Research Summary: A Dispersion-Theoretic account of neutralized diphthongization inventory in Chongqing Mandarin

Hailang Jiang

Summary. In Chongqing Mandarin (CM), there is an alternation that high vowels undergo utterance-finally where they are diphthongized. This study models the process of diphthongization of 5 high vowels as lengthening (also as fission) with partial lowering, which is motivated by the interaction of two independent general phonotactic constraints; it also explains the neutralization of the offglides of the diphthongized vowels using Dispersion Theory, which necessitates a scalar vowel system, and argues that lowered high vowels have a more narrow scale for F2 which does not accommodate contrastive qualities of front or back vowels, respectively.

1. Background. All descriptive data are collected and transcribed by me, a native speaker of CM. In CM, at the end of any utterances (incl. prosodic words in citation forms), 5 high vowels diphthongize, i.e., [i, y, u, ɯ, ɹ̩] → [iɪ, yɪ, uʊ, ɯʊ, ɹ̩ɛ] / __#. ([ɹ̩] is an 'apical vowel' and is an allophone of /i/ after alveolar fricatives or affricates, as a syllable nucleus.) 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'	[tʰə̹u^{33}.t͡sɹ̩ɛ^{35}]	'to invest'
	[fɯ^{22}.muʊ^{42}]	'parents'	[i̯ɑŋ^{34}.fɯʊ^{21}]	'foster father'
	[su^{35}.t͡sə̹u^{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 ([pin^{35}] 'ice') and the second part of an underlying diphthong ([pai^{213}] 'failure'), do not diphthongize utterance-finally.

Accounts are required of both: **(i.)** why high vowels in certain contexts diphthongize (while others do not) and how the quality of the offglides is determined; **(ii.)** why five high vowels have only three offglides with distinguished quality. Section 2 and 3-4 discuss them, respectively.

2. Motivation of diphthongization. I analyze diphthongization as lengthening. Note that non-high vowel monophthongs in (C)V syllables 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, following Féry (2003) using the constraint FINALLENGTH. I do not distinguish between a long vowel [iː] and vowel geminate [ii] here because such a long vowel or a vowel geminate does not occur elsewhere in CM. Thus, lengthening of [i_1] is also analyzed as fission with an offglide of identical quality [i_1i_1].

However, utterance-final high vowel geminates like [i_1i_1]# are not tolerated here due to a general OCP constraint *HH banning high vowel sequences in the domain of a syllable (see also Gong & Zhang 2021), making geminate vowels like *[i_1i_1] illicit. The evidence is that CM disallows rhymes like *ju, *ɥi, *wi, etc., e.g., 'New York' is borrowed as [ljə̹u.jo], not *[lju.jo]. Thus, the qualities of the 'offglides' of the diphthongized vowels should be changed to avoid violations of *HH, but only to a minimal degree due to faithfulness constraints. That means, only a slight change in height is allowed to guarantee the maximal resemblance between the surface forms of offglides and their inputs, e.g., [iɪ]# *[ie]#. Through the interaction of two independent constraints in CM, FINALLENGTH and *HH, I explain why pi# but not pi…#, pe#, pin#, pai# diphthongize. The tableau below shows the optimization for the utterance-final input /mau^{31}.pi^{31}/ 'writing brush'.

A tableau for an utterance-final high vowel

Input: /mau.pi$_1$/#	*HH	FINALLENGTH	Faithfulness(height)
☞ mau.pi$_1$		*!	
mau.pi$_1$i$_1$	*!		
mau.pi$_1$ɪ$_1$			*
mau.pi$_1$e$_1$			**!

We would expect five high vowels to have five offglides, but surprisingly, the offglides end up with 3 qualities only [ɪ, ʊ, ɛ], with some neutralized ([iɪ] and [yɪ], [uʊ] and [ɯʊ]).

3. Failure of a featural vowel system. I first tried to model this under a featural vowel system by integrating offglides into the system and derive offglides by removing specifications of certain features from high vowels. The attempt I made on front vowels is showed in Table 1. The problems are **(i.)** No more extra room for an 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] cannot make use of current features and has to be accompanied by additional features like [tense], but this is irrelevant to the crucial constraint at effect here banning a high vowel after another high vowel. **(ii.)** No phonetic motivation for neutralization, as the featural system is too coarse to capture it. 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] for [i] and [y], respectively.

To conclude, in Table 1, there are 3 levels of height with an additional feature [tense], with [tense] and [round] alternating between an unspecified value [0] and a specified value [+]/[–], without any articulatory or perceptual grounding. A phonetically grounded account is preferable.

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	ᴜ̠	ʊ	ʊ̠	2
/	/	e		ə	ə̢		o	/	/	3

Table 2. A scalar vowel system. ([ᴜ] = [ʊ̢])

4. Scalar modeling with Dispersion Theory. A scalar system is demonstrated in Table 2. The apical vowel [ɪ̟] is discussed in the next section. There are at least 6 units in the dimension of F2 from 1 to 6. For the dimension of F1, 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 shape of the formant space, i.e., the formant space is shaped like a triangle (e.g., Ladefoged 2006, Odijk & Gillis 2022, among many others), indicated by shaded areas in Table 2. It can be formalized into an *EFFORT constraint following Flemming (2004). The consequence of the narrowing of the 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 [i] and [y]'s lowering is that [ɪ] and [ʏ̟] are too close in the dimension of F2 and are not perceptually contrastive, violating Flemming's (2004) constraint MINDIST(F2) = 1. Therefore, only one of [ɪ] and [ʏ̟] can be retained. Similarly, the back vowels [ᴜ̠] and [ʊ] have to merge. Following Flemming (2004), given the number of contrast pairs (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 retained. Thus, the neutralization process is explained using a phonetically motivated explanation. 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	*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.

5. Open questions. 5.1. The mystery of the apical vowel. [ɹ̩] → [ɹ̩ɛ] / __#. 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: [ɹ̩] is phonologically high given its allophonic status of /i/, so it lengthens and then diphthongizes; [ɹ̩] 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 [ɹ̩ɛ] because the height level of [ɛ] (F1 = 4) is even lower than that of [e] (F1 = 3).

5.2. Phonetic evidence of the offglide merger. This study argues for the merger of the offglides based on preliminary acoustic analysis with data from several speakers as shown in Figure 1 and an impressionist consultation with some CM native speakers. The merger *may* be because the differences which do exist are not audible and are not discernible in the acoustic formant space. A more rigorous analysis may benefit from more acoustic data from more native speakers. A tentative analysis could be made on the F1 and F2 of the offglide [ɪ] in both [iɪ] and both [yɪ], to see if they have significantly different mean value when we allow random effects for individuals and items.

5.3 Why not lower the nuclei? To satisfy both *HH and FinalLength, a candidate can lower the first part of the geminate such as [i₁i₁] rather than the second part, e.g., [ɪ₁i₁] instead of [i₁ɪ₁]. I resort to prominence alignment (Crosswhite 2004), which argues that the syllabic prominence (peak > margin) has to align with the segmental prominence (sonority). Currently I assume that the offglides [ɪ, ʊ, ɛ] are less salient and less sonorous than high vowels because they are all centralized towards [ə] (as shown in Figure 1), but this analysis requires more phonetic evidence.

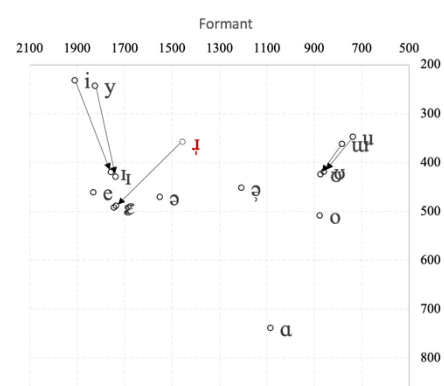

Figure 1 Formant space for vowels