

Yuè & the Sinitic Languages of the Pearl River Region — Unit 2.2

13th EACL Summer School, Universität Stuttgart

Hilário de Sousa 蘇沙

EHESS–CRLAO
membre associé

Tuesday 28 July 2026
丙午六月十五

hilario@bambooradical.com

Middle Chinese

The vast majority of linguistic material in Cantonese can be traced to Middle Chinese

“Early Middle Chinese” in Chinese historical phonology:

《切韻》 qièyùn / cit3 wan5 (rime book; 601 隋)

Only parts of it have survived; but there are many augmented versions

e.g. 《廣韻》 guǎngyùn / gwong2 wan5 (1008; official rime book of 宋)

Author of 《切韻》: 陸法言 lù fǎyán / luk6 faat3 jin4 鮮卑

隋 Suí (581-618) unified China in 589 CE after $\approx$ 300 years of disunity

Eight scholars (friends of 陸法言's father) were trying to create a new unified standard, based on a compromise of the “best” pronunciation of characters from the old northern and southern standards

Middle Chinese

切韻卷第四 去聲 五十六韻

凍 瀑 [雨] 又水名出發
鳩山 多貢反 三

[粵 多 do1 貢 gung3]

[MC 多 ta^A 貢 ku^wŋ^C =
tu^wŋ^C]

凍 冰

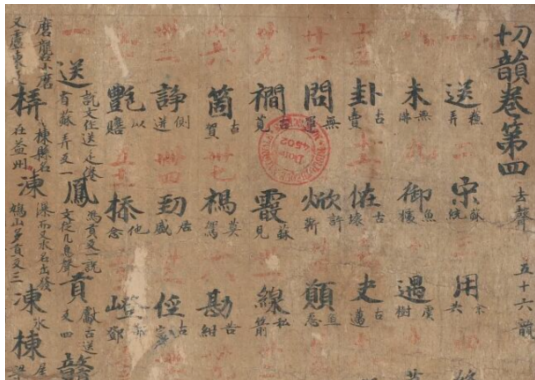
棟 屋梁

[XY 反, later XY 切]

Hence the name of this

method: 反切 faan2 cit3 ~

faan1 cit3 / fǎnqiè]



Middle Chinese

國 際 音 標 方 言 點	編 號	1911
	漢 字	法
	中 古 音	方法
	音 標	咸合三入 乏非
北 京	fa ²¹⁴	
廣 州 (市 區)	fat ³³	
香 港 (市 區)	fat ³³	
香 港 (新 界 錦 田)	fæk ³⁵	
澳 門 (市 區)	fat ³³	
番 禺 (市 橋)	fat ³³	
花 縣 (花 山)	fat ³³	
從 化 (城 內)	fat ²³	
增 城 (縣 城)	fak ³³	
佛 山 (市 區)	fat ³³	
南 海 (沙 頭)	p'et ³³	

法：咸合三入 乏非

- 咸：咸攝 (-am/-ap), 1 of 16 groups of finals
- 合：合 “rounded” vs. 開 (non-rn)
- 三：1234 “divisions” of finals
- 入：平上去入 Tones ABCD
- 乏：乏 -jəp^D final (廣韻 has $57^A + 55^B + 60^C + 34^D = 206$ finals)
- 非：非 pf- [LMC], 1 of about 40 initials
(非乏 pfjəp^D)

Middle Chinese

Transcriptions of Middle Chinese:
do not carry * (because not reconstructions)

Various transcription systems of MC, e.g.
Baxter's transcription for Middle Chinese

https://en.wikipedia.org/wiki/Baxter%27s_transcription_for_Middle_Chinese

Hill's Indological transcription of Middle Chinese

<https://doi.org/10.1163/19606028-bja10028>

Here, my own transcription of MC is used
(mostly based on Baxter's)

MC Tones

sonorant-ending			-p -t -k
A 平 'Level' ㄅㄟ	B 上 'Rising' ㄅㄟ	C 去 'Departing' ㄅㄟ	D 入 'Entering' ㄅㄟ
bjæŋ ^A ping4 píng	dʒjaŋ ^B soeng5 shǎng	k ^{hi} ə ^C heoi3 qù	nip ^D jap6 ⁽⁹⁾ rù

[https://en.wikipedia.org/wiki/Four_tones_\(Middle_Chinese\)](https://en.wikipedia.org/wiki/Four_tones_(Middle_Chinese))

MC Tones

Tone splitting:

e.g. Proto-Yue

/taj^C/ [taj[˥]]

/daj^C/ [daj[˥]]

tone-splitting may occur:

/taj^C/ [taj[˥]]

/daj^C/ [daj[˥]] (voiced onset lowering pitch)

In most, but not all, Sinitic languages, voiced initials have become voiceless

tone-splitting becomes phonemicised:

/taj^{C1}/ /taj[˥]/ Canto. 帶 daai3

/taj^{C2}/ /taj[˥]/ Canto. 大 daai6

cf. Cantonese tone 3 /˥/ 6 /˥/ both descended from MC tone C

cf. Standard Mandarin where MC tone C has not split:

帶 dài, 大 dài (as in 大夫 dài fu)

MC Tones

	* 平 *A	* 上 *B	* 去 *C	* 入 *D	
				長 L	短 S
陰 *voiceless 清 initial	1 ˩(˨˩)	2 ˩˨˩	3 ˩	3 ˩	1 ˩
陽 *voiced sonorant 次濁 initial	4 ˩˨˩	5 ˩	6 ˩	6 ˩	
陽 *voiced obstruent 全濁 initial					

Table 1: Tones in Standard Cantonese

	* 平 *A	* 上 *B	* 去 *C	* 入 *D
陰 *voiceless 清 initial	1 ˩	3 ˩	4 ˩	<i>random</i>
陽 *voiced sonorant 次濁 initial	2 ˩			4 ˩
陽 *voiced obstruent 全濁 initial				2 ˩

Table 2: Tones in Standard Mandarin

MC Tones

Examples of 全濁上歸去：

倚 MC gie^B ‘to stand’

Mand. ji4 ^(C)

Canto. kei5 ^(B2) 企 (MC khie^B) for ‘to stand’ is etymologically wrong

被 MC bie^B ‘to cover’

Mand. bei4 ^(C)

Canto. pei5 ^(B2) ‘quilt’; bei6 ^(C2) PASS

技 MC gie^B ‘technique’

Mand. ji4 ^(C)

Canto. gei6 ^(C2)

(cf. HK Hakka ki3 ^(B) Hakka has kept this character in tone B)

MC Tones

Lower tone D:

月 MC ɲwɟət^D ‘moon’

Canto. jyut6^(D2)

Mand. yue4^(C)

別 MC bjet^D ‘to part’

Canto. bit6^(D2)

Mand. bie2^(A2)

MC Tones

平 A	上 B	去 C	入 D
舒聲 smooth			入聲 checked
syu1 sing1			jap6 sing1
shū shēng			rù shēng
平聲 level	仄聲 oblique		
ping4 sing1	zak1 sing1		
píng shēng	zè shēng		

MC Tones

〈登鶴雀樓〉王之渙 688–742 唐/周/唐

白 日 依 山 盡	仄 仄 平 平 仄
黃 河 入 海 流	平 平 仄 仄 平
欲 窮 千 里 目	仄 平 平 仄 仄
更 上 一 層 樓	仄 仄 仄 平 平

pək↓ jət↓ əj↓ ʔan↓ tʃen↓
 hun↓ ho↓ nep↓ haj↓ ləw↓
 nuk↓ kun↓ tʃʰin↓ ləj↓ muk↓
 kɛŋ↓ ʔɛŋ↓ ət↓ tʃɛŋ↓ ləw↓

Bạch nhật y sơn tận
 Hoàng Hà nhập hải lưu
 Dục cùng thiên lý mục
 Cánh thương nhất tầng lâu

baak6 jat6 ji1 saan1 zeon6
 wong4 ho4 jap6 hoi2 lau4
 juk6 kung4 cin1 lei5 muk6
 gang3 soeng5 jat1 cang4 lau4

pik-jit i-san tsīn
 hông-hô jip-hái liū
 iòk-kiông tsian-lí-bòk
 kình siong it-tsing-liô

MC initials & Late Middle Chinese

Early Middle Chinese: 《切韻》 qièyùn / cit3 wan5 (601 隋)
(artificial: compromise between N and S Standard Chinese)

Organised by tones and finals
no organised information on the initials

Late Middle Chinese: rime tables 韻圖 yùntú / wan5 tou4
《韻鏡》 yùnjìng / wan5 geng3 ≈10th c 五代; 1161 南宋 ed.)
《七音略》 qīyīnlüè / cat1 jam1 loek6 (before 1161 南宋)

information on tones and finals: basically the same as 《切韻》
information on initials: current (10th century)

(There are rime tables like 《四聲等子》 (1039 北宋) and 《切韻指掌圖》
(1230 南宋) which reflected pronunciation at the time better)

MC initials & Late Middle Chinese

	齒音古		音		喉		音		齒	
	清	濁	清	濁	清	濁	清	濁	清	濁
東	○	龍	○	洪	○	洪	○	揭	○	叢
	○	隆	○	雄	○	雄	○	崇	○	崇
	○	融	○	融	○	融	○	嵩	○	嵩
董	○	龍	○	洪	○	洪	○	揭	○	叢
	○	隆	○	雄	○	雄	○	崇	○	崇
	○	融	○	融	○	融	○	嵩	○	嵩
送	○	弄	○	閑	○	閑	○	送	○	送
	○	○	○	趨	○	趨	○	銳	○	銳
	○	○	○	趙	○	趙	○	趙	○	趙
屋	○	祿	○	穀	○	穀	○	速	○	速
	○	肉	○	園	○	園	○	塾	○	塾
	○	育	○	育	○	育	○	肅	○	肅

	音		音		音		音		音	
	清	濁	清	濁	清	濁	清	濁	清	濁
牙	○	次	○	次	○	次	○	次	○	次
	○	空	○	空	○	空	○	空	○	空
	○	窮	○	窮	○	窮	○	窮	○	窮
舌	○	同	○	同	○	同	○	同	○	同
	○	蟲	○	蟲	○	蟲	○	蟲	○	蟲
	○	仲	○	仲	○	仲	○	仲	○	仲
唇	○	蒙	○	蒙	○	蒙	○	蒙	○	蒙
	○	蓬	○	蓬	○	蓬	○	蓬	○	蓬
	○	豐	○	豐	○	豐	○	豐	○	豐
內轉第一開	○	孔	○	孔	○	孔	○	孔	○	孔
	○	環	○	環	○	環	○	環	○	環
	○	動	○	動	○	動	○	動	○	動
牙	○	控	○	控	○	控	○	控	○	控
	○	貢	○	貢	○	貢	○	貢	○	貢
	○	燭	○	燭	○	燭	○	燭	○	燭
舌	○	洞	○	洞	○	洞	○	洞	○	洞
	○	仲	○	仲	○	仲	○	仲	○	仲
	○	中	○	中	○	中	○	中	○	中
唇	○	夢	○	夢	○	夢	○	夢	○	夢
	○	鳳	○	鳳	○	鳳	○	鳳	○	鳳
	○	賜	○	賜	○	賜	○	賜	○	賜
牙	○	哭	○	哭	○	哭	○	哭	○	哭
	○	獨	○	獨	○	獨	○	獨	○	獨
	○	秃	○	秃	○	秃	○	秃	○	秃
舌	○	逐	○	逐	○	逐	○	逐	○	逐
	○	蓄	○	蓄	○	蓄	○	蓄	○	蓄
	○	竹	○	竹	○	竹	○	竹	○	竹
唇	○	目	○	目	○	目	○	目	○	目
	○	伏	○	伏	○	伏	○	伏	○	伏
	○	蛟	○	蛟	○	蛟	○	蛟	○	蛟

Figure 1: first table in 韻鏡

MC initials & Late Middle Chinese

此圖每韻呼吸四聲字並屬之	齒音		舌音		喉音		齒音		舌音		唇音	
	清濁		清濁		清濁		清濁		清濁		清濁	
	日		來		喻		匣		曉		影	
	半微半商								喉音三獨立			
○三十六字母	疑		群		溪		見		泥		定	
									娘		澄	
									古上音		古頭音	
									唇音輕		唇音重	
○歸納助紐字	銀		勤		輕		經		盛		廷	
	言		皮		津		堅		年		田	

Figure 2: 36 字母 / initials in 韻鏡

MC initials & Late Middle Chinese

重 中 重	日	來	喻	匣	曉	影	邪	心	從	清	精	欽定四庫全書													疑	羣	溪	見	泥	定	透	端	明	並	滂	幫	內轉第一
							邪	審	未	穿	照	商													枕	羣	溪	見	泥	定	透	端	明	並	滂	幫	平
													宮													角	空	公	徵	同	通	東	蒙	達	降	風	
	東	龍	洪	烘	翁	摠	叢	葱	終	窮	穹	弓	蟲	忡	中	魯	馮	豐	風																		
	戎	隆	雄	設	碩	充	充	充	充	窮	穹	弓	蟲	忡	中	魯	馮	豐	風																		
	董	龍	融	頤	翁	嵩	庭	總														渴	孔	鏤	動	桶	董	蛟	擇	珠	上						
送	弄	哄	烘	翁	送	穀	刺	聰	衆														控	貢	鑣	洞	痛	凍	懔	槎	去						
										銃	衆														烤			仲	蟲	中	夢	鳳	賄	諷			
										趙	衆																										
										疾	緘	緘																									
										族	疾	緘	緘																								
										速	族	疾	緘	緘																							
										縮	族	疾	緘	緘																							
										叔	族	疾	緘	緘																							
										肅	族	疾	緘	緘																							
屋	祿	殺	屋	連	族	疾	緘	緘	哭	殺	獨	禿	穀	木	瀑	汴	卜	入																			
肉	六	圓	蒿	郁	肅	肅	肅	肅	哭	殺	獨	禿	穀	木	瀑	汴	卜	入																			

Figure 3: first table in 七音略

四等 sì děng / sei3 dang2 'division' or 'grade'

Early MC initials

Initials in Early Middle Chinese:

Information not displayed systematically in 《切韻》

陳澧 can4 lai5 / chén lǐ (1810—1882 Canton)

《切韻考》 he worked from 《廣韻》, ignored characters that were not in 《切韻》

「系聯法」 xìliánfǎ / hai6 lyun4 faat3, e.g.

(19th century Cantonese:)

深 šam1: 式 šik1 針 žam1 切

式 šik1: 賞 šoeng2 職 žik1 切

賞 šoeng2: 書 šyu1 兩 loeng5 切 etc...

i.e. 深式賞書... share the same initial

Discovery: EMC 《切韻》 Initials ≠ LMC rime table Initials

Early MC initials

全清	次清	全濁	全清	全濁	次濁	次濁
幫 [幫] bong1 [p] bāng	滂 pong1 [pʰ] pāng	並 [並] bing6 [b] bìng			明 ming4 [m] míng	
端 dyun1 [t] duān	透 tau3 [tʰ] tòu	定 ding6 [d] dìng			泥 nai4 [n] ní	來 loi4 [l] lái
知 zi1 [tʃ] zhī	徹 [徹] cit3 [tʰ] chè	澄 cing4 [dʒ] chéng			娘 noeng4 [ŋ] niáng	
精 zing1 [ts] jīng	清 cing1 [tsʰ] qīng	從 [从] cung4 [dz] cóng	心 sam1 [s] xīn	邪 ce4 [z] xié		
莊 [庄] zong1 [tʃʒ] zhuāng	初 co1 [tʃʰ] chū	崇 sung4 [dʒʒ] chóng	生 sang1 [ʃ] shēng	俟 zi6 [z] sì		
章 zoeng1 [tʃʒ] zhāng	昌 coeng1 [tʃʰʒ] chāng	常 soeng4 [dʒʒ] cháng	書 [书] syu1 [ʃʒ] shū	船 syun4 [ʒ] chuán	日 jat6 [ɲ] rì	以 ji5 (0j) yǐ
見 gin3 [k] jiàn	溪 kai1 [kʰ] xī	羣 [群] kwan4 [g] qún			疑 ji4 [ŋ] yí	
影 jing2 [ʔ] yǐng			曉 [晓] hieu2 [h] xiǎo	匣 hap6 [ɦ] xiá		云 wan4 (ɦj) yún

Early MC initials

Some notes:

- not all were phonemically contrastive, e.g. 泥 [n] 娘 [ŋ] were in complementary distribution
- in Canto., 清 voiceless = high tone, 濁 voiced = low tone (MC 全濁 voiced obstruents have become voiceless)
- Traditional 並 > Simplified 并; but linguists in Mainland China writes MC initial [b] as 並 (Simp. 并 < Trad. 并 MC [p] or 併 [p/b] or 並 [b])
- 俟 MC [z] only has two characters (俟𪔐) in 切韻, both merged into 崇 [dʒ] in 廣韻 (so [z] is often ignored)
- (excluding 俟 [z]) In Cantonese, MC retroflex 莊初崇生 (zong1 co1 sung4 sang1) often have back vowels
MC palatal 章昌常書船 (zoeng1 coeng1 soeng4 syu1 syun4) often have front vowels (rules not 100%)

Late MC initials

全清	次清	全濁	全清	全濁	次濁	次濁
幫 [幫] bong1 [p] bāng	滂 pong1 [pʰ] pāng	並 [並] bing6 [bʰ] bìng			明 ming4 [m] míng	
非 fei1 [f] fēi	敷 fu1 [f] fū	奉 fung6 [vʰ] fèng			微 mei4 [ŋ] wéi	
端 dyun1 [t] duān	透 tau3 [tʰ] tòu	定 ding6 [dʰ] dìng			泥 nai4 [n] ní	來 loi4 [l] lái
知 zi1 [t] zhī	徹 [徹] cit3 [tʰ] chè	澄 cing4 [dʰ] chéng			娘 noeng4 [ŋ] niáng	
精 zing1 [ts] jīng	清 cing1 [tsʰ] qīng	從 [从] cung4 [dzʰ] cóng	心 sam1 [s] xīn	邪 ce4 [zʰ] xié		
照 ziu3 [tʂ] zhào	穿 cyun1 [tʂʰ] chuān	牀 [床] cong4 [dʒʰ] chuáng	審 [审] sam2 [ʂ] shěn	禪 [禅] sim4 [ʐʰ] chán	日 jat6 [ʌ] rì	
見 gin3 [k] jiàn	溪 kai1 [kʰ] xī	羣 [群] kwan4 [gʰ] qún			疑 ji4 [ŋ] yí	
影 jing2 [ʔ] yǐng			曉 [晓] hui2 [h] xiǎo	匣 hap6 [ɦ] xiá		喻 jyu6 (0j) yù

Late MC initials

Some notes:

- voiced obstruents became breathy, and were on their way to become voiceless (in Northern Chinese)
- labiodentals 非敷奉微 developed out of bilabials 幫滂並明 (輕唇 hing1 seon4 vs. 重唇 cung5 seon4 respectively)
- 微 TW Mand. wéi, BJ Mand. wēi. However, when referring to the MC initial 微, N Chinese people may change their pronunciation to 微 wéi
- 非敷 were earlier [pf] [pʰ]; soon they merged to [f]
- Cantonese usually have /f/ for 非敷奉;
but some /p//pʰ/ are retained, e.g.
番 faan1 but 番禺 pun1 jyu4 (Pānyú)
伏 fuk6 but 伏匿匿 buk6 nei1 nei1 ‘hide and seek’

Late MC initials

- In Cantonese, LMC [f/v] 非敷奉 have developed, but LMC [m] 微 (Mand. [v] 微) has not (between EMC and LMC; $\approx 9^{\text{th}}-10^{\text{th}}$ C, late 唐 / 5 代)
- MC [t][tʰ][d] 知徹澄 have become Cantonese [ts][tsʰ] some exceptional plosive reflexes e.g. 秩 dit6 (cf. Mand. zhì)
- MC 知 [t] 照 [tʂ] series > Mand. [tʂ][tʂʰ][ʂ]
MC 精 [ts] series > Mand. [ts][tsʰ][s] or [tʃ][tʃʰ][ʃ]
Canto. all > [ts][tsʰ][s]; but the distinction existed earlier e.g. 牛池灣 Ngau Chi Wan (ngau4 či4 waan1) vs. 慈雲山 Tsz Wan Shan (cz4 wan4 šan1) (now both ci4)

Late MC initials

- MC 精 [ts] vs. 見 [k]:
Cantonese : 精 [tsɛŋ˧], 見 [kin˧˥]
BJ opera: 精 [tsin˧˥], 見 [tɕjɛn˧˥] (「分尖團」 zīān tuán)
BJ Mand.: 精 [tɕi˥˩], 見 [tɕjɛn˧˥] (「不分尖團」)
- MC [k^h] 溪: Canto. [k^h] / [h]
e.g. MC k^həj Canto. 慷^{C1} [k^hɔj˧˥] 開^{A1} [hɔj˧˥]
(with MC -w-) [k^{wh}] / [f]
e.g. MC k^hwəj Canto. 盍^{A1} [k^{wh}ɐj˧˥] 魁^{A1} [fuj˧˥] 傀^{B1} [faj˧˥]
- MC [ŋ] 疑 has remained [ŋ] in front non-high vowels
e.g. 牙 [ŋa˧˥] 牛 [ŋew˧˥] 我 [ŋɔ˧˥] 傲 [ŋou˧˥]

(etc.)

MC initials

MC voiced plosives / affricates: 並 b 定 d 澄 ɖ 從 dz 牀 ɖʑ 羣 g
e.g. MC [b] > [b]? [p]? [pʰ]? (other thing?)

MC >	牌 bɛ ^A ‘placard’	被 bje ^B ‘quilt’	病 bjæŋ ^C ‘sickness’	白 bæk ^D ‘white’
全州湘	bāj ^{A2}	bi ¹ ^C	biŋ ¹ ^C	bə ¹ ^{A2}
南寧平	paj ¹ ^{A2}	pəj ¹ ^{B2}	pən ¹ ^{C2}	pək ¹ ^{D2b}
柳州官	p ^h aj ¹ ^{A2}	pi ¹ ^C	pjen ¹ ^C	pə ¹ ^{A2}
廣州粵	p ^h aj ¹ ^{A2}	p ^h ej ¹ ^{B2}	pɛŋ ¹ ^{C2}	pak ¹ ^{D2}
香港客	p ^h aj ¹ ^{A2}	p ^h i ¹ ^{A1}	p ^h jaŋ ¹ ^C	p ^h ak ¹ ^{D2}

MC initials

Cantonese: AB aspirated, CD unaspirated

Mandarin: 「平送 仄不送」 i.e. A aspirated, BCD unaspirated
not entirely correct:

Standard Mand.	* 平 *A	* 上 *B	* 去 *C	* 入 *D
陰 *voiceless 清 initial	1 ʅ	3 ʅ	4 ʅ	<i>random</i>
陽 *voiced sonorant 次濁 initial	2 ʅ			4 ʅ
陽 *voiced obstruent 全濁 initial				2 ʅ

Few cases of *voiced obstruent initial still in tone B:

艇 ^th_iŋʅ, 勉強 mjɛn/ tɕ^hjaŋʅ, 窈窕 jao/ t^hjaoʅ,

否極泰來 p^hiʅ tɕiʅ t^hajʅ lajʅ

i.e. same as Cantonese: AB aspirated, CD unaspirated

MC initials

The trait of AB aspirated, CD unaspirated:

Cantonese got this pattern from Early Mandarin

13th century wars between Song China and the Mongols: Huge waves of refugees from/via Jiangxi arrived in the Pearl River Delta, bringing with them (native or non-native) Mandarin. Canto. and many Yue dialects inherited this immigrant pattern.

The older pattern in Yue/Pinghua is ABCD unaspirated

