Neutralized diphthongization inventory in Chongqing Mandarin

Hailang Jiang, University College London

Summary. This study models the process of diphthongization of 5 high vowels in Chongqing Mandarin (CM) using dispersion theory under a scalar vowel system. The main takeaways are that diphthongization takes place only when phonotactics requires a marked structure *[+hi][+hi] to be repaired, and the diphthongs neutralize because lowered [+hi] vowels have a more narrow scale for F2 accommodate contrastive qualities of front/back vowels, respectively.

Data. All discriptive data are collected and transcribed by myself who is a native speaker of CM. In CM, in utterance-final positions, 5 high vowels diphthongize with 3 qualities, i.e., [i, y, u, ɯ, ɹ̩[1]] → [iɪ, yɪ, uʊ, ɯʊ, ɹ̩ɛ] / __#. The examples in (1) illustrate the alternations of the five vowels (in red).

(1)	Utterance-medial high vowel		Utterance-final high vowel diphthongized	
	[pi^{33}.xoː^{31}]	‘pen case’	[mau^{33}.piɪ^{31}]	‘writing brush’
	[ly^{34}.zən^{31}]	‘female’	[mei^{34}.lyɪ^{42}]	‘beauty’
	[t͡sɹ̩^{35}. t͡ɕin^{33}]	‘funds’	[tʰə̣u^{33}.t͡sɹ̩ɛ^{35}]	‘to invest’
	[fɯ^{22}.muʊ^{42}]	‘parents’	[i̯ɑŋ^{34}.fɯʊ^{21}]	‘foster father’
	[su^{35}.t͡sə̣u^{33}]	‘Suzhou’ (a city)	[t͡ɕi̯ɑŋ^{35}.suʊ^{33}]	‘Jiangsu’ (a province)

Non-high monophthongs ([peː^{31}] ‘white’, [kʰoː^{213}] ‘class’, [paː^{31}] ‘eight’), a vowel in a closed syllable ([ɕin^{35}] ‘heart’) and the second part of an underlying diphthong ([pai^{213}] ‘failure’), do not diphthongize utterance-finally.

Motivation of diphthongization. I analyzes diphthongization as fission. Note that non-high vowels are lengthened utterance-finally (e.g., /pe^{31}/[peː^{31}] ‘white’), I argue that high vowels are lengthened as well, making all utterance-final syllables bimoraic. I do not distinguish between a long vowel [iː] and vowel geminate [ii] here because such a structure does not occur elsewhere in CM. However, utterance-final plain long high vowels like [ii]# are not tolerated here due to a general OCP constraint banning *[+high][+high] sequences in the domain of a syllable (following Gong & Zhang 2021), making ‘long’ or geminate vowels like *[ii] illicit. Evidence is that CM disallows rhymes like *ju, *ɥi, *wi, etc., e.g., ‘New York’ is borrowed as [ljə̣u.jo], not *[lju.jo]. Thus, the qualities of the ‘offglides’ of the diphthongized vowels are lowered to a minimal degree. That means, they resemble the corresponding original part (e.g., [i] vs [ɪ]), but surprisingly they end up with 3 qualities only, with some neutralized.

Failure of a featural vowel system. The difficulties to model this under the featural vowel system are: **(1)** No more room for additional distinguished height level. The three monophthongs [i, e, a] already occupy all possible levels under a system with two binary features of height, [high] and [low]; adding [ɪ] into it has to be companied by additional features like [tense], but this is irrelevant to the crucial contraint banning [+high] / [+high] __ at effect here. **(2)** No phonetic motivation for neutralization. To account for it, we can only specify [ɪ] by deleting a feature of [round] from [i, y] after being lowered, but it is not clear why we do not keep two unique offglides [ɪ] and [ʏ].

See Table 1 for a tentative messy featural vowel system for all front vowels, where there are 3 levels of height but with [tense], and [tense] and [round] alternate between an unspecified value [0] and a specified value [+]/[–], with no articulatory and perceptual ground at all. Thus a finer-grained scalar system, which accommodate phonetic mechanism is required.

[1] This [ɹ̩] is called an ‘apical vowel’. It is an allophone of /i/ which occurs after onsets {t͡s, t͡sʰ, s, z}, as a nucleus of the syllable. Given its allophonic status and the fact that it behaves together with the other four dorsal high vowels, it should be considered as a high vowel phonologically.

i [–rd][0tense], y [+rd][0tense]	[+hi][–lo]
ɪ [0rd][–tense]	[–hi][–lo]
e [+tense]	
a	[–hi][+lo]

Table 1. A featural vowel system.

6		5		4	3	2		1	F2/F1
i		y		…	…	ɯ		u	1
ɪ̟	ɪ	ʏ̟	ʏ	…	…	ᴜ[2] ᴜ̠	ʊ	ʊ̠	2
/	/	e		ə	ə̦	o	/	/	3

Table 2. A scalar vowel system.

Scalar modeling with Dispersion Theory. A scalar system is demonstrated in Table 2. The apical vowel [ɪ̟] is ignored for time being. There are at least 6 units in the dimension of F2 from 1 to 6. For dimension of F1, only the smallest 3 units suffice for current purposes. The essential phonetic property of this scalar system is that, for F1 = 2, the two peripheral units (2, 1) and (2, 6) are disallowed due to the articulatory property of the oral cavity, i.e., the articulatory space is shaped like a triangle (e.g., Ladefoged 2006, Odijk & Gillis 2022, among many others), indicated by shaded area in the right table. It can be formalized into an *Effort constraint using Flemming's (2004) terms. The consequence of the narrowing of F2 scale from F1 = 1 to F1 = 2 is that the minimal vertical lowering of [i] has to be accompanied by a horizonal shift inward, to satisfy *Effort, resulting in [ɪ] (2, 5.5). However, the minimal vertical lowering of [y] does not require an inward shift, resulting in [ʏ̟] (2,5); otherwise, a shifted [ʏ] (2, 4.5) violates a faithfulness constraint on horizontal position. The consequence of these two lowering is that [ɪ] and [ʏ̟] are too close in the dimension of F2 and cannot be perceptually contrastive, violating Flemming's (2004) MinDist (F2) = 1. Therefore, only one of [ɪ] and [ʏ̟] can be retained. Similarly, we have to drop of of the back vowels [ᴜ̠] and [ʊ]. Following Flemming (2004), given the number of contrast pair (1 here), contrast has to be as salient as possible, so [ɪ] and [ʊ] at both endpoints of the F2 scale (given F1 = 2) are reserved. Thus, the neutralization process is explained with widely established phonetic motivation. A tableau is given to demonstrate the rankings of relevant constraints.

A tableau for predicting vowel inventory at F1 = 2, as a result of high vowels' partial lowering

6-5-2-1	i-y-ɯ-u[3]	*Ef	M=1	M=2	M=3	M=4	M=5	MC	HF
5.5-1.5	☞ ɪ-ʊ						*	√	****
5.5-5-2-1.5	ɪ-ʏ̟-ᴜ̠-ʊ		*!*	**	**	*****	******	√√√√√√	**
5.5-4.5-2.5-1.5	ɪ-ʏ-ᴜ-ʊ			*!*	**	*****	******	√√√√√√	****
6-5-2-1	ɪ̟-ɪ-ʊ-ʊ̠	*!*		**	**	***	*****	√√√√√√	
5-2	ʏ̟-ᴜ̠					*!	*	√	**

*Ef = *Effort, M = MinDist(F2), MC = MaximizeContrast (note this is a positive constraint; the more √, the more harmonic the candidate is), HF = HorizontalFaithfulness.

The mystery of the apical vowel. [ɪ̟] → [ɪ̟ɛ] / __#. This is a very tricky case, as an apical vowel fissions with an additional dorsal vowel. Conventional acoustic formant analysis for dorsal vowels does not apply. Here is a tentative analysis: [ɪ̟] is phonologically high given its allophonic status of /i/, so it lengthens and then diphthongizes; [ɪ̟] forms a constriction near the alveolar ridge at the fronter area of the oral cavity (Lee-Kim 2014), so articulatorily the closest to it is a front vowel, probably [ɛ], but further measurement is required. I leave this issue open to future research. At least the analysis in the previous section at the level of F1 = 2 is not going to be threatened by any analysis of [ɪ̟ɛ] because the height level of [ɛ] (F1 = 4) is even lower than that of [e] (F1 =3).

[2] For the sake of brevity, I use the ad hoc notation [ᴜ] = [ɰ̞] in stardard IPA here, which can be considered as an unrounded version of [ʊ]. The two form a pair, as [ɪ] and [ʏ].

[3] It actually means the correspondents of [i-y-ɯ-u] when they are (partially) lowered to the level of F1 = 2.