

Yuè & the Sinitic Languages of the Pearl River Region — Unit 3.2

13th EACL Summer School, Universität Stuttgart

Hilário de Sousa 蘇沙

EHESS–CRLAO

membre associé

Wednesday 29 July 2026

丙午六月十六

hilario@bambooradical.com

Yue-Pinghua

Various ways of classifying dialects;
different classifications are good for different purposes.

Some questions:

1. Definition of Yue / Pinghua
2. Internal classification of Yue / Pinghua
 - 3a. Relationship between Yue & Pinghua
 - 3b. Boundary between Yue & Pinghua (if they are considered separate)

Yue–Pinghua LAC 2012



Figure 1: Some Southern Sinitic dialect groups per *Language Atlas of China*¹⁹⁸⁷

What is Yue/Pinghua? Not explicitly discussed in the LAC;
Preconceived categories of “Yue” & “Pinghua”. Sinitic dialects in and around
Guangdong/Guangxi which are not Hakka/Min/Mandarin/Xiang/“土話 Tuhua”.

Yue-Pinghua LAC 2012

Language Atlas of China (2nd ed.) B1-18 粤语 (by 伍巍)

粤语内部有几个相对统一的特点：

(1) 古全浊声母在多数方言中清化。清化后，有的读送气音，有的读不送气音，也有的平声、上声送气，去声、入声不送气。

[MC voiced obstruent initials have become voiceless in most dialects. After becoming voiceless, in some dialects they are aspirated, in some dialects they are unaspirated, in some dialects they are aspirated in tones AB and unaspirated in tones CD.]

(2) 古微母字今读双唇音，如“舞”、“文”、“万”、“网”等字今声母一律读 [m-]，舞 = 母、尾 = 美、文 = 门、万 = 慢。

[MC 微 [mw-/ɱ-] initial characters are pronounced bilabial in the modern language, e.g. 舞文万 all pronounced [m-] nowadays, e.g. [Cantonese] 舞 mou5 = 母 mou5, 尾 mei5 = 美 mei5, 文 man4 = 門 mun4, 萬 maan6 = 慢 maan6]

Yue-Pinghua LAC 2012

(3) 齿音声母一般只有一套，多数方言读舌叶音 [tʃ- tʃʰ- ʃ-]。

[There is usually only one set of sibilants, in most dialects they are laminal [tʃ- tʃʰ- ʃ-].]

[Actually not uncommon to have 2 series, especially e.g. /ʃ/ vs. /ʈ-/]

(4) 见组及晓母字不论洪细音，声母多不腭化，今读舌根音 [k- kʰ- ŋ- h-]，娇 ≠ 焦、溪 ≠ 妻、件 ≠ 践、惊 ≠ 精、脚 ≠ 雀、晓 ≠ 小。

[MC 見 [k-] series and 曉 [h-] initial characters, no matter whether it is followed by 洪 [open vowels / u] or 細 [i / y], the initial is usually not 腭化 palatalised, they are nowadays pronounced velar [k- kʰ- ŋ- h-], [Cantonese] 嬌 giu1 ≠ 焦 ziu1, 溪 kai1 ≠ 妻 cai1, 件 gin6 ≠ 踐 cin5, 驚 ging1 ≠ 精 zing1, 腳 goek3 ≠ 雀 zoek3, 曉 hiu2 ≠ 小 siu2]

部分方言流、深、臻三摄牙音细音字声母有齿化现象。[In some dialects, there is dentalisation with MC 流 [-iw], 深 [-im], 臻 [-in] groups of finals in front of 細音 [i / y]

e.g. 玉林九 tʃau˧ 球 tʃau˨ 金 tʃam˧ [Canto. 九 kɐu˧ 球 kʰɐu˧ 金 kɐm˧]

Yue-Pinghua LAC 2012

(5) 古溪母开口常用字多有读清喉擦音 [h-] 的现象；古溪母合口字（多与古晓母合口字合流）今多读 [f-]，如广州话：苦 = 虎 = 府 [fu˧]，宽 = 欢 [fun˧]。

[With 常用字 common characters, MC 開口 non-rounded 溪 [kʰ-] initial often has the phenomenon of being pronounced as 清喉擦 voiceless glottal fricative [h-]; MC 合口 rounded 溪 [kʰ-] initial is (mostly merged with MC 合口 rounded 曉 [h-] initial) now often pronounced [f-], e.g. Canton dialect 苦 = 虎 = 府 [fu˧], 寬 = 歡 [fun˧].]

[Canto. 開 hoi1 口 hau2 坑 haang1]

Yue-Pinghua LAC 2012

(6) 除邕潯粤语外，大多数方言无舌尖元音韵母。绝大多数粤语止撮开口齿音字韵母为舌面元音 [-i]，部分粤语止撮开口精、庄组字韵母为 [-y] 或 [-u]。

[Except 邕潯 Yōngxún Yue / Jung1 Cam4 Jyut6, most dialects do not have 韻母 finals with 舌尖元音 apical vowels. In the vast majority of Yue dialects, MC 開口 non-rounded 止 [i] group of finals with 齒音 coronal initial is pronounced with 舌面元音 “normal” vowel [-i], in some Yue dialects MC 開口 non-rounded 止 [i] group of finals with 精 [ts-] 莊 [tʂ-] series of initials is pronounced as [-y] or [-u].]

[資 廣州 Canto. tsi˧ 桂平 Canto. tsɿ˧ 百色 Canto. ɬɿ˧ 順德大良 tsy˧]

(7) 长短元音 [a] 与 [ɐ(e)] 基本成系统对立，如广州话：街 kai˧ ≠ 鸡 kei˧，三 ɟam˧ ≠ 心 ɟem˧，蛮 man˧ ≠ 民 men˧，彭 pʰaŋ˧ ≠ 凭 pʰeŋ˧。在音质上，[a、ɐ] 不但有时值长短的区别，也同时有舌位高低的区别。

[Long and short vowels [a] and [ɐ(e)] basically form a systematic contrast, for example Cantonese 街 kai˧ ≠ 鸡 kei˧，三 ɟam˧ ≠ 心 ɟem˧，蛮 man˧ ≠ 民 men˧，彭 pʰaŋ˧ ≠ 凭 pʰeŋ˧. In terms of quality, [a, ɐ] not only have a temporal length difference, there is also a difference in tongue height]

Yue-Pinghua LAC 2012

(8) 流摄、臻摄开口一等与三等字不分洪细音，多数方言同韵。

[MC 流 -iw and 臻 -in groups of finals, first (uw ən) and third division (juw jiw (j)in jin) characters, no matter whether it is followed by 洪 [open / u] or 細 [i/y] vowels, in most dialects they are the same final]

[Canto. 樓 = 流 lau4, 走 = 酒 zau2, 狗 = 九 gau2, 根 = 斤 gan1]

(9) 古阳声韵和入声韵今分别读 [-m、-n、-ŋ] 和 [-p、-t、-k] 两套阳、入配的韵尾。

[MC 陽聲韻 nasal-ending finals and 入聲韻 plosive-ending finals are now read with [-m -n -ŋ] [-p -t -k] two sets of nasal and plosive codas]

[Canto. 針 zam1 文 man4 江 gong 1 汁 zap1 物 mat6 覺 gok3]

Yue-Pinghua LAC 2012

(10) 平、上、去、入四调均分阴阳，入声调中阴入上下两分，有的阳入也上下两分，所以入声调通常有 3 个或 4 个。

[Tones ABCD all split into upper and lower, tone D1 is also split into two, in some dialects tone D2 is also split into two, so there are usually 3 or 4 tone D's.]

[Canto. tone DS1 必 bit1 (*bat1), tone DL1 鰲 bit3, tone D2 別 bit6]

(11) 常用词汇中，家多称“屋企”，东西多称“嘢”，公鸡、母鸡说成“鸡公”、“鸡𨵿”，什么说成“乜”，视觉动词为“睇”，宰杀称“割”，判断动词多为“系”。

[Amongst common vocabulary, “home” is 屋企 uk1 kei2, “thing” is 嘢 je5, rooster and hen are 雞公 gai1 gung1, 雞𨵿 gai1 naa2, “what” is 乜 mat1, visual verb is 睇 tai2, “to butcher” is 割 tong1, copula is usually 係 hai6.]

Yue-Pinghua LAC 2012

(12) 量词独用具有指示功能，如“件衫好靚”（这件衣服很漂亮）、“部车坏左”（这辆车坏了）、“啲人走晒”（这些人跑光了），其中“件”、“部”、“啲”分别相当于北京话“这件”、“这辆”、“这些”。表示动作行为或现象持续的助词用“紧”，如“讲紧”（正在说话），“落紧雨”（正下着雨）；表示动作行为回复的助词一般都用“翻”，如“找翻两文我”（找回我两块钱）、“做翻先前 [啲嘢]”（重新做起以前的那些工作）。

[A classifier on its own has deitic function, e.g. 件衫好靚 gin6 saam1 hou2 leng3 (CLF shirt) very beautiful) ‘The shirt is beautiful’, 部車壞咗 bou6 ce1 waai6 zo2 (CLF vehicle bad PFV) ‘The car broke down’, 啲人走晒 di1 jan4 zau2 saai3 (CLF:MASS person leave COMPL) ‘All the people left’. 件 部 啲 is equivalent to Beijing dialect 這件 這輛 這些 respectively. The progression of an action or state is expressed by the particle 紧 gan2, e.g. 講紧 gong2 gan2 (speak PROG) ‘X is speaking’, 落紧雨 lok6 gan2 jyu5 (descend PROG rain) ‘it is raining’; returning action is expressed by the particle 翻 faan1, e.g. 找翻兩文我 zaau2 faan1 loeng5 man4*1 ngo5 (give.change RETURN two \$ I) ‘give me change of \$/¥2 back’, 做翻先前啲嘢 zou6 faan1 sin1cin4 di1 je5 (do RETURN before CLF:MASS thing ‘Doing things [which I did] before again’.

Yue-Pinghua LAC 2012

Approach in the LAC: top-down

Preconceived dialect groups, then describe traits common in the group (not necessarily traits which are highly exclusive or representative)

When dealing with atypical cases, often 1) lump into existing dialect groups, or 2) put into a “waste basket” category

Unclassified Sinitic dialects in S Hunan: 湘南土話

Unclassified Sinitic dialects in N Guangdong: 粵北土話

“Everything” in Guangxi: 「平話」 “Pinghua”

(All Sinitic varieties which are not considered Yue, Mandarin, Hakka, Xiang, or Min)

Yue-Pinghua LAC 2012

Phonology, typologically:

Southern Pinghua 桂南平話 varieties are like Yue,
conservative with codas, e.g. (南寧位子淥):

六 luk↓

七 tʃ^het↓

心 ɬəm↓

門 mun↓

Northern Pinghua 桂北平話 varieties are like Mandarin,
usually no *-m -p -t -k* codas, e.g. (富川秀水):

六 liou↓

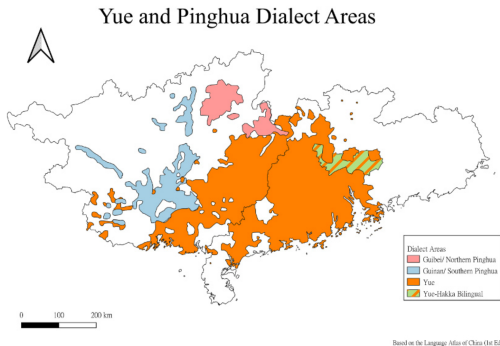
七 ts^hai↓

(local trait) sometimes nasal codas are totally deleted, e.g.

心 sie↓

門 mə↓

Yue–Pinghua LAC 2012



(map from Sung 2026)

1st ed. LAC:

N 平話: 賀州/桂林

pref. in NE 廣西

S 平話: others,
including 柳州/河池
pref. in mid-N 廣西

2nd ed. LAC:

(Non-官) dialects in
柳州/河池 doubly
counted as (勾漏) 粵
and S 平話

Figure 2.12: Geographical distribution of Yue, Guinan and Guibei Pinghua dialects (based on LAC 1st Edition)

Yue–Pinghua LAC 2012

In LAC:

“Yue” dialects in Guangxi:

- 1. related to migration from Guangdong, or areas redrawn from Guangdong*
- 2. remained typologically like Yuè dialects of Guangdong*

“Pínghuà”:

- 1. Han ancestors mostly came from or via Hunan*
- 2. LAC includes Mandarinised-Yue dialects in NE Guangxi as N Pinghua (e.g. 陈 2009, 覃 2019)*

Chén, Ránrán 陈然然. 2009. 灵川杨家话语音研究. In 郑作广 & 周本良 (eds.), 广西汉语方言珍稀方言语音研究, 343–391. Nanning: Guangxi Ethnic Publishing House 广西民族出版社.

Qín, Yuǎnxióng 覃远雄. 2019. Guǎngxī Pínglè, Zhōngshān “Běndìhuà” de xìngzhì 广西平乐、钟山“本地话”的性质 ‘The Local Vernacular’ in Pingle County and Zhongshan County. Fāngyán 方言 2019(4). 484–495.

Yue–Pinghua Sung 2026

Dialectometry, e.g. (Sung 2026) (: 38)

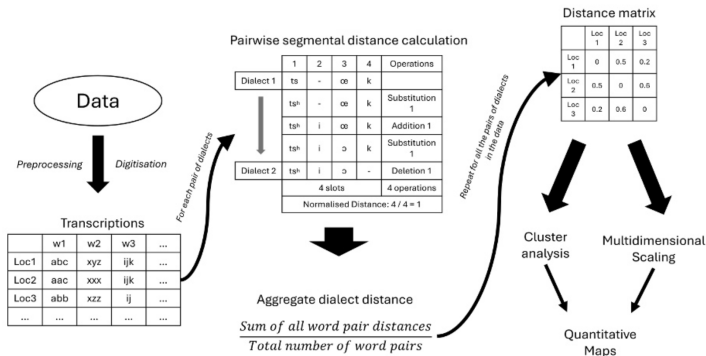


Figure 3.2: Workflow of Dialectometry using Levenshtein Distance

Yue-Pinghua Sung 2026

Pairwise segmental distance calculation, e.g. 雀:

Dialects	Onset	Medial	Nucleus	Coda	Operations	Costs
Guangzhou	ts	-	œ	k		
	ts^h	-	œ	k	Substitution from ts to ts ^h	1
	ts ^h	i	œ	k	Insertion of i	1
	ts ^h	i	ɔ	k	Substitution from œ to ɔ	1
	ts ^h	i	ɔ	-	Deletion of k	1
Bao'an	ts ^h	i	ɔ	-		4

Figure 3.3: Illustration of Levenshtein Distance

Yue-Pinghua Sung 2026

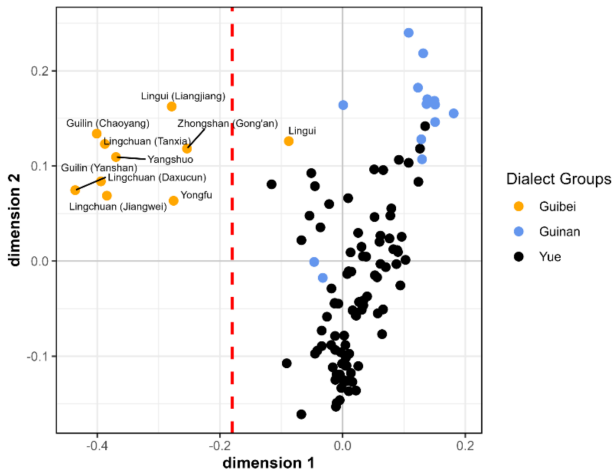


Figure 5.1: MDS plots of the segmental distances of 113 Yue and Pinghua dialects (Levenshtein distance), $r^2 = 0.59$

“Lingui” =
臨桂五通

Sung (2026:66)

Yue-Pinghua Sung 2026

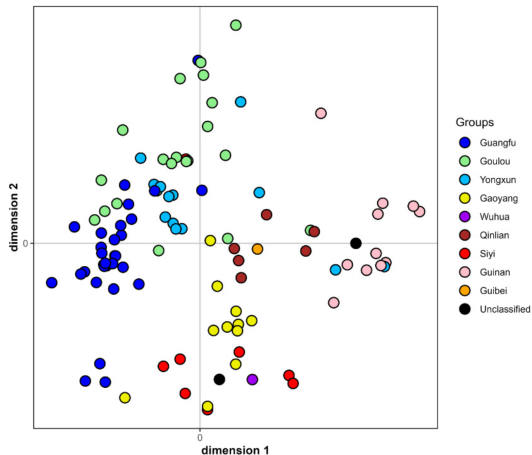


Figure 5.5: MDS plot of 104 Yue dialects (Figure 5.3) with LAC dialect group annotation (Dimension 1 & 2)

Yue-Pinghua Sung 2026

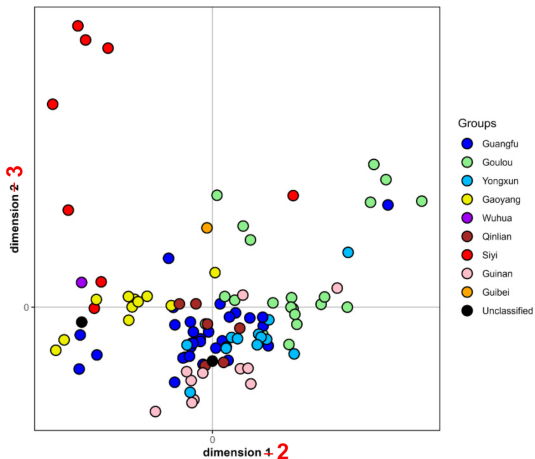


Figure 5.6: MDS plot of 104 Yue dialects (Figure 5.3) with LAC dialect group annotation (Dimension 2 & 3)

Yue–Pinghua Sung 2026

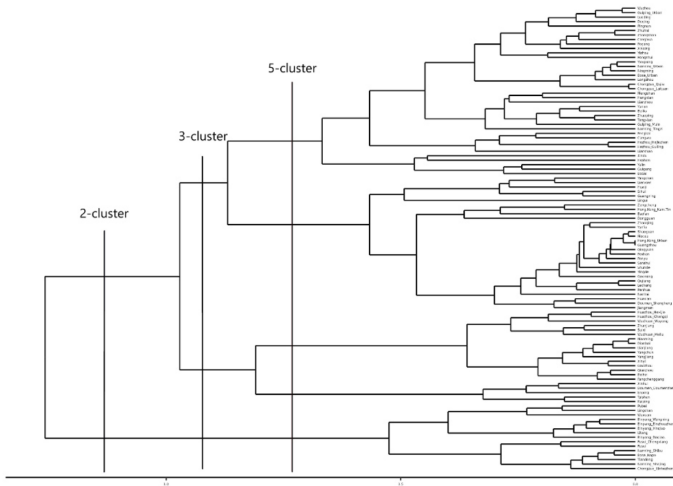
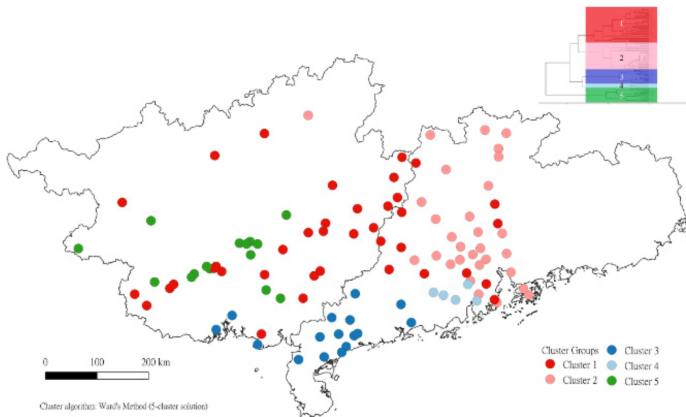


Figure 5.12: Dendrogram of 104 Yue segmental distances using Ward's method

Yue-Pinghua Sung 2026



(c) *5-cluster solution*

Figure 5.13: Cluster maps based on the Ward's method

Yue-Pinghua Sung 2026



Tones:

Cantonese or similar

core 四邑

Sung (2026:163)

Figure 8.3: MDS plot of aggregate tone distances between Yue dialects, $r^2 = 0.49$

Yue-Pinghua Sung 2026

Amongst the many merits of Sung (2026):

- criticisms of the subgrouping of Yue and Pinghua dialects in the LAC
- dealing with a very large number of features (no cherry-picking of which features should have more weighting)
- [...]

One criticism: Used the LAC's definition of Yue and Pinghua as a starting point. My questions are, e.g.

- Why should the 四邑 dialects be considered Yue?
- Should some of the Hakka / Hakka-like dialects be considered Yue? (i.e. is the Yue-Hakka boundary as clear as assumed?)

(Sung & de Sousa forthcoming)

Yue-Pinghua Chan 2025

Historical linguistics (i.e. using only common innovations),
e.g. Chan (2025):

15 features of 泛粵 Pan-Yue , most of them innovations

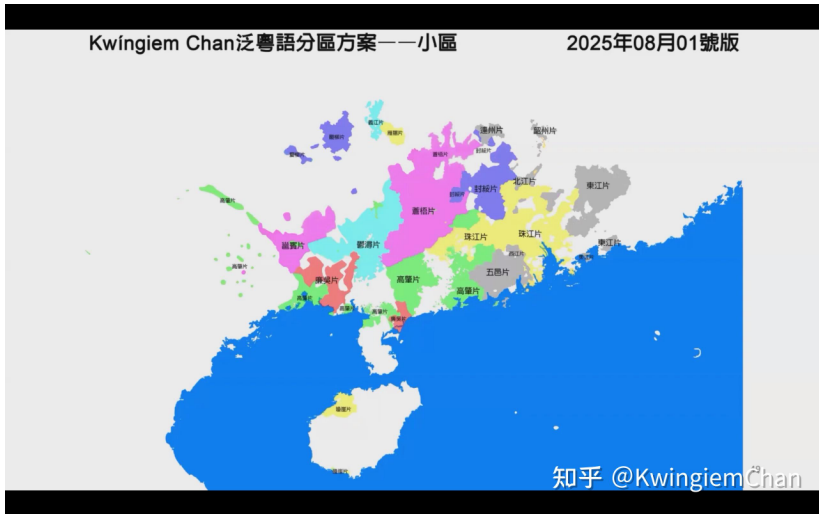
(We will only talk about some of these; examples provided by me)

“泛粵 Pan-Yue” includes the following out of LAC’s dialect groups:

- Yue 粵 (**excluding** 四邑 (a.k.a. 五邑))
- 儋州 Dānzhōu / Daam1 Zau1 & 邁 Má / Maai4 of Hainan
- S Pinghua
- Parts of N Pinghua (W/S of 桂林 & most of 賀州)
- Parts of N Guangdong Patois 粵北土話 (around 連山)

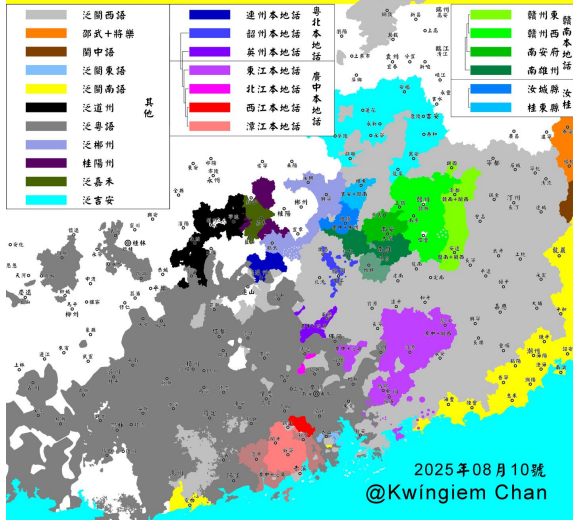
“Mixed areas”: e.g. 五邑、東江

Yue-Pinghua Chan 2025

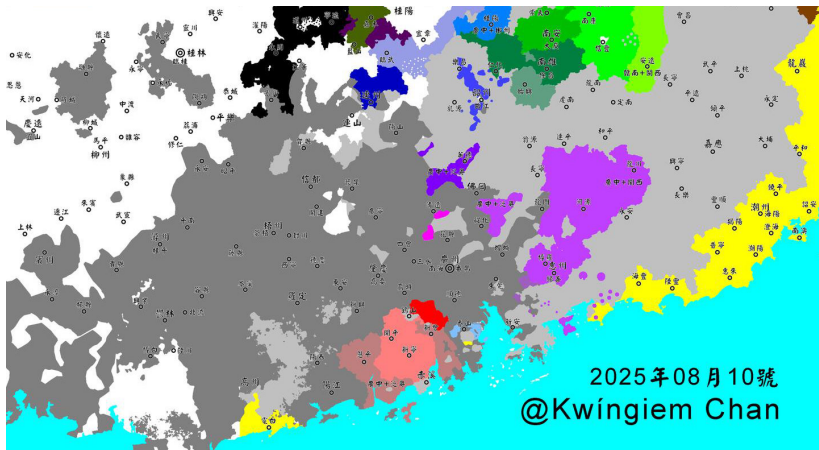


Yue-Pinghua Chan 2025

贛粵本地話家族分佈圖



Yue-Pinghua Chan 2025



Yue-Pinghua Chan 2025

1) 內轉一等韻三等韻主元音央低化 (至少不在 i u 的位置上) :
此點指九狗巾根金互吉乞急刻等字的主元音不高於 ə 。 [...]
主元音為 [ɐ 或 ə 或 ʊ 的] 現象。

Basically: i/u splitting & lowering, e.g. -iC > -jəC (> -əC)

MC	狗 kuw ^B	九 kjuw ^B	陳 ɕin ^A	十 dzip ^D	金 kim ^A
香港客	kiw↓	kiw↓	ts ^h in↓	sip↓	kim↓
惠城	kjaw↓	kiw↓	tɕ ^h in↓	ɕip↓	kim↓
台山	kew↓	kiw↓	ts ^h in↓	sip↓	kim↓
儋州 ^文	kɔw↓	kjew↓	sɔn↓	sip↓	kim↓
Sn-Viet	< cẩu >	< cửu >	< trần >	< thập >	< kim >
儋州 ^白	kaw↓	kɔw↓	sɔn↓	tɔp↓	kɔm↓
Canto.	kew↓	kew↓	ts ^h en↓	sep↓	kem↓
陽江	kew↓	kjew↓	ts ^h en↓	sep↓	kjem↓
南寧平	kew↓	kəw↓	tʃen↓	ʔep↓	kem↓

Yue-Pinghua Chan 2025

2) 外轉元音縮合 [...] 元音為 i、y 和 u 的現象。

[Basically, e.g. -jeC > -iC, -woC > -uC, -wjeC > -wiC (> i/u/yC)]

MC	鳥 tew ^B	占 tɕjem ^A	軟 n̩wjen ^B	葉 jep ^D	半 pwan ^C
香港客	tjaw↓	tsam↓	n̩jon↓	n̩jap↓	pan↓
惠城	njew↓	tɕjem↓	n̩ɣen↓	jep↓	pɔn↓
台山	njaw↓	tsjam↓	n̩un↓	jap↓	p ^u ɔn↓
儋州 ^文	njew↓	tɕjem↓	[ɜwan↓]	zjep↓	ɓɔn↓
Sn-Viet	<điêu>	<chiêm>	<nhuyễn>	<diệp>	<bán>
儋州 ^白	(ne↓)	tɕim↓	[ɜun↓]	zip↓	ɓɔn↓
Canto.	niw↓	tsim↓	jyn↓	jip↓	pun↓
陽江	niw↓	tsim↓	jin↓	j/hip↓	pun↓
南寧平	niw↓	tɕim↓	win↓	hip↓	pun↓

Yue-Pinghua Chan 2025

6) 魚虞合併前有一層虞入模層次。[Before the merger of MC 魚 jə 虞 ju, there is a layer of MC 虞 ju merger into MC 模 u]

(you expect Canto. yu/eoi, but some words are pronounced u/ou)

At the time of 切韻 (601 隋)

Standard N Chinese (middle Yellow) 魚 jə = 虞 ju

Standard S Chinese (lower Yangtze) 魚 jə ≠ 虞 ju

據 MC ɲjə^C: Sn-Viet <cư>, 潮州 /kɯ1/, Sn-Kor 거 /kʌ/

句 MC ɲju^C: Sn-Viet <cú>, 潮州 /ku1/, Sn-Kor 구 /ku/

Canto. 據 = 句 = geoi3 (perhaps true of all Pan Yue dialects)

Nonetheless, typically Pan Yue dialects have a few characters which have kept this distinction, with MC 虞 ju pronounced like MC 模 u.

	MC	(Canto. norm)	Canto.
徐	zjə ^A	yu/eoi	ceoi4
[須] 須	sju ^A	yu/eoi	seoi1
[須] 鬚	sju ^A		sou1
酥	su ^A	u/ou	sou1

Yue-Pinghua Chan 2025

Some more examples in Canto. (other Pan Yue dialects are similar)

	MC	Canto. norm	Canto.
[須] 鬚	sju^A	yu/eoi	sou1
芋	hju^C	yu/eoi	wu6
文 娶	ts ^h ju ^B	yu/eoi	ceoi2
白 娶	ts^hju^B	yu/eoi	cou2
雛	dzju ^A	o	co1
鋤	dzjə ^A	o	co4
數	sju^{B/C}	o	sou2/3

See more e.g. Kwok 郭 (2004) on Yue, Chiang 江 (2014) on Sino-Vietnamese

Kwok, Bit-Chee 郭必之. 2004. 從虞支兩韻「特字」看粵方言跟古江東方言的聯繫 Evidence for an Ancient Jiangdong layer in the Yue dialects as revealed in “special words” of the Yu and Zhi Rhymes. *Language and Linguistics* 《語言暨語言學》5(3). 583–614

Chiang, Chia-lu 江佳路. 2014. 析論越南漢字音魚虞分韻的歷史層次 Discussion on the phonological strata of Sino-Vietnamese as reflected in the distinction between rhymes Yu (魚) and Yu (虞). *Language and Linguistics* 《語言暨語言學》15(5). 613–634.

Yue-Pinghua Chan 2025

[We will talk about these later in §3.3:]

14) 聲調大致平上去入，陰陽未分化，濁上未歸去：
具有長短元音對立或長短入聲分化 (語音特徵)。
(或有曾分的痕跡)

[[Proto-Yue:] Tones are basically ABCD, tones not yet split into Upper and Lower, [morphemes in tone] Lower B has not shifted to C: there is a contrast of long vs. short vowels or split into L vs. S tone D (phonological characteristics). (or there exist traces thereof)]

15) 大量使用粵語特色詞彙，排除現代借詞仍有大量非漢源詞彙

[Using many words emblematic to Yue, after excluding modern loans [from Modern Written Chinese] there are still many words of non-Sinitic origin]