

Yuè & the Sinitic Languages of the Pearl River Region — Unit 2.1

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

Tones

Tone values 調值 (phonetics) vs.
Tone categories 調類 (phonology)

Tone values 調值: phonetic shape of tones
People commonly use Chao's tone letters or tone numerals
(Chao Yuen Ren 趙元任)

e.g. Standard Mandarin

官話 /kwan¹ xwa⁴/ ~ /kwan⁵⁵ xwa⁵¹/

你好 /ni²¹³ xao²¹³/ ~ /ni^{213>25} xao²¹³/

Tone categories 調類:

tones in phonological system (tonemes, allotones)

(names like) tones A B C D, tones 1 2 3 4, etc.

e.g. Mandarin 旗 /tɕʰi2/ [tɕʰi²⁵], Cantonese 恥 /tsʰi2/ [tsʰi²⁵],
Taiwanese 持 /tʃʰi5/ [tʃʰi²⁵]

(with numerals, be careful whether people mean tone categories or tone values)

Tones

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

Table 1: Tones in Standard Cantonese

Tones

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

cong⁻¹ syu⁻² zeoi⁻³ m⁻⁴ laan⁻⁵ do⁻⁶
倉 鼠 最 唔 懶 惰

canto.hk

倉鼠最唔懶惰 cong1 syu2 zeoi3 m4 laan5 do6
三碗半牛腩麵 saam1 wun2 bun3 ngau4 naam5 min6

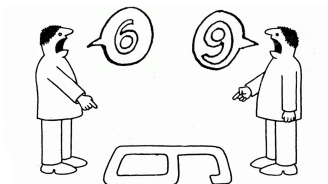
Tones

	*A	*B	*C	*D	
				L	S
陰	1 ˩(˨˩)	2 ˩~˩	3 ˩	3 ⁽⁸⁾ ˩	1 ⁽⁷⁾ ˩
陽	4 ˩~˩	5 ˩	6 ˩	6 ⁽⁹⁾ ˩	

6 tones? 9 tones?:

6 tonemes

「九聲六調」



Chinese linguistics traditionally counts tones in sonorant-ending syllables and obstruent-ending syllables (-p -t -k) separately

Tones

	*A	*B	*C	*D	
				L	S
陰	1 ˩(˨˩)	2 ˩˨˩	3 ˩	3 ⁽⁸⁾ ˩	1 ⁽⁷⁾ ˩
陽	4 ˩˨˩	5 ˩	6 ˩	6 ⁽⁹⁾ ˩	

詩 si1 'poem' 唧 zit1⁽⁷⁾ 'tickle'
 史 si2 'history'
 試 si3 'to try' 節 zit3⁽⁸⁾ 'festival'
 時 si4 'time'
 市 si5 'city'
 事 si6 'matter' 截 zit6⁽⁹⁾ 'to sever'

Tones

	*A	*B	*C	*D	
				L	S
陰	1 𠂔(～𠂔)	2 𠂔～𠂔	3 𠂔	3 ⁽⁸⁾ 𠂔	1 ⁽⁷⁾ 𠂔
陽	4 𠂔～𠂔	5 𠂔	6 𠂔	6 ⁽⁹⁾ 𠂔	

<i>sonorant-ending</i>						<i>obstruent-ending</i>		
cong1	syu2	zeoi3	m4	laan5	do6	bit1 ⁽⁷⁾	faat3 ⁽⁸⁾	lik6 ⁽⁹⁾
倉	鼠	最	唔	懶	惰	必	發	力
saam1	wun2	bun3	ngau4	naam5	min6	jat1 ⁽⁷⁾	baak3 ⁽⁸⁾	dip6 ⁽⁹⁾
三	碗	半	牛	腩	麵	一	百	碟

See more examples at @ 語言學貓

https://youtu.be/0ebW9QF1K0s?si=cqwoMFQa_BdgQG1A&t=73

(From now on: tones 7 8 9 will usually be notated as tones 1 3 6, following Jyutping)

Tones

	*A	*B	*C	*D	
				L	S
陰	1 ˩(˩)	2 ˩~˩	3 ˩	3 ⁽⁸⁾ ˩	1 ⁽⁷⁾ ˩
陽	4 ˩~˩	5 ˩	6 ˩	6 ⁽⁹⁾ ˩	

Obstruent-ending syllables, with suprafix tone 2:

碟 dip6*2 [ti:p˩] ‘plate’

賊 caak6*2 [tsʰa:k˩] ‘thief’

Obstruent-ending syllables, tone 4 in ideophones:

嚙嚙聲 syut4 syut2 seng1 sound of slurping e.g. noodle

軟揸揸 jyun5 det4 det4 weak and flimpsy 廣州話方言詞典

濕笛笛 sap1 det4 det4 ~ zet4 zet4 ~ det6 det6

fully soaked and dripping

Suprafixes

Suprafix: suprasegmental affix (affix in the form of e.g. tone, stress)
In the case of Cantonese: tonal affixes

suprafix $\neq$ **tone-sandhi** although both are called 變調 in Chinese

Tone-sandhi:

(largely) automatic & predictable process of tones changing,
given a particular phonological / morphosyntactic environ.

e.g. Standard Mandarin tone 3 sandhi rule

買馬 mǎi mǎ > mǎi mǎ (= 埋馬 mǎi mǎ)

e.g. Shanghainese

三杯 se˩ pe˩ > se˩ pe˩

四杯 sz˩ pe˩ > sz˩ pe˩

Suprafixes

Tone Sandhi in Cantonese?

e.g. Cheung 張 (1969)

different sandhi rules concerning ˩ > ˨˩ in Canton vs Hong Kong

However, it is very hard to find people nowadays who contrast ˩ and ˨˩ consistently

Hong Kong is nowadays basically all ˨˩

Most Cantonese speakers do not have tone sandhi

Cheung, Yat-Shing 張日昇. 1969. 香港粵語陰平調及變調問題 A study on the upper even tone and changed tones in Cantonese as spoken in Hong Kong. *The Journal of The Institute of Chinese Studies of The Chinese University of Hong Kong* 香港中文大學中國文化研究所學報 2: 81-107.

<https://www.cuhk.edu.hk/ics/journal/articles/v02p081.pdf>.

Suprafixes

(Suprafixes in Standard Cantonese are tone-substitutive)

Grammatical vs. lexical suprafixes:

Grammatical suprafix:

e.g. verbal:

食 sik6 ‘eat’ > sik6*2 (= sik6 zo2 食咗 eat PFV)

nei5 sik6*2 faan6 mei6 aa3?

2SG eat:PFV rice not.yet Q

‘Have you eaten yet?’

𢦏 tam3 ‘to coax’, ‘to cheer s.o. up’

tam3*2~tam3 (= tam3 jat1 tam3 𢦏一𢦏 coax:DELIM)

tam3*2~tam3 keoi5

coax:DELIM 3SG

‘Coax him/her a little bit.’

Suprafixes

Grammatical Suprafix:

e.g. adjectival:

滑 waat6 ‘smooth’ > waat6*2~waat6

(= waat6 jat1 waat6 滑一滑 ‘extraordinarily smooth’)

sik6 dou3 pei4fu1 waat6*2~waat6

eat till skin smooth:extraordinary

‘Eat this and the skin becomes smooth as hell.’

<https://hkpltravel.com/23966/foods/> 【滑溜溜豆腐花 ！】 /

紅 hung4 ‘red’ > hung4~hung4*2-dei2 ‘somewhat red’

Suprafixes

Lexical (Nominal) Suprafix: see also de Sousa (2024)

a.k.a. 小稱變調 siu2 cing1 bin3 diu6 ‘diminutive tone-change’

However: Cantonese lexical suprafixes have been used so much that they often do not have a diminutive meaning (*small / endearment meaning*).

e.g.

頭 tau4 ‘head’

糖 tong4 ‘sugar’

阿頭 aa3 tau4*2 ‘boss, chief’

糖 tong4*2 ‘lolly, candy’

comparable with Northern Mandarin -兒 -r, e.g.

頭 tóu ‘head’

頭兒 tóur ‘boss, chief’

de Sousa, Hilário. 2024. Some observations on the Cantonese lexical suprafixes. *Languages* 9 (10). 311. (Special issue “Typology of Chinese Languages: One Name, Many Languages”, edited by Umberto Ansaldi & Pui Yiu Szeto.)
<https://doi.org/10.3390/languages9100311>

Suprafixes

Cantonese lexical suprafix:

Derivative morphology (word building)

Usually applied to last syllable of a noun / noun stem, e.g.

魚 jyu4*2 ‘fish’, 蝴蝶 wu4 dip6*2 ‘butterfly’

(cf. 魚生 jyu4 saang1 ‘raw fish’, 蝴蝶灣 wu4 dip6 waan1 ‘butterfly bay’)

Usually do not apply to other word classes, e.g.

鉗 kim4*2 ‘pliers’, 鉗 kim4 ‘to grab with pliers’

碟 dip6*2 ‘plate’, 碟 dip6 CLF

Not quite predictable, e.g.

法蘭西 faat3 laan4*1 sai1 ‘France’, 荷蘭 ho4 laan4*1 ‘Holland’,

but 芬蘭 fan1 laan4 ‘Finland’, 紐西蘭 nau2 sai1 laan4 ‘New Zealand’

澳門 ou3 mun4*2 ‘Macau’,

but 屯門 tyun4 mun4 ‘Tuen Mun’, 江門 gong1 mun4 ‘Kongmoon /

Jiangmen’, 廈門 haa6 mun4 ‘Amoy / Xiamen’

Suprafixes

	*A	*B	*C	*D	
				L	S
陰	1 ˩(˨˩)	2 ˨˩˨˩	3 ˨˩	3 ˨˩	1 ˩
陽	4 ˩˨˩	5 ˩	6 ˨˩	6 ˨˩	

Lexical suprafixes in Standard Cantonese:

- tone 2 [˨˩] (most common)
- tone 1 [˩] (rarer)
- tone 4 [˩˨˩] (only in some informal personal names, e.g. 達哥 daat6-go1*4 ‘Brother Tat’; or place names, e.g. 長沙灣 coeng4 saa1 waan1*4 ‘Cheung Sha Wan’)

Suprafixes

	*A	*B	*C	*D	
				L	S
陰	1 ㄋ(ㄣ)	2 ㄥ~ㄥ	3 ㄥ	3 ㄥ	1 ㄋ
陽	4 ㄣ~ㄣ	5 ㄥ	6 ㄥ	6 ㄥ	

Suprafix tone 2 [ㄥ] can be applied to
morphemes with citation tones 3 4 5 6

片 pin3*2 'film', 房 fong4*2 'room', 女 neoi5*2 'daughter', 妹 mui6*2 'younger sister',
鴨 aap3*2 'duck', 碟 dip6*2 'plate'

Suprafix tone 1 [ㄋ] can be applied to
morphemes with citation tones 2 3 4 5 6

自己 zi6 gei2(*1) 'self', -派 -paai3(*1) 'faction', 欄 laan4*1 'animal pen', -尾 -mei5*1
'last', -妹 -mui6*1 'girl', 掠水 loek6*1 sei2 'rob money'

Suprafixes

People usually know the citation tone, as the citation tone morpheme appears in other environments, e.g.:

蛋 daan6*2 ‘egg’

雞蛋 gai1 daan6*2 ‘chicken egg’

雞蛋仔 [gai1 daan6*2] -zai2 ([fowl egg] DIM)

雞蛋仔 gai1 daan6 zai2 ‘egg waffle’

https://en.wikipedia.org/wiki/Egg_waffle

蛋撻 daan6 taat1 ‘egg tart’

https://en.wikipedia.org/wiki/Egg_tart

蛋糕 daan6 gou1 ‘cake’

(蛋 daan6 is not used as an independent morpheme.)

Suprafixes

However, the citation tone is not always known by speakers.

(The citation tone version of the morpheme is no longer used)

e.g. daai6 mou6 saan1 highest peak in HK

is sometimes wrongly written as 大霧山 'big fog mountain'

Should be: 大帽山 'big hat mountain'; 帽 mou6*2 'hat'

(the citation tone version 帽 mou6 is no longer known)

橙 caang?*2 'orange'



Suprafixes

	*A	*B	*C	*D	
				L	S
陰	1 ㄣ(ㄣ)	2 ㄣ~ㄣ	3 ㄣ	3 ㄣ	1 ㄣ
陽	4 ㄣ~ㄣ	5 ㄣ	6 ㄣ	6 ㄣ	

Quiz for Cantonese speakers:

What is the citation tone of these syllables pronounced **jyun2**?

		Cantonese	Mandarin	Sino-Viet
garden	園	jyun?	yuán	viên / vườn
pellet	丸	jyun?	wán	hoàn
courtyard	院	jyun?	yuàn	viện
(surname)	阮	jyun?	ruǎn	nguyễn
garden	苑	jyun?	yuàn	uyễn / uồn

Suprafixes

On Tone 1 [˩] [˨˩] again:

Note: Modern speakers still have two words which must be [˨˩] (or maybe this is boundary intonation):

1.

先 sin1 [˨˩], sentence final particle expressing reservations in a question, vs.

先 sin1 [˩]^{usually}, adverb “first, before, earlier”

係	唔	係	你	去	先	先	?
hai6	m4	hai6	nei5	heoi3	sin˩	sin˨˩	?
be	NEG	be	2SG	go	first	SFP	

‘Is it that you are going first...?’

Suprafixes

2.

㗎 ce1 [ʏ] interjection expressing dismissal, vs.

車 ce1 [ɿ]^{usually} ‘vehicle’

Otherwise, nowadays [ɿ] [ʏ] are free variations for most speakers.

e.g. 根 gan1 [ɿ]~[ʏ] ‘root’

Hong Kong: tone 1 most usually [ɿ]

Suprafixes

For people who still make a distinction between /ŋ/ and /ɿ/
(Cheung 1969):

Tone 1 [ŋ] vs. Tone *1 [ɿ]

Suprafix tone *1: primarily on nouns, can be applied to
morphemes in any tone ^{other than tone 7 [ɿ]}, e.g.:

忠 zung1 ‘loyal(ty)’	鐘 zung1*1 ‘clock’
好酸 hou2 syun1 ‘very sour’	好孫 hou2 syun1*1 ‘good grandchild’
四方 sei3 fong1 ‘four directions’	四方 sei3 fong1*1 ‘square’
門門 saan1 mun4 ‘close door’	門門 mun4 saan1*1 ‘latch’

Female given names can end in tone *1 (DIM) but not tone 1

Male name varies

Cheung, Yat-Shing 張日昇. 1969. 香港粵語陰平調及變調問題. 香港中文大學中國文化研究所學報 2: 81-107.

<https://www.cuhk.edu.hk/ics/journal/articles/v02p081.pdf>

Suprafixes

Tone 2:

Cheung 張 (1969: 96-97): only one tone 2 in Hong Kong

Gāo 高 (1980: 7): Tone *2 rises higher than Tone 2 in Canton

Chao (1947: 34-35): Tone 2 [1], Tone *2 [1]

Ball (1888: xxxi) “third rising tone” [Tone *2]: “banning lower and ascending higher than [the other two rising tones]”, something like [√]

訪 fong2 ‘to visit’

房 fong4*2 ‘room’

挾 joeng2 ‘to shake’

樣 joeng6*2 ‘pattern’

嚙屎 gau6 si2 ‘lump faeces’

舊時 gau6 si4*2 ‘the past’

Gāo, Huáníán 高華年. 1980. 廣州方言研究. Hong Kong: The Commercial Press 商務印書館.

Chao, Yuen Ren. 1947. *Cantonese Primer*. Cambridge and Massachusetts: Harvard University Press.

Ball, James Dyer. 1888. *Cantonese Made Easy: A Book of Simple Sentences in the Canton Dialect, with Free and Literal Translations, and Directions for the Rendering of English Grammatical Forms in Chinese*, 2nd ed. Hong Kong: China Mail Office.

Suprafixes

Canton & Hong Kong at the forefront of the development of Cantonese suprafixes.

Other Cantonese varieties, e.g.

In 北海 Bak1 Hoi2 / Běihǎi Canto., out of 2179 lexical items:

例子很少，在我们的词汇表里，我们一共记录了七个有变调的词。(Chen & Chen 2005: 26)

In 南寧 Naam4 Ning4 / Nánning Canto., the tone 2 suprafix:

使用远无广州、香港白话那么普遍，基本可以穷尽。

(Lin & Qin 2008: 90)

Chén, Xiǎojīn 陈晓锦, and Tāo Chén 陈滔. 2005. 广西北海市粤方言调查研究. Beijing: China Social Science Press 中国社会科学出版社.

Lín, Yì 林亦, and Fèngyú Qín 覃凤余. 2008. 广西南宁白话研究. Guilin: Guangxi Normal University Press 广西师范大学出版社.

Suprafixes

e.g.

	南寧 Canto.	Older 澳門 Canto.	香港 Canto.
白話	baak6 waa6	baak6 waa6*2	baak6 waa6*2
帽	mu6	mou6	mou6*2
枱	toi4	toi4(*2)	toi4*2
枱布	toi4 bu3	toi4 bou3	toi4*2 bou3
乞兒	hat1 ji4*1	haat1 ji4*1	hat1 ji4*1
乞兒仔	?	haat1 ji4 zai2	hat1 ji4*1 zai2
澳門	u3 mun4	ou3 mun4(*2)	ou3 mun4*2

In Macau, /ou3 mun4/ is on average

more formal than /ou3 mun4*2/

e.g. 澳門大學 'Universidade de Macau' more likely pronounced /ou3 mun4 daai6 hok6/ in formal speech.

(/ou3 mun4*2 daai6 hok6/ is also OK)

Suprafixes

What is the origin of the lexical suprafixes in Cantonese?

We don't actually know for sure.

Ball (1888) describes the suprafix rising tone as a long tone.

Presumably there was a suffix originally.

cf. grammatical suprafixes

*sik6*2 (eat:PFV) < 食咗 sik6 zo2 (or 食嘅 sik6 hiu1)*

Lexical suprafix: ?

Best theory so far: < DIM suffix -兒 -ji4*1

Suprafixes

兒 MC ne^A ‘male child’ later ‘child’

南寧位子潯平話 (Pinghua, south central Guangxi)

雞兒 $kɛj\downarrow ni\downarrow$ (fowl-DIM) 刀兒 $taw\downarrow ni\downarrow$ (knife-DIM)

車兒 $tʃ^hɛ\downarrow ni\downarrow \sim \downarrow$ (vehicle-DIM)

鑊兒 $tʃ^hɛŋ\downarrow ni\downarrow$ (wok-DIM)

桂平白話 (Cantonese Yue, central eastern Guangxi)

only $-ŋi\downarrow$, e.g

雞兒 $kɛj\downarrow ŋi\downarrow$

信宜 (Yue, southwestern Guangdong)

豬 $tʃy\downarrow >$ 豬兒 $tʃyn \nearrow$ (pig:DIM)

豬兒 $tʃy\downarrow ni\downarrow$ ‘piglet’ $>$ 豬兒兒 $tʃy\downarrow nin \nearrow$ (pig-DIM:DIM)

Further east: -兒 suffix no longer productive

Suprafixes

Standard Cantonese, fossilised cases of -兒 ji4*1

乞兒 hat1 ji4*1 'beggar'

捉兒人 zuk1 ji4*1 jan4*1 'hide and seek'

貓兒 maau1 ji4*1 'kitten' (moribund)

老	狗	嫩	貓	兒
lou5	gau2	nyun6	maau1	ji2*1
食	死	冇	人	知
sik6	sei2	mou5	jan4	zi1

so... perhaps mui6*2 & -mui6*1 < 妹兒 mui6 ji4*1 (?)

(The 兒 theory is the best theory so far; see also de Sousa (2024) for criticisms of this theory)

Suprafixes

仔 zai2 (DIM): another possible origin of the lexical suprafixes.
However:

- 仔 zai2 was in ordinary tone 2 and not suprafix tone *2
- 仔 zai2 explains tone *2 [4] reasonably, but tone *1 [1] not well
e.g. mui6 + zai2 > mui6*2 is OK, but mui6*1 (?)
- lexical suprafixes and -仔 -zai2 already coexisted in the 19th century Cantonese textbooks, e.g. 檯仔 toi4*2 zai2 (table DIM)

However, 仔 zai2 could have inspired further suprafixed forms, e.g. 麻疹 maa4 zai2 > maa4*2 ‘measles’ (Kwok 2016: 307)