventions of the original Yale scheme, the ‘simplified’ Yale version, the predecessor/successor of the ILE scheme, and the original schemes of Rao’s and Wong’s systems are not shown in the tables. They are addressed in the preceding paragraphs.

References

- Ball, James Dyer (1883) *Cantonese Made Easy*, Hong Kong: China Mail Office.
- Bauer, Robert S. and Paul K. Benedict (1997) *Modern Cantonese Phonology*, Berlin: Mouton de Gruyter.
- Bauer, Robert S., Cheung Kwan-Hin, and Cheung Pak-Man (2003) ‘Variation and Merger of the Rising Tones in Hong Kong Cantonese’, *Language Variation and Change* 15(2): 211–25.
- Bridgman, Elijah Coleman (1841) *Chinese Chrestomathy in the Canton Dialect*, Macau: S. W. Williams.
- Chalmers, John (1859) *An English and Cantonese Pocket-dictionary, for the Use of Those Who Wish to Learn the Spoken Language of Canton Province*, Hong Kong: The London Missionary Society’s Press.
- Chao, Yuen-Ren (1947) *Cantonese Primer*, Cambridge, Mass.: Harvard University Press.
- Cheng, Siu-Pong (2014) ‘Gang-Ao liang di de zhengfu Yueyu pinyin’ (‘The Government Romanization of Cantonese in Hong Kong and Macao’), *Zhongguo yuwen tongxun (Newsletter of Chinese Language)* 93(1): 27–38.
- Cheung, Kwan-Hin (1986) ‘The Phonology of Present-day Cantonese’, Doctoral Dissertation, University College London.
- Cheung, Kwan-Hin (1994) ‘Xianggang Yuyanxue Xuehui Yueyu Pinyin Fang’an de yuanqi, sheji yuanze he tedian’ (‘The Genesis, Design Principles, and Characteristics of the LSHK Cantonese Romanization Scheme’), in *Yueyu Pinyin Zibiao (Guide to LSHK Cantonese Romanization of Chinese Characters)*, Hong Kong: Linguistic Society of Hong Kong, ix–xix.
- Cheung, Kwan-Hin (2003) ‘Guangzhouhua Pinyin Fang’an xilie de wanshanhua guocheng’ (‘A Series of Refinements to the Cantonese Romanization Scheme 1960’), in Tang Keng-Pan (ed) *Proceeding of 6th International Conference on Yue Dialects*, Macau: Associação de Literatura Chinesa de Macau, 13–19.

- Eitel, Ernst John (1877) *A Chinese Dictionary in the Cantonese Dialect*, London: Trübner and Company.
- Guangdongsheng sizhong fangyan pinyin fang'an (Four schemes of dialectal phonetic symbols of the Guangdong Province) (1960) *Yuwen jianshe (Language Planning)* 15: 21–5.
- Huang, Parker Po-Fei (1962) *Speak Cantonese: Book II*, New Haven: Institute of Far Eastern Languages, Yale University.
- Huang, Parker Po-Fei (1970) *Cantonese Dictionary: Cantonese–English, English–Cantonese*, New Haven: Yale University Press.
- Huang, Parker Po-Fei and Gerald P. Kok (1958) *Speak Cantonese*, New Haven: Institute of Far Eastern Languages, Yale University.
- Institute of Language in Education (1990) *Changyongzi Guangzhouhua duyinbiao (List of Cantonese Pronunciation of Commonly-used Chinese Characters)*, Hong Kong: Government Printer.
- Kataoka, Shin (2014) 'Xianggang zhengfu pinyin: yi ge luanzhongyouxu de xitong' ('Finding Order in Disorder: An Investigation into the Hong Kong Government Cantonese Romanization'), *Zhongguo yuwen tongxun (Newsletter of Chinese Language)* 93(1): 9–25.
- Kataoka, Shin and Cream Lee (2008) 'A System without a System: Cantonese Romanization Used in Hong Kong Place and Personal Names', *Hong Kong Journal of Applied Linguistics* 11(1): 79–98.
- Lam, Yan-Yan (2009) 'Xianggang Yueyu biao'in de xianzhuang' ('The Current Situation of Hong Kong Cantonese Sound Notation'), *Zhongguo yuwen yanjiu (Studies in Chinese Linguistics)* 27(1): 59–66.
- Lau, Sidney (1965) *'Kwang Tung Wah': Cantonese by Radio, Intermediate. Listeners' Guides to Accompany Lessons 1 to 117*, Hong Kong: Government Printer.
- Matthews, Stephen and Virginia Yip (1994) *Cantonese: A Comprehensive Grammar*, London and New York: Routledge.
- Meyer, Bernard F. and Theodore F. Wempe (1934) *The Student's Cantonese–English Dictionary*, Hong Kong: St. Louis Industrial School Printing Press.
- Morrison, Robert (1828) *A Vocabulary of the Canton Dialect*, Macau: The Honorable East India Company's Press.
- Williams, S. Wells (1856) *A Tonic Dictionary of the Chinese Language in the Canton Dialect*, Guangzhou: The Office of the Chinese Repository.
- Wong, S. L. (1941) *Yueyin yunhui (A Chinese Syllabary Pronounced according to the Dialect of Canton)*, Shanghai: Zhonghua.
- Wu, Man-Ying (1997) 'Yueyu Luoma pinyin fang'an de bijiao' ('Comparison of Cantonese Romanization Schemes'), Master Thesis, The Hong Kong Polytechnic University.
- Yu, Ping-Chiu (1971) *Tongyin zihui (Rhyming Glossary)*, Hong Kong: Kwong Wah.
- Zee, Yun-Yang Eric (1999) 'Chinese (Hong Kong Cantonese)', in International Phonetic Association (ed) *Handbook of the International Phonetic Association*, Cambridge: Cambridge University Press, 58–60.
- Zhan, Bohui (ed) (2002) *Guangzhouhua zhengyin zidian (The Dictionary of Standard Cantonese Pronunciation)*, Guangzhou: Guangdong Renmin Chubanshe.