



Fu, Boer. 2022. Contrast preservation in Mandarin r-suffixation: A comparative study of Beijing and Liaoning dialects. *Glossa: a journal of general linguistics* 7(1). pp. 1–42. DOI: <https://doi.org/10.16995/glossa.5801>



Open Library of Humanities

## Contrast preservation in Mandarin r-suffixation: A comparative study of Beijing and Liaoning dialects

Boer Fu, Massachusetts Institute of Technology, US, [boerfu@mit.edu](mailto:boerfu@mit.edu)

Two dialects of Mandarin Chinese, Beijing and Liaoning, display different patterns in r-suffixation, a morphophonological process. This paper investigates the low vowel rhymes: /an/, /a/, /aŋ/ in the two dialects. In Beijing, /an/ and /a/ undergo neutralization with the addition of the diminutive /-r/ suffix (Zhang 2000). But in Liaoning, the two rhymes maintain their contrast in the suffixed forms by each employing a different /r/ allophone for the suffix. /an/ selects for the retroflex rhotic [ɻ], while /a/ opts for the non-retroflex, or bunched, [ɹ] in Liaoning. I present a contrast preservation analysis to account for the difference between the two dialects. The analysis is set in the framework of Flemming's (1995) Dispersion Theory. It employs MinDist constraints and PreserveContrasts-BD, which is a modified version of Flemming's MaximizeContrasts, adapted to refer to output-output mapping. I argue that the overarching objective for both dialects is to maintain salient contrasts between the suffixed forms, but they differ in whether they prioritize preserving the three-way contrasts of low vowel rhymes or maintaining enough auditory distance between the suffixed rhymes. Liaoning prioritizes preserving the three-way contrasts, ranking PreserveContrasts-BD alongside MinDist, whereas Beijing is less preoccupied with keeping three suffixed forms, ranking PreserveContrasts-BD below MinDist. This is why the contrast between /an/ and /a/ is preserved in Liaoning via /r/ allophony, but neutralized in Beijing. Acoustic evidence for the two allophones of the r-suffix is provided. Segmental processes involving the low vowel and the nasal coda are also discussed.

*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by the Open Library of Humanities. © 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

OPEN ACCESS



## 1. Introduction

Several dialects of Mandarin Chinese employ an /-r/ suffix to derive diminutive nouns. This process is called *r-suffixation*, *erhua*, or *er-suffixation*. The phenomenon has attracted much attention from phonologists over the years (see Wang & He 1985; Lin 1993; Zhang 2000; Lee 2005; Cheng 2014; Hsieh et al. 2019; Xu 2020), since it is the only morphophonological process in Mandarin that involves segmental change. Such morphophonological alternations are commonplace in world languages, providing a key source of evidence for underlying representations. However, in Mandarin, morphophonological alternation is rare. It is found only in tone sandhi, a suprasegmental phenomenon, and r-suffixation.

Previous research on the diminutive form has mainly focused on the Beijing dialect, on which Standard Chinese, or *Putonghua*, is based. It has been reported that after r-suffixation, the rhymes /an/ and /a/ neutralize to [aŋ] (Wang & He 1985; Zhang 2000). This is indeed true for the Beijing dialect. However, the two rhymes do not neutralize in a dialect of Mandarin spoken in the province of Liaoning, northeast of Beijing. Liaoning forms one branch of Northeastern Mandarin, or *Dongbeihua*. In Liaoning, /an/ becomes [äŋ] while /a/ becomes [är] after suffixation, as seen in (1). Here, [ä] denotes a centralized low vowel. The two diminutive forms differ in their choice of allophone for the /-r/ suffix. [ŋ], with a curling tail, is a retroflex approximant, whereas [r] is a non-retroflex approximant.

(1)	Rhyme	Beijing Diminutive	Liaoning Diminutive
	(a) /an/	(c) äŋ	(e) äŋ
	(b) /a/	(d) är	(f) är

There is a wealth of studies on the diminutive forms of non-standard dialects of Mandarin, although the primary focus is on patterns that are drastically different from Beijing, like Yanggu l-infixation and Jiyuan rime change (see Lin 1993). A dialect like the northeastern Liaoning is considered to be near identical to Beijing, and thus given less attention in the literature. Notable exceptions include a descriptive and historical study of Northeastern Mandarin by Tan (2018) written in French, and *electromagnetic articulography* (EMA) studies conducted by Hsieh, Jiang, and Chang (see Hsieh et al. 2019; Jiang et al. 2019). Examining the articulatory gestures of Liaoning speakers, they found two distinct /r/ allophones. One is a retroflex [ŋ] and the other is a bunched [r]. To my knowledge, there has been no *Optimality Theory* (OT) analysis of any Northeastern Mandarin dialects in the literature, even though Beijing has been accounted using OT many times before (see Da 1996; Zhang 2000; Tian 2009). This paper will serve as a first attempt at an OT analysis of the dialect.

Another shortcoming of the previous literature is that open-syllable monophthong stems are treated as a separate phenomenon from nasal stems (see Lee 2005; Liu 2017). The open-syllable /a/ rhyme is usually compared to other open-syllable monophthong stems to see how their

formant values are affected by r-suffixation. Researchers group the nasal rhymes /an/ and /aŋ/ together to investigate nasal stop deletion and vowel denasalization.

In this paper, I compare three contrasting rhymes involving the low vowel: /an/, /a/, and /aŋ/. The mid vowel rhymes /ən/, /ə/, and /əŋ/ undergo the same phonological processes as their low vowel counterparts in Beijing and Liaoning. I only include an analysis for the low vowel rhymes in this paper, but it should be understood that the same analysis can account for the mid vowel rhymes as well. The high vowel and diphthong rhymes, on the other hand, are beyond the scope of this study, but a brief discussion is provided in Section 5.4.

I include both Beijing and Liaoning in my comparative study. The acoustic features of the low stem vowel /a/ in the suffixed forms, as well as the rhotic suffix itself, are examined. It is shown that in Liaoning, /an/ and /a/ do not neutralize. I argue that the avoidance of neutralization in Liaoning is a direct result of contrast preservation: the dialect resorts to /r/ allophony in order to preserve the contrast from the stem forms in the suffixed forms. Acoustic cues such as vowel backness and vowel nasalization also play a role in aiding contrast preservation.

My analysis is set in the framework of Flemming's (1995) *Dispersion Theory*. The original framework is modified to accommodate multiple acoustic cues, which are modeled as individual dimensions in a multi-dimensional rhyme space. Each surface rhyme has a coordinate that signifies its feature value corresponding to each acoustic cue. A *Rhyme Distance* is calculated between each pair of rhymes using the Euclidean distance formula.

I argue that Dispersion Theory can account for the divergent behavior of the two dialects in r-suffixation. Beijing opts for neutralization, in order to avoid having too small a contrast between two suffixed rhymes, by virtue of the constraint *MinimalDistance*. Whereas Liaoning's strategy is to enhance the contrast, so that the other key constraint in Dispersion Theory, *MaximizeContrasts*, can be satisfied. Previous OT accounts of Beijing do not engage with the issue of neutralization, nor do they compare it with a dialect that averts neutralization like Liaoning.

The paper is organized as follows. Section 2 presents data from Beijing and Liaoning with acoustic and articulatory details. Section 3 provides a brief introduction to Flemming's (1995) Dispersion Theory. I present my contrast preservation analysis of r-suffixation in Section 4. Section 5 is a discussion of the role of *Base-Derivative* (BD) faithfulness constraints, several alternative analyses, diachronic sound change, high vowel and diphthong rhymes. Section 6 concludes.

## 2. R-suffixation in Beijing and Liaoning

The low vowel stem forms are the same in Beijing and Liaoning, as shown in (2). The vowel quality of /a/ is subject to *Rhyme Harmony* (Duanmu 2007), which dictates that the backness of the vowel must match the backness of the nasal coda. Specifically, an alveolar nasal follows a front vowel [a] (2a), and a velar nasal follows a back vowel [ɑ] (2c). A non-high vowel in an

open syllable is central. The low central vowel is transcribed here as [ä], as seen in (2b), with the umlaut diacritic marking centralization. Both nasal codas nasalize the preceding vowel.

(2)	<b>Beijing/Liaoning</b>	<b>English</b>
	(a) pān	‘half’
	(b) pä	‘handle’
	(c) pāŋ	‘club’

When the diminutive suffix /-r/ is added, several changes take place. The suffixed forms in Beijing and Liaoning are listed side by side in (3). Also included are the stem forms and their assumed *Underlying Representation* (UR). In the two nasal rhymes, (3a&c), the nasal codas are deleted after adding the /-r/ suffix. The deletion of the nasal coda has varying effects on the preceding vowel, depending on the place of articulation of the nasal coda. In (3a), where the vowel precedes /n/, nasalization is removed in the suffixed forms (3d&g). But in (3c), the back nasalized vowel preceding /ŋ/ retains its nasalization in its diminutive forms (3f&i).

(3)	<b>Stem Form</b>	<b>Beijing Suffixed</b>	<b>Liaoning Suffixed</b>	<b>UR</b>	<b>English</b>
	(a) pān	(d) pāɿ	(g) pāɿ	/pan-r/	‘half.DIM’
	(b) pä	(e) päɿ	(h) päɿ	/pa-r/	‘handle.DIM’
	(c) pāŋ	(f) pāɿ	(i) pāɿ	/paŋ-r/	‘club.DIM’

The vowel quality of /a/ is also subject to change by r-suffixation. The rhotic suffix has a coarticulatory effect that centralizes the vowel (Liu 2017), backing the front vowel [a] and the central vowel [ä]. The front vowel becomes a central one in (3d&g), whereas the central vowel, despite some backing, remains categorically a central vowel, as shown in (3e&h). The back vowel [ä] in (3f&i) remains back in both dialects. The deletion of the nasal coda, vowel denasalization, as well as vowel backing, apply to both Beijing and Liaoning.

The two dialects differ in the distribution of the /-r/ allophones across the three low vowel rhymes. There are two rhotic allophones in Mandarin, the retroflex [ɿ] and the non-retroflex [ɻ] (Hsieh et al. 2019). In Beijing, both /an/ and /a/ rhymes choose the retroflex [ɿ] as suffix (3d&e), while /aŋ/ uses the non-retroflex [ɻ] (3f). In Liaoning, /a/ and /aŋ/ rhymes employ the non-retroflex [ɻ] (3h&i), whereas /an/ opts for the retroflex [ɿ] (3g). To put it more succinctly, Beijing and Liaoning only differ in the choice of rhotic allophone for the open-syllable rhyme /a/. It is [ɿ] in Beijing (3e), and [ɻ] in Liaoning (3h). For the nasal rhymes /an/ and /aŋ/, the two dialects agree on the choice of suffix. /an/ uses the retroflex rhotic, and /aŋ/ uses the non-retroflex one.

As seen in (3d&e), Beijing neutralizes /an/ and /a/ rhymes after r-suffixation, where they both surface as [pāɿ]. But Liaoning shows no neutralization, with /pan-r/ surfacing as [pāɻ]

and /pa-r/ surfacing as [p̥äɪ]. The two rhymes are distinguished from each other in the suffixed forms, due to a difference in the rhotic allophone.

Two questions arise from the difference between the stem forms and the suffixed forms: (i) why does vowel denasalization treat alveolar nasals and velar nasals differently? (ii) what accounts for the different distribution of /r/ allophony between the two dialects? Before addressing these two questions, acoustic and articulatory evidence for the changes in r-suffixation transcribed in (3) is provided. Section 2.1 explores vowel nasalization. In Section 2.2, I present acoustic evidence for vowel backing. Section 2.3 is a discussion of the two /r/ allophones.

## 2.1 Vowel nasalization

Between the two nasal rhymes, only one undergoes vowel denasalization in r-suffixation. In the stem forms, both /n/ and /ŋ/ nasalize the preceding vowel. During r-suffixation, the nasal coda is deleted to make way for the addition of the /r/ suffix, in order to avoid a complex coda (Zhang 2000). The conditioning environment for vowel nasalization is therefore lost. Indeed, in a nasal rhyme that ends in /n/, the preceding vowel loses its nasalization, and surfaces as an oral vowel in [p̥äɪ] (3d&g). But in a nasal rhyme ending in /ŋ/, the vowel retains its nasalization, as seen in (3f&i) [p̥äɪ̃]. This is true for both Liaoning and Beijing. Why is the process of vowel denasalization sensitive to the place of articulation of the deleted nasal coda? In other words, why does vowel denasalization only target alveolar nasal rhymes?

Zhang (2000) pointed out that the answer can be found in the stem forms of the nasal rhymes. In an aerodynamics study, Zhang measured the duration of nasal airflow in nasalized vowels in the stem forms produced by two Beijing speakers. The durational ratio of nasal airflow to oral airflow in the production of a nasalized vowel is taken as an indication of the degree of nasalization. The higher the nasal-to-oral airflow duration ratio is, the more nasalized the vowel is. It was found that /CVŋ/ rhymes show a much higher nasal-to-oral airflow duration ratio than /CVn/ rhymes do, which suggests that the velar nasal triggers more nasalization on a preceding vowel than the alveolar nasal does. Zhang used double tilde to mark the highly nasalized vowels in [CṼŋ]. I adopt this notation. The data in (3) are updated in (4) to reflect the notational change in vowel nasalization.

(4)	Stem Form	Beijing Suffixed	Liaoning Suffixed	UR	English
	(a) p̃än	(d) p̥äɪ	(g) p̥äɪ	/pan-r/	'half.DIM'
	(b) p̃ä	(e) p̥äɪ	(h) p̥äɪ	/pa-r/	'handle.DIM'
	(c) p̃äŋ	(f) p̥äɪ̃	(i) p̥äɪ̃	/paŋ-r/	'club.DIM'

According to Zhang, it is the strength of the nasal trigger that determines whether the vowel surfaces as a nasalized or an oral vowel in the r-suffixed form. The strong nasalization

triggered by a velar nasal is perceptually more salient, and therefore more likely to surface in the absence of the triggering nasal. Whereas the weak nasalization triggered by an alveolar nasal is less salient, and therefore more easily deleted when the environment for nasalization is lost. Zhang used a gradient input-output **Max[ +Nas]** constraint to capture this generalization. There is a **Max[ +Nas]** for every nasalization-triggering segment. The **Max[ +Nas]** associated with the stronger nasalization trigger always outranks the one associated with the weaker trigger. In the case of Mandarin, **Max[ +Nas]<sub>ŋ</sub>** ranks above **Max[ +Nas]<sub>n</sub>**. The two constraints are defined in (5), alongside a markedness constraint **\*V<sub>nas</sub>**, which punishes surface nasalized vowels found without a nasal trigger. The markedness constraint is ranked between the two faithfulness constraints, as seen in (5d), to derive the pattern seen in Mandarin diminutive forms. Vowels preceding /ŋ/ retain their nasalization in the suffixed forms (**Max[ +Nas]<sub>ŋ</sub>** » **\*V<sub>nas</sub>**), whereas those preceding /n/ get denasalized (**\*V<sub>nas</sub>** » **Max[ +Nas]<sub>n</sub>**). The ranking arguments are illustrated with the tableaux of (6). To put the limelight on vowel denasalization, only the subset of candidates that have undergone nasal coda deletion are considered in (6).

- (5) (a) **Max[ +Nas]<sub>ŋ</sub>**: If /ŋ/ is in the input, then [+nasal] must be in the output.  
 (b) **Max[ +Nas]<sub>n</sub>**: If /n/ is in the input, then [+nasal] must be in the output.  
 (c) **\*V<sub>nas</sub>**: No nasalized vowel is allowed in non-nasal environments  
 (d) **Max[ +Nas]<sub>ŋ</sub>** » **\*V<sub>nas</sub>** » **Max[ +Nas]<sub>n</sub>**

Zhang (2000)

- (6) (a) **Max[ +Nas]** constraints:

/CVn-ɿ/	<b>Max[ +Nas]<sub>ŋ</sub></b>	<b>*V<sub>nas</sub></b>	<b>Max[ +Nas]<sub>n</sub></b>
a. C $\tilde{V}$ ɿ		*!	
☞ b. CVɿ			*

- (b)

/CVŋ-ɿ/	<b>Max[ +Nas]<sub>ŋ</sub></b>	<b>*V<sub>nas</sub></b>	<b>Max[ +Nas]<sub>n</sub></b>
☞ a. C $\tilde{V}$ ɿ		*	
b. CVɿ	*!		

Zhang's **Max[ +Nas]** account places much importance on the preservation of individual features from input to output. The rhyme with a stronger degree of [+nasal] gets to preserve it in the output. I argue that it is actually the contrast between rhymes that is preserved in the suffixed output. The original contrast between the two inputs /CVn-ɿ/ and /CVŋ-ɿ/ is the place of articulation of the nasal coda. But it cannot be manifested in the suffixed

output, due to nasal coda deletion. Therefore the two rhymes have to resort to enhancing their contrast in vowel nasalization, by fully nasalizing one rhyme and denasalizing the other. This is achieved in coordination with many other features of the suffixed forms, to preserve the contrast between the low vowel rhymes. A more detailed analysis is presented in Section 4.

## 2.2 Vowel backing

In the stem forms, the backness of the low vowel is determined by the following nasal coda (or the lack of it). After r-suffixation, the conditioning environment for rhyme harmony is lost via nasal coda deletion. The low vowel is now situated next to the suffixed rhotic coda, which has a coarticulatory backing effect. The backing effect of a rhotic has been observed in many dialects of English (see Tunley 1999; Lilienthal 2009; Chung & Pollock 2014). In Mandarin, the vowels in the stem forms [ãn] and [ä] undergo vowel backing by the rhotic coda. The stem [ãŋ] is not affected by vowel backing, since its vowel is already back.

The coarticulatory vowel backing effect is observed in the recorded speech of a Beijing speaker and a Liaoning speaker (author). The second formant (F2), an indication of vowel backness, is measured across stem and suffixed forms. A front vowel has a high F2, while a back vowel has a low F2. It is found that for the rhymes /an/ and /a/, the F2 of the vowel is lower in the suffixed form than in the stem form. But the vowel in /aŋ/ rhymes have similar F2 for suffixed and stem forms.

Beijing Speaker		Estimate (Hz)	Std. Error (Hz)	Pr(>  t )
/an/	(Intercept)	1647.75	39.58	< 2e-16
	Formsuffixed	<b>-301.52</b>	39.58	2.46e-07
	Onsetch	142.04	48.48	0.00828
	Onsetf	-21.01	48.48	0.66943
/a/	(Intercept)	1512.01	62.66	2.92e-16
	Formsuffixed	<b>-211.78</b>	62.66	0.00298
	Onsetch	262.46	76.74	0.00271
	Onsetf	77.72	76.74	0.32329
/aŋ/	(Intercept)	1570.94	32.70	< 2e-16
	Formsuffixed	<b>-30.67</b>	32.70	0.3594
	Onsetch	27.84	40.05	0.4950
	Onsetf	-124.60	40.05	0.0055

**Table 1:** Change in F2 between Beijing stem and suffixed forms.

Liaoning Speaker		Estimate (Hz)	Std. Error (Hz)	Pr(>  t )
/an/	(Intercept)	1756.54	36.71	< 2e-16
	Formsuffixed	<b>-326.66</b>	36.71	2.15e-11
	Onsetch	98.10	44.96	0.0345
	Onsetf	-22.51	44.96	<i>0.6191</i>
/a/	(Intercept)	1598.26	45.21	< 2e-16
	Formsuffixed	<b>-153.34</b>	45.21	0.00148
	Onsetch	42.86	55.38	<i>0.44312</i>
	Onsetf	-33.98	55.38	<i>0.54262</i>
/aŋ/	(Intercept)	1531.18	41.21	< 2e-16
	Formsuffixed	<b>-53.16</b>	41.21	<i>0.204</i>
	Onsetch	-63.98	50.47	<i>0.212</i>
	Onsetf	-75.63	50.47	<i>0.141</i>

**Table 2:** Change in F2 between Liaoning stem and suffixed forms.

The recording of the two speakers, both female, was conducted in a soundproof booth, with a sampling rate of 44.1Khz, 32 bits. Measurements were taken using Praat (Boersma & Weenink 2017). The recording materials include minimal triplets shaped as Can vs. Ca vs. Caŋ, in both stem and suffixed forms. There are 3 minimal triplets, each with a different onset. They include a voiceless stop /p/, a fricative /f/, and an affricate /tʃ<sup>h</sup>/. The tones are consistent within each minimal triplet, but not controlled between triplets. All /p/-initial syllables are in the falling tone, while all syllables that begin with /tʃ<sup>h</sup>/ have a low dipping tone. To elicit the diminutive forms, a disyllabic or trisyllabic noun compound was formed using the low vowel rhyme syllable as the final morpheme, where the rhotic coda is attached to. The compounds were recorded as part of a carrier sentence [wo xwej ʃwo \_\_\_ tʃəkə ts<sup>h</sup>i ], which can be glossed as ‘I can say \_\_\_ the word’. Each compound was recorded in its stem form and diminutive form. Each sentence was repeated 4 times. There are altogether 2 (speakers) × 4 (repetitions) × 2 (forms) × 3 (rhymes) × 3 (onsets) = 144 tokens. For each token, the F2 value is taken at the visible vowel midpoint.

In the statistical analysis, the stem form and suffixed form of the same rhyme spoken by the same speaker are compared. For example, the Beijing speaker’s diminutive form /Can-r/ is compared against the stem form /Can/. A linear regression model on the vowel midpoint F2 is ran using RStudio (R Core Team 2018). Independent variables include morphological form (stem or suffixed) and onset type. Significance level is taken at  $p < 0.05$ . The results are summarized in **Table 1** (Beijing) and **Table 2** (Liaoning). The intercept is taken to represent the mean F2 of

the stem rhymes. The estimate for “Formsuffixed” indicates the change in F2 between stem and suffixed forms. A negative change in F2 indicates that the vowel F2 is lower in the suffixed form than in the stem form.

In both Beijing and Liaoning, the *p*-value for change in F2 between stem and suffixed forms have reached significance for /Can/ and /Ca/ forms. This shows that vowel backing has taken place for /an/ and /a/ rhymes after r-suffixation. However, F2 is not significantly different between stem and suffixed /Caŋ/ forms, which indicates that the vowel is not backed. Onset does not appear to significantly affect F2.<sup>1</sup>

The acoustic data support the description that the front [a] in the stem form (4a) [pān] is backed into a central [ä] in the suffixed form of (4d&g). The central [ä] in the open-syllable stem (4b) [pā], though also backed in r-suffixation, remains categorically central in the suffixed form in (4e&h), since it has roughly the same F2 realization as the central vowel derived from the [pān] stem (see mean F2 values in **Table 3** & **4**). The back vowel [ɑ] in the stem (4c) [pāŋ] remains the same in (4f&i).

I also list the mean vowel F2 of each rhyme form in **Table 3** (Beijing) and **Table 4** (Liaoning). In each table, the middle block lists the stem rhyme F2 values, whereas the rightmost block has the suffixed rhyme F2 values. Each row shows the mean F2 of a minimal triplet of rhymes that share the same onset and tone. Each cell represents the mean F2 of 4 tokens of the same lexical item.

Beijing Speaker			Stem Rhymes			Suffixed Rhymes		
Onset	Tone	F2(Hz)	/an/	/a/	/aŋ/	/an-r/	/a-r/	/aŋ-r/
/p/	Falling	Mean	1605.82	1433.95	1545.62	1388.15	1378.29	1565.58
		s.d.	104.83	127.88	153.93	24.21	30.71	43.97
/tʂ <sup>h</sup> /	Low	Mean	1828.00	1882.02	1628.09	1450.05	1455.14	1538.79
		s.d.	129.12	40.82	19.66	31.47	28.47	25.45
/f/	High	Mean	1630.46	1560.24	1442.34	1321.50	1407.44	1419.65
		s.d.	110.22	289.46	77.85	112.55	63.80	77.52

**Table 3:** mean F2 and for stem and suffixed rhymes in Beijing.

<sup>1</sup> Occasionally, the onset coded as “ch”, /tʂ<sup>h</sup>/, might display a low *p*-value. This might be due to a formant measurement issue. All /tʂ<sup>h</sup>-initial items happen to be in the dipping tone, which comes with laryngealization. This results in sporadic spectrograms interspersed with silence. It is difficult to identify vowel-coda boundaries. The vowel midpoint might also fall on periods of silence, making F2 measurement less than accurate.

Liaoning Speaker			Stem Rhymes			Suffixes Rhymes		
Onset	Tone	F2(Hz)	/an/	/a/	/aŋ/	/an-r/	/a-r/	/aŋ-r/
/p/	Falling	Mean	1836.84	1679.49	1441.07	1542.04	1594.62	1466.12
		s.d.	110.31	43.52	39.93	74.05	56.74	153.63
/tʂ <sup>h</sup> /	Low	Mean	1962.68	1501.67	1422.34	1524.51	1418.94	1173.24
		s.d.	94.25	221.25	87.34	33.02	58.71	21.37
/f/	High	Mean	1826.63	1549.92	1428.39	1504.23	1432.83	1425.49
		s.d.	78.10	33.08	135.36	136.71	31.17	222.23

**Table 4:** mean F2 and for stem and suffixed rhymes in Liaoning.

There are some unexpected F2 values<sup>2</sup> in **Tables 3** and **4**. For instance, in **Table 3**, the vowel in the /a/ rhymes are not consistently central for the Beijing speaker. In /p/-initial syllables, it has a similarly low mean F2 as the back vowel in /aŋ/. But in /tʂ<sup>h</sup>/-initial syllables, it is as front as the front vowel in /an/. It could be the case that the vowel in an open syllable is not necessarily specified as central, but more likely to be unspecified in backness. Without a following nasal coda to dictate its surface representation, it can vary in its F2 value. Although such a wide range of variation in F2 is not found in the open-syllable rhyme produced by the Liaoning speaker. As to the suffixed rhymes, they usually have low F2 values across the board. But it is interesting to note that for the Beijing speaker, the suffixed /an/ and /a/ rhymes sometimes have lower F2 than the originally back vowel in the /aŋ/ rhymes. Since the vowel in /aŋ-r/ is heavily nasalized, it is possible that it might have colored the F2 measurement, thus making the comparison with other rhymes difficult.

### 2.3 /r/ allophones

Up to this point, I have discussed phonological processes that are common between Beijing and Liaoning in r-suffixation. They both undergo nasal coda deletion, vowel denasalization, and vowel backing. In this section, I introduce the major point of difference between the two dialect — the distribution of the /r/ allophone across the three rhymes.

---

<sup>2</sup> Some reviewers have pointed out the unexpected F2 values might be due to the sample size being too small. It is hard to see whether difference in behavior between the two speakers can be attributed to dialectal difference or between-speaker variation. In order to get a clearer picture of the F2 values of stem and suffixed rhymes, tokens produced by more speakers need to be collected in future research.

The /-r/ suffix has two surface realizations: the retroflex [ɻ] and the non-retroflex, or bunched, approximant [ɹ] (Shuwen Chen et al. 2017; Hsieh et al. 2019; Jiang et al. 2019). The retroflex [ɻ] is produced with the tongue tip raised and bent towards the palate. [ɹ], on the other hand, is not a retroflex rhotic. During its production, the tongue tip is located in the lower region of the oral cavity, pointing downwards, behind the back of the lower teeth. There is no discernible upwards curving of the tongue tip.

Jiang et al. (2019) have identified two types of rhotic codas in an EMA study of the speech of Liaoning speakers. They compared the rhotic coda in two environments. One is the diminutive suffix /-r/ that attaches to a monosyllabic open-syllable stem like /a/. The other is the coda /r/ found in the monomorphemic word [ɕ] ‘child’, where it is not an affix. The tongue position of the speaker was tracked throughout the production of the rhyme in both cases. It was discovered that the affixational /-r/ found after monophthong stems mainly involves tongue-body movement, whereas the monomorphemic /r/ in [ɕ] is essentially produced with tongue-tip movement. The former is the non-retroflex [ɹ] used next to /a/ rhymes in Liaoning, while the latter corresponds to the retroflex [ɻ] found in /an/ stems. Jiang et al. suggest that the difference between the two rhotic allophones might be quite similar to that between the bunched and the retroflex rhotic found in English (see Delattre & Freeman 1968). The same group of researchers also compared the rhotic suffix between 3 Beijing speakers and 3 northeastern speakers. They found that after an open-syllable stem, Beijing speakers produce a retroflex suffix, whereas northeastern speakers have a non-retroflex suffix (Hsieh et al. 2019).

In Liaoning, /an/ rhymes select the retroflex [ɻ] as the diminutive suffix (7d), while /a/ rhymes select the non-retroflex [ɹ] (7e). But in Beijing, both /an/ and /a/ rhymes employ the retroflex [ɻ] for r-suffixation, as seen in (7a&b). When it comes the velar nasal /aŋ/ rhyme, the non-retroflex [ɹ] is used in both dialects, as shown in (7c&f).

(7)	Beijing Suffixed	Liaoning Suffixed	UR	English
	(a) pǎɻ	(d) pǎɻ	/pan-r/	‘half.DIM’
	(b) pǎɻ	(e) pǎɹ	/pa-r/	‘handle.DIM’
	(c) pǎ̃ɹ	(f) pǎ̃ɹ	/paŋ-r/	‘club.DIM’

The difference between the retroflex [ɻ] and the non-retroflex [ɹ] can be observed in formant transitions. In a production experiment, it was found that at the end of a suffixed rhyme with a retroflex [ɻ], F1 lowers and F2 rises. While in a suffixed rhyme with a non-retroflex [ɹ], F1 and F2 stay roughly the same towards the end of the rhyme.

4 female Liaoning speakers and 2 Beijing speakers (1 male, 1 female), who now live in Boston, were recruited for the experiment. They were all born in Liaoning or Beijing and have

lived there for most of their childhood. They were recorded producing suffixed forms in isolation as well as in carrier sentences. The recordings were made in a soundproof booth. The sampling rate was 44.1Khz, 32 bits.

The test items include 100 disyllabic or trisyllabic words (one item has 4 syllables). Among them, 73 are of the shapes /Can/, /Ca/, or /Caŋ/, all beginning with a stop consonant. The rest are filler items. The test items consist of nouns and verb or adjective compounds that contain a noun as the rightmost morpheme. Some words are selected on the basis that they sound extremely unnatural with a diminutive suffix. For instance, [ts<sup>h</sup>wen tʂaŋ] ‘head of village’, is a title for someone in a position of authority. It is very unlikely to appear in the diminutive form. These words are placed in the experiment to make sure that speakers do not blindly judge every suffixed form to be natural sounding.

During the experiment, for each test item, the speakers are first asked to read out loud a sentence shown on a computer screen. The sentence is in the form of a question: [ni tʂetə k<sup>h</sup>əji tʂja ə xwajin ma], ‘Do you think \_\_\_\_\_ can take on an erhua suffix?’ Next, they are encouraged to attempt to produce the suffixed form. Speakers can use this opportunity to figure out whether the suffixed form is natural-sounding to them or not. During this period, isolated forms are recorded as well. Afterwards, the speakers are presented with a choice between two answer sentences displayed on the screen. The first answer reads: ‘I think \_\_\_\_\_ can take on an erhua suffix.’ ([wo tʂetə \_\_\_\_\_ k<sup>h</sup>əji tʂja ə xwajin]). The other reads: ‘I think \_\_\_\_\_ cannot take on an erhua suffix.’ ([wo tʂetə \_\_\_\_\_ pu k<sup>h</sup>əji tʂja ə xwajin]). The speakers are asked to make a choice based on their judgement of the naturalness of the suffixed form they have just attempted to produce, and to read that answer out loud. The answer sentence not only serves the purpose of recording suffixed forms in a carrier sentence, but also ensures that the suffixed tokens collected are actually natural sounding to the individual speaker.

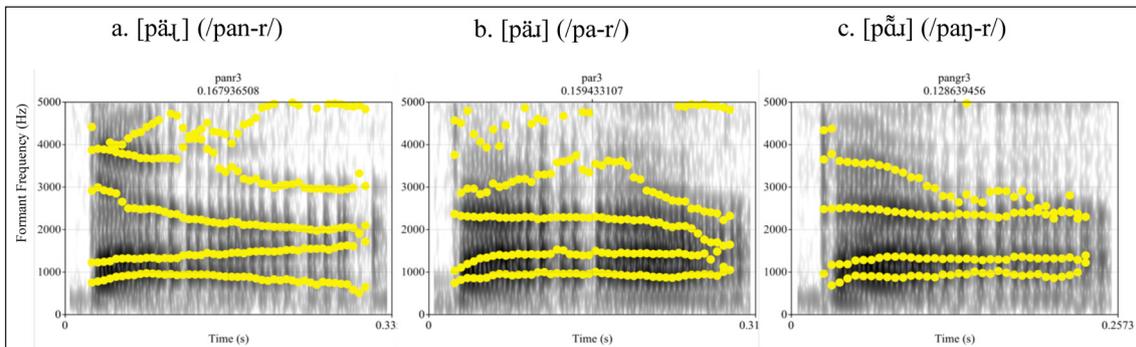
The procedure is repeated for the 100 nouns. Of the 73 low vowel rhymes that are of interest, only those that pass the naturalness test are measured for their formant values.<sup>3</sup> This means that there is no fixed number of tokens analyzed for each speaker, since there is a lot of between-speaker variation in the naturalness judgement of the suffixed forms.

**Figure 1** shows the spectrograms of suffixed forms produced by the same female Liaoning speaker. The spectrograms are generated using Praat. Each spectrogram represents one of the

---

<sup>3</sup> Only natural-sounding items are collected because we want to find out how speakers produce suffixed rhymes already present in their lexicon. To add the suffix to a new word item, speakers might be inadvertently changing their articulation. For example, lengthening the vowel or exaggerating the rhotic tongue movement.

three rhymes, /an/, /a/, and /aŋ/, in suffixed forms. They all begin with a bilabial unaspirated stop /p/, and are all in the third tone.



**Figure 1:** Spectrograms of low vowel rhymes produced by a female Liaoning speaker.

**Figure 1.a** is the spectrogram of [pãɿ] (/pan-r/), and **Figure 1.b** displays that of [pãɿ] (/pa-r/). In **Figure 1.a**, where there is a retroflex coda [ɿ], F1 and F2 diverge towards the end of the rhyme. F1 lowers and F2 rises. Compare this with the formant transition at the end of the rhyme in **Figure 1.b**, where F1 and F2 both stay around the same level for a non-retroflex coda [ɿ]. The same can be said for **Figure 1.c**, where [pãɿ] (/paŋ-r/) also ends in a non-retroflex coda.

The F1-F2 gaps at the end of suffixed rhymes in isolation are measured. There are 99 Liaoning tokens from 4 speakers, and 55 tokens from the 2 Beijing speakers. **Table 6** displays the mean F1-F2 gap of each suffixed rhyme produced by the Liaoning speakers. The retroflex [ɿ] has a big F1-F2 gap of nearly 900Hz, whereas the non-retroflex [ɿ] has a small F1-F2 gap at about 550Hz. This shows that F1-F2 gap is a good acoustic indicator that distinguishes between retroflex and non-retroflex rhotic codas. **Table 5** shows the mean F1-F2 gaps produced by the Beijing speakers. Both /an-r/ and /a-r/ display a big F1-F2 gap at around 1000Hz, indicating a retroflex [ɿ]. /aŋ-r/ has a small F1-F2 gap below 700Hz for the non-retroflex [ɿ].

Beijing Speakers	Suffixed Rhyme	Rhyme UR	R	Mean F1-F2 Gap	S.D.
	ãɿ	/an-r/	ɿ	1020.23Hz	160.50Hz
	ãɿ	/a-r/		965.29Hz	148.32Hz
	ãɿ	/aŋ-r/	ɿ	681.84Hz	160.56Hz

**Table 5:** Mean F1-F2 gap of suffixed rhymes produced by Beijing speakers, measured at end of rhymes.

Liaoning Speakers	Suffixed Rhyme	Rhyme UR	R	Mean F1-F2 Gap	S.D.
	äɿ	/an-r/	ɿ	884.06Hz	42.21Hz
	äɪ	/a-r/	ɪ	569.84Hz	48.49Hz
	ä̃ɪ	/aŋ-r/		554.61Hz	44.90Hz

**Table 6:** Mean F1-F2 gap of suffixed rhymes produced by Liaoning speakers, measured at end of rhymes.

### 3. Dispersion theory

The analysis for Mandarin r-suffixation presented in this paper is based on Flemming's (1995) Dispersion Theory, in which the realization of contrast is taken as the driving force of phonological processes. This framework is adopted by many recent studies of phonology (see Varis 2011; Dmitrieva 2012; Luo 2016; Petersen 2016; Wang 2017; Stanton 2018; Magri & Storme 2019).

According to Flemming, the contrastive sound inventory of a language is the result of two competing forces. One is the imperative for the language to realize as many contrasts as possible, and the other is to make sure that each contrast is perceptually salient. In Dispersion Theory, a language manifests its preference to have as many contrasts as possible in a constraint called *MaximizeContrasts*. For a human language to be able to encode a long list of different meanings into distinct units in the lexicon, it needs as many different sound sequences at its disposal as possible. Imagine a language with only one contrast in its sound inventory, say, between a bilabial stop /b/ and a low vowel /a/. An infinite number of sound sequences, or words, can be derived by creating strings of various lengths with the two phonemes. One can encode meanings with strings such as /ba/, /ab/, /abba/, /bababababa.../, etc.. But this is a very inefficient language from a communication point of view, for words and sentences might take too much time to produce. A simple way to solve this problem is to increase the number of phonemes that can be used for word formation, which means that a greater number of contrasts between different sounds is preferred.

At the other end of the spectrum, is a language with too many contrasts. A contrast-abundant language can indeed derive a sizable lexicon in which words are all of a reasonable length. But it might very well be the case that two words with different meanings are perceptually too similar for speakers to discern. For example, if a language has a contrast between a word /ba<sub>1</sub>/ and another word /ba<sub>2</sub>/, where /a<sub>1</sub>/ has a F2 of 1800Hz and /a<sub>2</sub>/, 1850Hz. The two words would sound too similar for the listener to distinguish. Miscommunications ensue. To counter the problem, a language will only establish a contrast if it is perceptually distinguishable. To capture this preference, Flemming uses a set of constraints he termed *MinimalDistance* (or

*MinDist*). They act on specific dimensions of auditory properties, and dictate a specific minimum distance on the dimension between two contrasting sounds. For example, a *MinDist* constraint on the F2 dimension can punish two low vowels, /a<sub>1</sub>/ and /a<sub>2</sub>/, for having an F2 contrast of mere 50Hz.

When a contrast is too small, sometimes the grammar opts for not realizing the contrast at all, therefore resulting in neutralization. This is indeed what happens in the r-suffixation of Beijing Mandarin, where /an/ and /a/ neutralize in the suffixed forms as [päɹ] to avoid an imperceptible contrast. In Liaoning, however, priority is placed on *MaximizeContrasts*. /an/ and /a/ each chooses a different rhotic allophone to manifest the contrast in the suffixed forms, as [päɹ] and [päɹ̥] respectively.

#### 4. The contrast preservation analysis

The contrast preservation analysis presented in this section has two goals. One is to derive the sounds used for contrast in diminutive forms of the Mandarin low vowel rhymes. The other is to account for the maintenance of identity between stem and diminutive forms. For the former, I adopt Flemming's (1995) *Dispersion Theory*. For the latter, I incorporate *BD faithfulness constraints*.

Faithfulness constraints are absent in Flemming's formulation of *Dispersion Theory*. Neither input-output mappings nor output-output mappings are evaluated in this approach. All that *Dispersion Theory* is concerned with is producing the right inventory for a specific phonological environment. I depart from Flemming's original framework by incorporating *BD faithfulness constraints* in my analysis. This is because Mandarin r-suffixation is a morphophonological process, where the derived diminutive forms are required to stay as faithful as possible to the stem forms. Without reference to any output-output correspondence, the grammar will only produce two perfectly contrastive inventories, one for the stem forms, and one for the suffixed forms, with no way for the speaker to know which stem form maps onto which suffixed form.

Another modification to Flemming's original framework involves a change in the unit of speech evaluated for contrast preservation. Usually the constraints act on individual segments. In Mandarin r-suffixation, however, the basic unit of contrast preservation is the rhyme. Each segment in the rhyme contributes to the preservation of between-rhyme contrast from stem forms to suffixed forms.

In the UR of the three low-vowel rhymes, /an/, /a/, /aŋ/, the source of contrast solely resides in the nasal coda (or the absence of it). In the stem forms, the low vowel displays variation to enhance the nasal coda contrast. The place of articulation of the nasal determines the backness of the vowel via rhyme harmony, and affects the degree of nasalization of the vowel. In the

suffixed forms, the nasal coda is deleted to make way for the diminutive suffix, thus the original contrast in the UR cannot surface. The burden of contrast is moved to the remaining vowel and the newly introduced /-r/ suffix. This transfer of contrast is summarized in **Table 7**. The bold IPA symbols in the rhyme UR column point to the segments that originally constitute the contrast. An empty set symbol “ $\emptyset$ ” denotes the absence of a coda. In the SR columns, the segments that show allophonic variation are marked in italics. The two suffixed SR columns have IPA symbols that are both in italics and bold. This means that these segments show allophonic variation, but they have now taken on the burden of expressing contrast in the suffixed forms, where the segmental source of the contrast has been deleted.

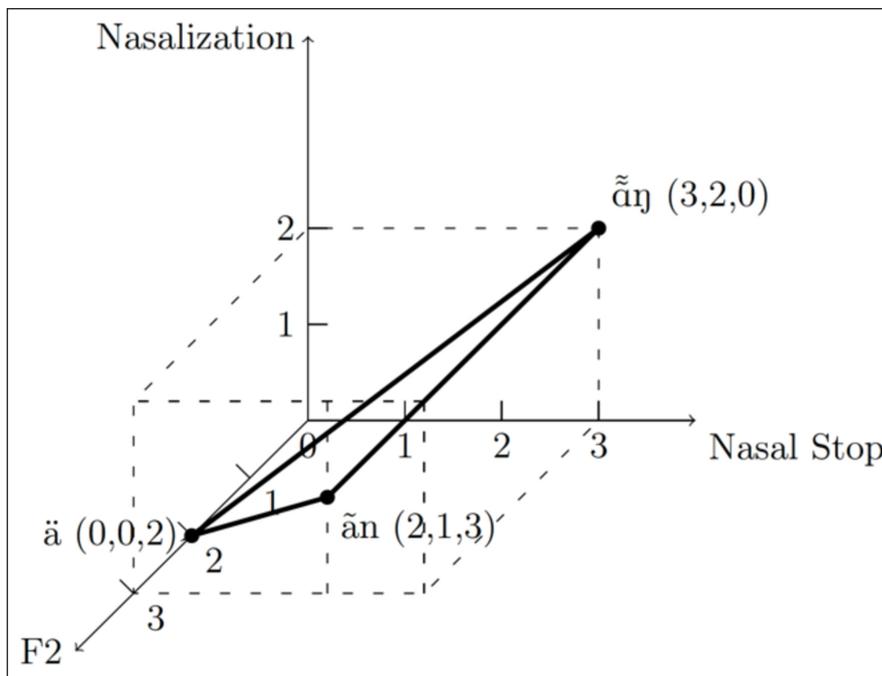
Rhyme UR	Stem SR	Suffixed SR	
		Beijing	Liaoning
/an/	[ <i>ã</i> n]	[ <i><b>ã</b></i> ɹ]	[ <i><b>ã</b></i> ɹ]
/a $\emptyset$ /	[ <i>ä</i> $\emptyset$ ]	[ <i><b>ä</b></i> ɹ]	[ <i><b>ä</b></i> ɹ]
/aŋ/	[ <i><b>ã</b></i> ŋ]	[ <i><b>ã</b></i> ɹ]	[ <i><b>ã</b></i> ɹ]

**Table 7:** Transfer of contrast.

In the stem and suffixed SR’s, the contrast between rhymes is expressed via a series of auditory cues contributed by the vowel and the coda. They include vowel F2, degree of vowel nasalization, the nasal coda’s oral closure duration, as well as F1-F2 gap of the rhotic suffix. In the contrast preservation analysis, each auditory cues corresponds to a dimension in a rhyme space. Each rhyme SR can be identified with a set of coordinates in this rhyme space, composed of individual values each corresponding to a particular auditory cue dimension. The Euclidean distance between any two rhymes is calculated by the grammar, to determine if MinDist requirement has been met. In Section 4.1, I introduce the mechanics of the three-dimensional rhyme space of stem rhymes. Section 4.2 shows the suffixed rhyme space. Contrast preservation constraints, MinDist=RD:2 and PreserveContrasts-BD, are introduced in Section 4.3. A discussion of /r/ allophony constraints is presented in Section 4.4. Section 4.5 introduces the F2 constraints, and Section 4.6 the nasal constraints.

#### 4.1 The three-dimensional rhyme space

Three auditory cues are used to distinguish between the stem SR of the low vowel rhymes: nasal stop’s oral closure duration, the degree of vowel nasalization, and vowel F2. The three cues correspond to the x-axis, the y-axis, and the z-axis in a three-dimensional rhyme space, respectively. This is illustrated in **Figure 2**, with coordinates for each rhyme labeled in parenthesis.



**Figure 2:** Three-dimensional stem rhyme space.

The z-axis has abstract values of vowel F2 expressed as integers. The bigger the integer, the higher the F2. A difference of 1 on the F2 axis corresponds to roughly 100Hz. This is based on Zee & Lee's (2001) measurements of the low vowel produced by speakers of Beijing Mandarin in different rhymes. They report that for 10 female speakers, the mean F2 of the front low vowel found in [ãŋ] is 1730Hz. The central vowel in [ä] has a mean F2 of 1594Hz, while the back vowel in [ã] has a mean of 1382Hz. The difference between the front [a] and central [ä] is roughly 100Hz, and that between the central [ä] and back [a], about 200Hz. This is why the central [ä] corresponds to a value of 2 on the F2 axis, 1 unit lower than the front [a] (3), and 2 units higher than the back [a] (0).

The y-axis of vowel nasalization is an abstraction from Zhang's (2000) observation that there are two degrees of nasalization in the stem forms. The number on this axis corresponds to the number of nasalization tilde marks. The weakly nasalized [ã] has a value of 1 on the nasalization axis, while the strongly nasalized [ã̃] has a value of 2. The oral vowel [ä] sits low at 0 on this axis.

The numbers on the x-axis of nasal stop are based on the duration of the oral closure of the two nasal codas. Li & Cheng (2014) reported that the two nasal codas of Mandarin differ in their duration of oral closure. /n/'s oral closure is proportionally shorter than that of /ŋ/, when compared to the duration of the entire nasal rhyme. Therefore I assign a higher numerical value

for /ŋ/ (3) on the nasal stop *x*-axis, and a lower one for /n/ (2). The open-syllable rhyme has a value of 0 on the nasal stop axis. Note that the distance between /n/ and the absence of a nasal coda is 2, larger than that between /n/ and /ŋ/. This is because perceptually, the difference between a nasal coda and the total absence of any nasal segment, is greater than that between two nasal stops of different lengths. The assignment of values for each dimension is summarized in **Table 8**.

The choice of oral closure duration as the auditory cue corresponding to the nasal stop *x*-axis might be a controversial one. It could be argued that vowel-to-nasal transition F2 is a better indicator for the place of articulation of the nasal coda (see Liberman et al. 1954; Malécot 1956; Recasens 1983; Harding & Meyer 2003). Indeed, in Mandarin, vowel-to-nasal transition F2 can be used to distinguish between alveolar and velar nasals. Faytak et al. (2020) measured the formant values at the vowel endpoint preceding a nasal coda in Standard Mandarin. They found that between /n/ and /ŋ/ codas, there is approximately a 300Hz difference in F2. I concur that vowel-to-nasal transition F2 is a good auditory cue for nasal place of articulation. However, it cannot be used to model the difference between open-syllable rhymes and nasal rhymes. After all, the nasal stop *x*-axis is required to provide a coordinate for each of the three low vowel rhymes, /an/, /a/, /aŋ/. Had vowel-to-nasal transition F2 been used as the auditory cue for this axis, it would not be possible to represent the open-syllable rhyme /a/ on the axis. /a/ does not have a nasal coda, and therefore there is no vowel-to-nasal transition F2. It is for this reason that I have decided to use oral closure duration as the auditory cue for the nasal stop axis. The open-syllable rhyme, despite having no nasal coda, can be interpreted as having an oral closure duration of zero. Thus one can compare /a/ with the two nasal rhymes along the nasal stop *x*-axis.

Axis	Auditory Cue	Numerical Values			
		0	1	2	3
<i>x</i> -axis	Nasal Stop	∅		n	ŋ
<i>y</i> -axis	Vowel Nasalization	V	Ṽ	Ṽ̃	
<i>z</i> -axis	Vowel F2	ɑ		ä	a

**Table 8:** Numerical values on the auditory cue dimensions.

Using the criteria of **Table 8**, the rhyme [ãŋ] has the coordinates (2,1,3) in the three-dimensional rhyme space. It has a value of 2 on the *x*-axis for /n/, 1 on the *y*-axis for one degree of vowel nasalization, and 3 on the *z*-axis for having the highest F2. Similarly, [ä] has coordinates of (0,0,2), and [ãŋ], (3,2,0). The three rhyme SR's, with their coordinates, can be found in the rhyme space of **Figure 2**. The locations of the rhymes are labeled with black dots. The dashed

lines are drawn to help visualize their locations in the three-dimensional space. Concrete lines connect each pair of rhymes, showing their distance, which are calculated in (8).

$$(8) \quad d(/an/, /a/) = d(\tilde{a}n_{(2,1,3)}, \tilde{a}_{(0,0,2)}) = \sqrt{(2-0)^2 + (1-0)^2 + (3-2)^2} = \sqrt{6} \approx 2.45$$

$$d(/a/, /aŋ/) = d(\tilde{a}_{(0,0,2)}, \tilde{a}ŋ_{(3,2,0)}) = \sqrt{(0-3)^2 + (0-2)^2 + (2-0)^2} = \sqrt{17} \approx 4.12$$

$$d(/aŋ/, /an/) = d(\tilde{a}ŋ_{(3,2,0)}, \tilde{a}n_{(2,1,3)}) = \sqrt{(3-2)^2 + (2-1)^2 + (0-3)^2} = \sqrt{11} \approx 3.32$$

The smallest distance between any pair of rhymes is that between [ãn] and [ã], at approximately 2.45, which satisfies a **MinDist = RD:2** constraint, as defined in (9).

- (9) **MinDist = RD:2:** The Euclidean distance between two rhymes must be bigger than or equal to 2 in the three-dimensional rhyme space. (RD stands for Rhyme Distance.)

## 4.2 Suffixed rhyme space

After r-suffixation, the SR rhyme space takes on a different look, for the nasal stop x-axis ceases to be relevant for contrast representation due to nasal coda deletion. The nasal coda deletion process itself can be accounted for using the following three constraints: **\*ComplexCoda**, **RealizeAffix**, and **Template**, adopting the analysis in Zhang (2000). Their definitions can be found in (10). In (11), **\*ComplexCoda** rules out the faithful candidate (a) [C $\tilde{V}$ nɿ], which contains both the nasal coda and the /-r/ suffix. One of the coda consonants has to be deleted, and **RealizeAffix** makes sure it is the stem consonant that gets deleted. The constraint places a priority on realizing the affix over the stem, based on Lin's (1993) *Affix Manifestation Principle*. Thus candidate (d) [C $\tilde{V}$ n] is ruled out. Finally, there is also the possibility of realizing the /-r/ suffix in a separate syllable, as in candidate (b) [C $\tilde{V}$ n.ɿ]. **Template** dictates that the diminutive rhyme should be realized as a single syllable. All 3 constraints are ranked above Max, resulting in the win of the nasal-coda-deletion candidate (c) [CVɿ].

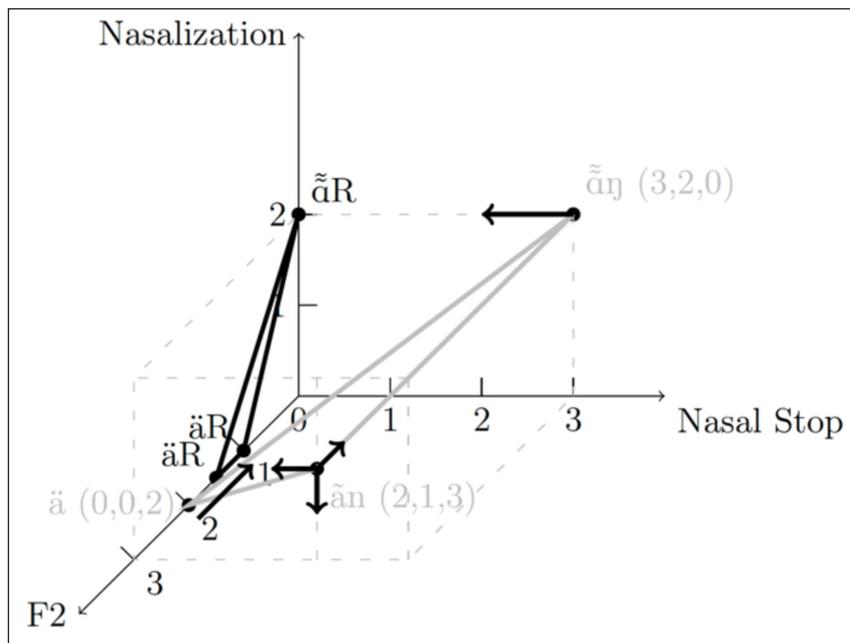
- (10) (a) **\*ComplexCoda:** No complex coda is allowed.  
 (b) **RealizeAffix:** Affixes must be realized.  
 (c) **Template:** The suffixed form must be one syllable.  
 Zhang (2000)

- (11) Nasal coda deletion

/CVn-ɿ/	*ComplexCoda	RealizeAffix	Template	Max
a. C $\tilde{V}$ nɿ	*!			
b. C $\tilde{V}$ n.ɿ			*!	
☞ c. CVɿ				*
d. C $\tilde{V}$ n		*!		*

Zhang (2000)

With the deletion of the nasal coda, all three rhymes now have a value of 0 on the nasal stop  $x$ -axis. It can be visualized as all three rhymes being squashed to the  $yz$ -plane on the left hand side, as illustrated in **Figure 3**. The original locations of the stem rhymes are marked in gray. The movement the rhymes undergo after  $r$ -suffixation are shown in arrows, which include vowel backing along the  $F2$  axis, and denasalization along the nasalization axis. The  $/-r/$  suffix is written with an abstract  $R$  here, since its SR is specific to each dialect.

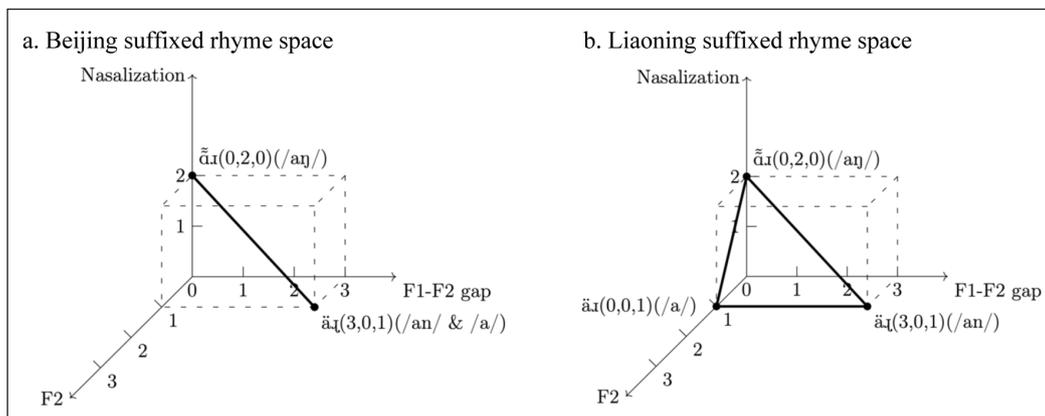


**Figure 3:** The rhyme space after nasal coda deletion.

In the suffixed rhyme space, a fourth dimension needs to be taken into consideration, to include auditory cues to the rhotic suffix. This is the  $F1$ - $F2$  gap dimension. Recall that the  $F1$ - $F2$  gap at the end of a rhyme is an indicator for whether the rhotic suffix is retroflex or not (see Section 2.3). **Figure 4** shows the suffixed rhyme space with the added  $F1$ - $F2$  gap dimension. **Figure 4.a** is the suffixed rhyme space of Beijing. **Figure 4.b** represents Liaoning. Here, I reuse the  $x$ -axis for the  $F1$ - $F2$  gap dimension, since the nasal stop dimension is no longer relevant in the suffixed forms.<sup>4</sup> As shown in **Table 6**, the retroflex rhotic suffix  $[-ɻ]$  has an  $F1$ - $F2$  gap of approximately 850-900Hz at the end of rhymes, whereas the non-retroflex  $[-ɹ]$  has one of about 550Hz. The two types of  $/-r/$  allophones are separated by approximately 300Hz. Recall that each unit on the  $F2$  axis corresponds to a difference of 100Hz. The same

<sup>4</sup> This is done for ease of display, and should not be taken to suggest that the  $F1$ - $F2$  gap dimension and the nasal stop dimension are the same.

rule of thumb is applied to the F1-F2 gap axis as well. Therefore, the 300Hz difference is represented on the x-axis as a difference of 3 between the non-retroflex [ɿ] at 0, and the retroflex [ɿ̣] at 3.



**Figure 4:** Suffixed Rhyme Space.

As shown in **Figure 4.a**, in Beijing, the rhymes /an-r/ and /a-r/ share the same suffixed form, occupying the same location in the suffixed rhyme space. [äɿ] has the coordinates (3,0,1), where 3 on the F1-F2 gap axis points to a retroflex suffix [-ɿ], 0 on the nasalization axis indicates that it is an oral vowel, and the 1 on the F2 axis is a slightly backed central vowel [ä]. The rhyme /aŋ-r/ surfaces as [ä̃ɿ], with coordinates of (0,2,0). The first 0 is for the non-retroflex [ɿ] on the F1-F2 gap axis, 2 signifies the degree of nasalization on the vowel, and the second 0 indicates a back vowel on the F2 axis. The rhyme distance is calculated in (12). Since there are only two surface forms in Beijing suffixed rhymes, there is only one rhyme distance.

(12) *Suffixed rhyme distance, Beijing*

$$d(/a/, /aŋ/) = d(/an/, /aŋ/) = d(\tilde{a}_ɿ(3,0,1), \tilde{a}_ɿ(0,2,0)) = \sqrt{(3-0)^2 + (0-2)^2 + (1-0)^2} = \sqrt{14} \approx 3.74$$

Liaoning (**Figure 4.b**), on the other hand, retains the contrast between /an-r/ and /a-r/ in the suffixed forms. /an-r/ surfaces as [ä̃ɿ], with a retroflex [ɿ̣], which is a 3 on the F1-F2 gap axis. Whereas /a-r/ surfaces as [äɿ], with a non-retroflex [ɿ], corresponding to 0 on the F1-F2 gap axis. The /aŋ-r/ rhyme has a value of 0 on the F1-F2 gap axis, just like its counterpart in Beijing. The rhyme distance between each pair of suffixed rhymes in Liaoning is calculated in (13).

(13) *Suffixed rhyme distance, Liaoning*

$$\begin{aligned} d(/an-r/, /a-r/) &= d(\tilde{a}_ɿ(3,0,1), \tilde{a}_ɿ(0,0,1)) = \sqrt{(3-0)^2 + (0-0)^2 + (1-1)^2} = 3 \\ d(/a-r/, /aŋ-r/) &= d(\tilde{a}_ɿ(0,0,1), \tilde{a}_ɿ(0,2,0)) = \sqrt{(0-0)^2 + (0-2)^2 + (1-0)^2} = \sqrt{5} \approx 2.24 \\ d(/aŋ-r/, /an-r/) &= d(\tilde{a}_ɿ(0,2,0), \tilde{a}_ɿ(3,0,1)) = \sqrt{(0-3)^2 + (2-0)^2 + (0-1)^2} = \sqrt{14} \approx 3.74 \end{aligned}$$

### 4.3 Contrast preservation constraints

As seen in the calculations of suffixed rhyme distances in (12) and (13), neither Beijing nor Liaoning violates  $\text{MinDist} = \text{RD}:2$ . The smallest rhyme distance in Beijing is 3.74, and 2.24 in Liaoning. However, the number of contrasts remaining in the suffixed forms differ between the two dialects. In Beijing, the /an/ and /a/ rhymes are neutralized after r-suffixation. Therefore, there are only two contrastive suffixed rhymes in Beijing (**Figure 4.a**). Liaoning, on the other hand, retains all three contrastive rhymes in the suffixed forms, as shown in **Figure 4.b**. There is no neutralization between any of the three rhymes.

**PreserveContrast-BD**, as defined in (14), rewards Liaoning for retaining all three contrastive rhymes in the suffixed forms, while penalizing Beijing for its neutralization. It is adapted from Flemming's (1995) *MaximizeContrasts*, to evaluate mappings from stem forms to suffixed forms.

- (14) **PreserveContrasts-BD**: The number of contrasts in the derived forms should match the number of contrasts in the base forms. Assign a check mark for each contrastive base form that is realized in the derived forms.

Similar to *MaximizeContrasts*, the constraint assigns checkmarks instead of violation marks. But *PreserveContrasts-BD* departs from Flemming's constraint in that it evaluates the number of contrasts in the derived forms in relation to the base forms. Instead of saying "have as many contrasts as you can", *PreserveContrasts-BD* says "keep as many contrasts from the base forms as you can in the derived forms".

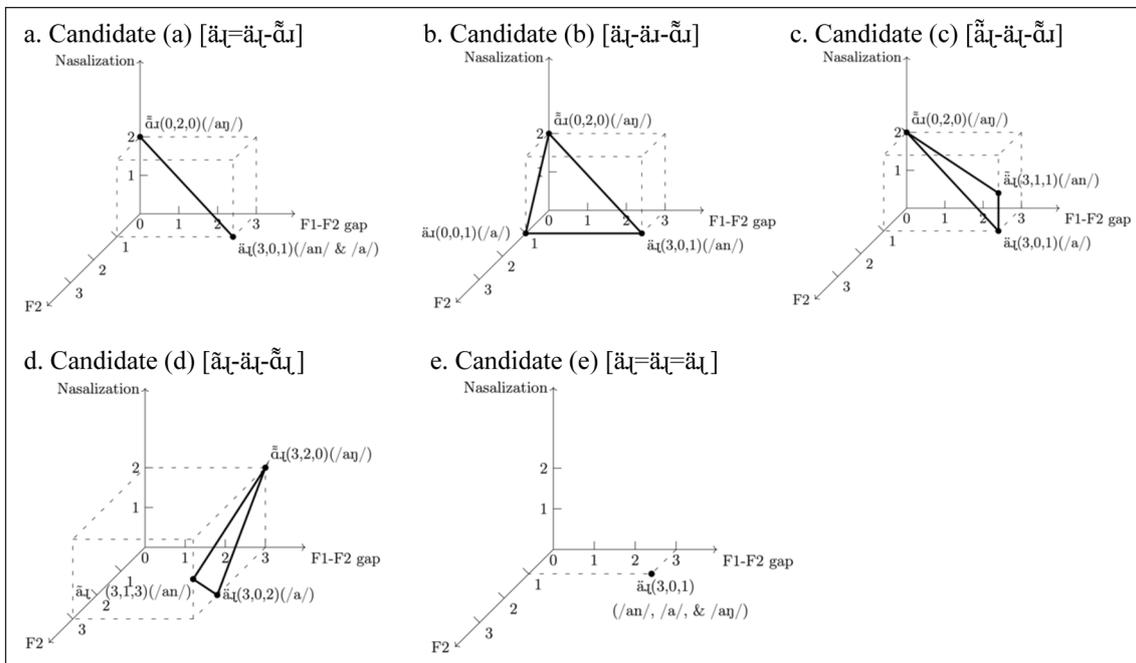
*PreserveContrasts-BD* has a different ranking in Beijing and Liaoning. In Beijing, it is ranked below  $\text{MinDist} = \text{RD}:2$  and  $\text{Ident}[\text{Retroflex}]-\text{IO}$  (more on this constraint in Section 4.4), for the dialect ensures that the rhyme distance between each suffixed rhyme is big enough, at the expense of losing contrasts between rhymes. This is shown in the tableau of (15a). In Liaoning, however, *PreserveContrasts-BD* is ranked as high as  $\text{MinDist} = \text{RD}:2$ . This is because Liaoning puts as much importance on retaining rhyme contrasts as maintaining rhyme distances. The ranking of Liaoning is shown in (15b).

The candidates of the tableaux in (15) are not individual output forms, but sets of output forms. Each candidate represents a suffixed rhyme space, illustrated separately in **Figure 5**. They are evaluated against the stem rhyme space. In each candidate, the suffixed rhymes, from left to right, correspond to /an-r/, /a-r/, and /aŋ-r/. This is the way to read tableau candidates throughout this paper.

The five candidates are the same in (15a) and (15b). Candidate (a) is Beijing, with neutralization between the /an-r/ and /a-r/ rhymes into [äŋ] (see **Figure 5.a**). In terms of notation, the neutralization output form [äŋ] is written twice, to show their one-to-one mapping with the stem rhymes. Between the two instances of [äŋ], an equality sign "=" is used instead of a dash, to emphasize that the output forms are identical. Candidate (b) is Liaoning, where

the /a-r/ rhyme surfaces as [äɪ], with a non-retroflex coda, and every contrast is kept (see **Figure 5.b**).

Candidate (c) also maintains the three-way contrast, but violates  $\text{MinDist} = \text{RD}:2$ . Here, the /an-r/ rhyme retains its nasalization, surfacing as [ä̃ɪ], and reduces its distance from the /a-r/ rhyme to a mere 1, as shown in the calculation in (16a). Candidate (d) has all three vowels maintaining their stem values in F2 and nasalization, while employing identical retroflex [-ɻ] suffix. It is ruled out by  $\text{MinDist} = \text{RD}:2$ , for a rhyme distance of 1.41 between /an-r/ and /a-r/, as shown in (16b). Candidate (e) displays complete neutralization between all three suffixed rhymes, converging at [äɪ] (see **Figure 5.e**). It is disfavored by both dialects, due to losing too many contrasts.



**Figure 5:** Candidate rhyme spaces in tableaux (15a&b).

(15) (a) Beijing:  $\text{MinDist} = \text{RD}:2 \gg \text{PreserveContrasts-BD}$

	[än-ä-ä̃ɲ]	$\text{MinDist} = \text{RD}:2$	Id[Retro]-IO	PresCont-BD
a.	[äɪ = äɪ - ä̃ɪ]		*	✓✓
b.	[äɪ - äɪ - ä̃ɪ]		**!	✓✓✓
c.	[ä̃ɪ - äɪ - ä̃ɪ]	*!	*	✓✓✓
d.	[ä̃ɪ - ä̃ɪ - ä̃ɪ]	*!		✓✓✓
e.	[äɪ = ä̃ɪ = ä̃ɪ]			✓!

(b) Liaoning: MinDist = RD:2, PreserveContrasts-BD

	$[\tilde{a}n\text{-}\tilde{a}\text{-}\tilde{a}\eta]$	MinDist = RD:2	PresCont-BD	Id[Retro]-IO
a.	$[\tilde{a}\downarrow = \tilde{a}\downarrow\text{-}\tilde{a}\downarrow]$		✓✓!	*
b.	$[\tilde{a}\downarrow\text{-}\tilde{a}\downarrow\text{-}\tilde{a}\downarrow]$		✓✓✓	**
c.	$[\tilde{a}\downarrow\text{-}\tilde{a}\downarrow\text{-}\tilde{a}\downarrow]$	*!	✓✓✓	*
d.	$[\tilde{a}\downarrow\text{-}\tilde{a}\downarrow\text{-}\tilde{a}\downarrow]$	*!	✓✓✓	
e.	$[\tilde{a}\downarrow = \tilde{a}\downarrow = \tilde{a}\downarrow]$		✓!	

(16) (a) Candidate (c) rhyme distance:

$$d(/an\text{-}r/, /a\text{-}r/) = d(\tilde{a}_{(3,1,1)}, \tilde{a}_{(3,0,1)}) = \sqrt{(3-3)^2 + (1-0)^2 + (1-1)^2} = 1$$

$$d(/a\text{-}r/, /a\eta\text{-}r/) = d(\tilde{a}_{(3,0,1)}, \tilde{a}_{(0,2,0)}) = \sqrt{(3-0)^2 + (0-2)^2 + (1-0)^2} = \sqrt{14} \approx 3.74$$

$$d(/a\eta\text{-}r/, /an\text{-}r/) = d(\tilde{a}_{(0,2,0)}, \tilde{a}_{(3,1,1)}) = \sqrt{(0-3)^2 + (2-1)^2 + (0-1)^2} = \sqrt{11} \approx 3.32$$

(b) Candidate (d) rhyme distance:

$$d(/an\text{-}r/, /a\text{-}r/) = d(\tilde{a}_{(3,1,3)}, \tilde{a}_{(3,0,2)}) = \sqrt{(3-3)^2 + (1-0)^2 + (3-2)^2} = \sqrt{2} \approx 1.41$$

$$d(/a\text{-}r/, /a\eta\text{-}r/) = d(\tilde{a}_{(3,0,2)}, \tilde{a}_{(3,2,0)}) = \sqrt{(3-3)^2 + (0-2)^2 + (2-0)^2} = \sqrt{8} \approx 2.83$$

$$d(/a\eta\text{-}r/, /an\text{-}r/) = d(\tilde{a}_{(3,2,0)}, \tilde{a}_{(3,1,3)}) = \sqrt{(3-3)^2 + (2-1)^2 + (0-3)^2} = \sqrt{10} \approx 3.16$$

In both dialects, MinDist=RD:2 is ranked high, ruling out candidate (c). The remaining two candidates, the neutralizing candidate (a) and contrast-preserving candidate (b) both satisfy MinDist. Liaoning (15b) ranks PreserveContrasts-BD as high as MinDist=RD:2, and therefore favors the contrast-preserving candidate (b) over the neutralizing candidate (a). In Beijing, however, PreserveContrasts-BD ranks below MinDist=RD:2 and Ident[Retroflex]-IO. The neutralizing candidate (a) has one less violation of Ident[Retroflex]-IO than the contrast-preserving candidate (b), and thus is chosen by Beijing as the optimal candidate. The lowly-ranked PreserveContrasts-BD does not get a say on the matter.

Beijing satisfies MinDist = RD:2 at the expense of violating PreserveContrasts-BD. This can be observed in the candidate pair (a) and (c). The two candidates are equally acceptable according to Ident[Retroflex]-IO, incurring one violation each. But MinDist = RD:2 favors the neutralizing candidate (a), whereas PreserveContrasts-BD favors candidate (c). For the dialect, MinDist = RD:2 is ranked higher, thus the neutralizing candidate (a) wins, despite not preserving all contrasts from the stem forms.

In Liaoning, on the other hand, the two constraints are equally high. Therefore, neither candidate (c) nor the neutralization candidate (a) can surface. The dialect places as much priority on preserving the three-way contrasts from the stem forms, as making sure each contrast is salient enough.

## 4.4 /r/ allophony constraints

Up until now, I have used the constraint **Ident [Retroflex] -IO** without explanation. This section provides the reasoning behind this constraint, arguing against an alternative markedness constraint account. I argue that the retroflex [-ɻ] is the default diminutive suffix. I also address why the /aŋ/ rhyme would opt for the non-default [-ɹ] as its suffix.

### 4.4.1 IO faithfulness constraint

There is overwhelming evidence suggesting that the retroflex [-ɻ] is the default diminutive suffix. First of all, the source of the diminutive suffix contains a retroflex coda. The diminutive suffix derives its meaning from the monomorphemic word [əɻ] ‘child’, spelled as *er* in pinyin. The process of diminutivization is called *erhua* in Mandarin, where *hua* means ‘-ization’. As the diminutive suffix, the schwa is deleted, leaving only the rhotic coda to attach to noun stems. Evidence that the rhotic coda in [əɻ] is indeed a retroflex can be found in Jiang et al. (2019) EMA study of Liaoning speakers. They have discovered that the rhotic coda in the monomorphemic [əɻ] is distinct from the one used alongside open-syllable stems like /a/. The rhotic coda in [əɻ] involves mainly the movement of the tongue tip, rather than the tongue body. In other words, it is a retroflex rhotic. When it comes to the rhotic diminutive suffix, it is not unreasonable to think that it ought to be subject to an input-output faithfulness constraint. It should surface as a retroflex [-ɻ] unless suggested otherwise.

A previous version of the analysis includes a markedness constraint \*ɹ to capture the default status of the retroflex suffix. In the original analysis, the retroflex [-ɻ] is less marked than the non-retroflex [-ɹ]. However, an anonymous reviewer has pointed out that there is no cross-linguistic universal that can support the claim that the retroflex rhotic is less marked than a bunched rhotic. If anything, the retroflex is more marked, since retroflex consonants are rare in world languages (Hamann 2003). It was also suggested by the same reviewer that an input-output faithfulness constraint that links up the diminutive suffix with its source morpheme [əɻ] ‘child’ can account for the pattern equally well.

### 4.4.2 BD faithfulness constraint

The relative ranking between **PreserveContrasts-BD** and **Ident [Retroflex] -IO** can account for the fact that /an-r/ and /a-r/ display different rhotic suffixes in Liaoning, but it fails to explain which rhyme uses which suffix. In other words, why do we observe a faithful retroflex [ɻ] in /an-r/ and an unfaithful non-retroflex [ɹ] in /a-r/, and not the other way around? In addition, why do /aŋ/ rhymes also opt for [ɹ]? There is no incentive for /aŋ/ rhymes to use an unfaithful rhotic suffix for contrast preservation purposes, since there is already enough distance between /aŋ/ and the other two rhymes. Interestingly, in Beijing, a dialect in which **PreserveContrasts-BD** is

ranked low, the /aŋ/ rhyme also uses the unfaithful [ɿ] as its suffix. There must be some other factor at play.

I argue that the two nasal rhymes, /an/ and /aŋ/, select their rhotic suffix in order to be faithful to the acoustic properties in their stem forms. This is true in both Liaoning and Beijing. Specifically, it is the formant transition of F1 and F2 at the end of the rhyme that /an/ and /aŋ/ rhymes pay attention to. Marilyn Chen (2000) noted that when a Mandarin low vowel is transitioning into an [n] coda, its F2 is higher than that of a low vowel transitioning into an [ŋ] coda. Conversely, the low vowel's F1 is lower when transitioning into [n] than [ŋ]. In other words, the difference between F1 and F2, or the F1-F2 gap, is larger at the vowel endpoint in [ãŋ] than that in [ãŋ].

The rhotic allophone chosen for the suffix is simply a replication of the F1 and F2 transition into the nasal codas in the stem forms. If there is a big F1-F2 gap at the vowel endpoint in the stem form, then there also needs to be a big F1-F2 gap at the end of the suffixed rhyme, and vice versa. To capture this generalization, a BD faithfulness constraint **Ident[FormantTransition]-BD** is proposed in (17). It is ranked above the input-output faithfulness constraint **Ident[Retroflex]-IO**.

- (17) **Ident[FormantTransition]-BD**: If the base form contains a vowel and a consonant coda, and if the same vowel is followed by a consonant coda in the derived form, then the formant transition from the vowel into the consonant coda in the derived form must be identical to that of the base form, even if the consonant codas are different between the two forms.

Note that the definition of **Ident[FormantTransition]-BD** restricts its application to the two nasal rhymes. The open-syllable rhyme /a/ is not subject to the constraint. This is because the BD faithfulness constraint evaluates the similarity of the vowel-to-coda formant transition between the stem form and the suffixed form. There is no coda in the open-syllable stem [ä]. Therefore, the constraint does not apply. Due to /a/'s immunity to **Ident[FormantTransition]-BD**, it is allowed to vary in its choice of rhotic suffix allophone between Beijing and Liaoning. The nasal rhymes /an/ and /aŋ/, on the other hand, are consistent in their choice of rhotic suffix across the two dialects, for **Ident[FormantTransition]-BD** is very highly-ranked, as shown in the tableaux for Beijing and Liaoning in (18).

- (18) (a) Beijing: /r/ allophony constraints:

[ãŋ-ä-ãŋ]	Id[FormTrans]-BD	MinDist = RD:2	Id[Retro]-IO	PresCont-BD
a. [äɿ = äɿ-ãɿ]	*!			✓✓
☞ b. [äɿ = äɿ-ãɿ]			*	✓✓
c. [äɿ-äɿ-ãɿ]			**!	✓✓✓
d. [äɿ = äɿ-ãɿ]	*!		***	✓✓

(b) Liaoning: /r/ allophony constraints:

	[ã̃n-ä-ã̃ŋ]	Id[FormTrans]-BD	MinDist = RD:2	PresCont-BD	Id[Retro]-IO
a.	[äɿ = äɿ-ã̃ɿ]	*!		✓✓	
b.	[äɿ = äɿ-ã̃ɿ]			✓✓!	*
☞ c.	[äɿ-äɿ-ã̃ɿ]			✓✓✓	**
d.	[äɿ = äɿ-ã̃ɿ]	*!		✓✓	***

Again, the candidates are the same for the two tableaux. Candidate (a) has all its rhymes taking the retroflex [-ɿ] suffix. It violates the highly-ranked Ident[FormantTransition]-BD, incurred by the /aŋ-r/ rhyme. Here, /aŋ-r/ has a retroflex [ɿ], which shows a rising F1 and falling F2 transition, as opposed to the stable F1 and F2 in the base form [ã̃ŋ]. Similarly, candidate (d), where every suffixed rhyme takes on a non-retroflex [-ɿ] suffix, is also ruled out. Here, the Ident[FormantTransition]-BD-violating rhyme is /an/. The base form [ã̃n] has a rising F1 and falling F2 transition, but the suffixed form [äɿ] has flat F1 and F2 transitions.

Both candidates (c) and (d) satisfy Ident[FormantTransition]-BD, for their /an-r/ rhymes and /aŋ-r/ rhymes display the appropriate rhotic suffixes that match their base form formant transitions. The two candidates also meet the MinDist = RD:2 requirement. Their difference lies in the choice of the rhotic suffix of the open-syllable /a/ rhyme. In candidate (b), /a/ takes on the retroflex [-ɿ] suffix, whereas in candidate (c), /a/ uses the non-retroflex [-ɿ]. For each dialect, the choice between the two candidates comes down to the ranking between PreserveContrasts-BD and Ident[Retroflex]-IO. Beijing ranks Ident[Retroflex]-IO above PreserveContrasts-BD. Therefore, candidate (b), which has one less [ɿ] than candidate (c), wins. This can be seen in (18a). Liaoning, on the other hand, prioritizes PreserveContrasts-BD over Ident[Retroflex]-IO. The three-way contrast candidate (c) wins, despite having more instances of [ɿ] in its output. The crucial ranking is summarized in (19).

- (19) (a) Beijing: Ident[FormantTransition]-BD » Ident[Retroflex]-IO » PreserveContrasts-BD  
 (b) Liaoning: Ident[FormantTransition]-BD, PreserveContrasts-BD » Ident[Retroflex]-IO

#### 4.4.3 The role of IO and BD faithfulness constraints

As demonstrated by the /r/ allophony constraints, contrast preservation constraints alone cannot derive the suffixed patterns of Beijing and Liaoning. IO and BD faithfulness constraints are also needed. The two contrast preservation constraints, MinDist = RD:2 and PreserveContrasts-BD, evaluate the entire suffixed rhyme space, or more precisely, the relation between rhymes. The /r/ allophony constraints, on the other hand, only focus on individual rhymes. The IO faithfulness constraint Ident[Retroflex]-IO evaluates the mapping between the suffix and its

original diminutive morpheme, whereas the faithfulness constraint  $\text{Ident}[\text{FormantTransition}]\text{-BD}$  compares a suffixed rhyme with its stem form. In Section 4.5 (vowel backing) and 4.6 (vowel denasalization), I introduce more BD faithfulness constraints.

#### 4.5 F2 constraints

The vowel backing of the front [a] and central [ä], triggered by the addition of a rhotic suffix, can be accounted for with the markedness constraint  $\text{F2} < 2/_R$ , as defined in (20a). The front vowel [a], with an F2 value of 3, violates the constraint. So does the slightly more front central vowel [ä] seen in the open-syllable stem. By stipulating that before a rhotic suffix, a vowel cannot have a value bigger than or equal to 2 on the F2 axis, the constraint explains why both the front [a] in the /an-r/ rhyme and the central [ä] in the /a-r/ rhyme undergo coarticulatory backing.

(20) F2 constraints:

- (a)  $\text{F2} < 2/_R$ : A vowel ought to have a value smaller than 2 on the F2 axis if it precedes a rhotic coda. Assign a violation if a vowel preceding a rhotic coda has an F2 value bigger than or equal to 2.
- (b)  $\text{Ident}[\text{F2}]\text{-BD}$ : The F2 value of the derived form of a segment should be the same as that of its correspondent in the base form. Assign one violation if the F2 value of the derived form of the segment is different from that of the correspondent in the base form by a degree of 1 on the F2 axis. Assign two violations if they are different by a degree of 2, etc.
- (c)  $\text{F2} < 2/_R \gg \text{Ident}[\text{F2}]\text{-BD}$

A markedness constraint is often the triggering force behind modifications in the output form that departs from the input form, or another output form. But its power needs to be checked by a counterforce: the faithfulness constraint, which requires the output form to be identical to the input form, or another output form. This is why I propose the BD faithfulness constraint  $\text{Ident}[\text{F2}]\text{-BD}$  (defined in (20b)) to ensure that the backing process triggered by  $\text{F2} < 2/_R$  does not push the vowels all the way to 0 on the F2 axis. The markedness constraint is ranked above the BD faithfulness constraint, as shown in (20c).

The relative ranking between the two F2 constraints and the contrast preservation constraints are illustrated with the tableau for Beijing in (21a), and the tableau for Liaoning in (21b). Between the two tableaux, candidates with the same letter share the same F2 values for each suffixed rhyme, differing minimally on the choice of /r/ allophone. For instance, candidate (a) of (21a) [a<sub>1</sub>-ä<sub>1</sub>-ã<sub>1</sub>] and candidate (a) of (21b) [a<sub>1</sub>-är<sub>1</sub>-ã<sub>1</sub>] have the same values on the F2 axis: [3-2-0]. **Figure 7.a** illustrates the position of each suffixed rhyme in candidate (a) on the F2 axis, with R as an abstracted rhotic suffix. Every pair of near-identical candidates in (21a) and (21b) corresponds to an F2 diagram in **Figure 7**. **Figure 6** shows the stem rhyme F2 values for comparison.

In order to focus on F2 constraints, I only include candidates that have the same nasalization pattern and choice of rhotic suffix as the winning candidate for each dialect. Only F2 values vary between candidates. Therefore, /r/ allophony constraints and nasal constraints can be omitted from the tableau.

In (21a), candidate (a) has F2 values that are completely faithful to the base forms. It incurs two violations of the markedness constraint  $F2 < 2/_R$ . Candidate (b) (Figure 7.b) has partial backing, where the front vowel in the /an/ rhyme moves from 3 to 2 on the F2 axis, and the central vowel in /a/ from 2 to 1. One violation of  $F2 < 2/_R$  is incurred by /an/. Candidate (c) has both front and central vowels moving back to the position of 1 on the F2 axis (see Figure 7.c). The requirement of  $F2 < 2/_R$  is satisfied. Candidate (d) has the central vowel in /a/ moved back one step further, showing up at 0 on the F2 axis in Figure 7.d. But this causes a violation of the equally highly ranked  $MinDist = RD:2$ , for the rhyme distance between the suffixed /an-r/ and /a-r/ rhymes is only 1. Candidate (e) violates neither  $F2 < 2/_R$  nor  $MinDist = RD:2$ , by having a back vowel [ɑ] in all three rhymes, and neutralizing the contrast between /an-r/ and /a-r/. But it violates the gradient BD faithfulness constraint  $Ident[F2]-BD$  5 times, whereas candidate (c) only causes 3 violations of this constraint. Therefore, candidate (c) is the optimal candidate. Both  $F2 < 2/_R$  and  $MinDist = RD:2$  are highly ranked, for the winning candidate cannot violate either.

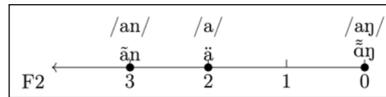
(21) F2 constraints:

(a) Beijing:

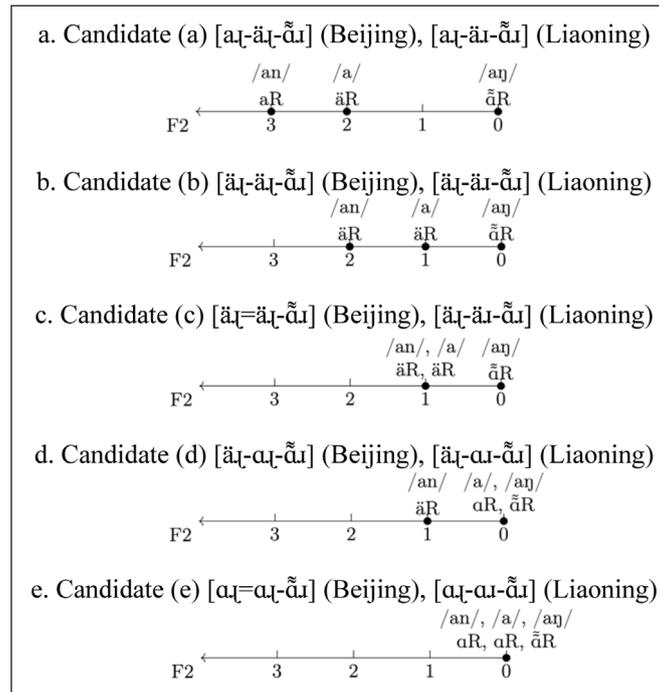
	[ān-ä-ā̃ŋ]	$F2 < 2/_R$	$MinDist = RD:2$	$Id[F2]-BD$	$PresCont-BD$
a.	[aɿ-äɿ-ā̃ɿ]	*!*	*		✓✓✓
b.	[äɿ-äɿ-ā̃ɿ]	*!	*	**	✓✓✓
☞ c.	[äɿ = äɿ-ā̃ɿ]			***	✓✓
d.	[äɿ-ɑɿ-ā̃ɿ]		*!	****	✓✓✓
e.	[ɑɿ = ɑɿ-ā̃ɿ]			****!*	✓✓

(b) Liaoning:

	[ān-ä-ā̃ŋ]	$F2 < 2/_R$	$MinDist = RD:2$	$PresCont-BD$	$Id[F2]-BD$
a.	[aɿ-äɿ-ā̃ɿ]	*!*		✓✓✓	
b.	[äɿ-äɿ-ā̃ɿ]	*!		✓✓✓	**
☞ c.	[äɿ-äɿ-ā̃ɿ]			✓✓✓	***
d.	[äɿ-ɑɿ-ā̃ɿ]			✓✓✓	****!
e.	[ɑɿ-ɑɿ-ā̃ɿ]			✓✓✓	****!*



**Figure 6:** Stem rhyme F2 values in tableaux (21a&b).



**Figure 7:** Candidate rhyme F2 values in tableaux (21a&b).

The F2 constraints derive the same vowel F2 values for Liaoning, as shown in tableau (21b). Here, the candidates are minimally different from the Beijing tableau in their choice of rhotic suffix for the open syllable rhyme, while all F2 values remain the same as the Beijing candidates. The difference in constraint ranking lies in the fact that Liaoning ranks PreserveContrasts-BD as high as  $\text{MinDist} = \text{RD}:2$ . Candidates (a) and (b) are ruled out for violations of the markedness constraint  $\text{F2} < 2/_R$ .  $\text{Ident}[\text{F2}]\text{-BD}$  stops candidates (d) and (e) from surfacing, both of which have incurred too many violations of the BD faithfulness constraint. Candidate (c), where both the front vowel and the central vowel are backed into the position of 1 on the F2 axis, wins, just like they do in Beijing.

#### 4.6 Nasal constraints

As previewed in Section 2.1, I argue that vowel denasalization is not motivated by the markedness constraint  $*V_{\text{nas}}$ , but by contrast preservation. In the stem forms, the varying degrees of nasalization seen in the two rhymes  $[\ã\text{n}]$  and  $[\ã\text{ŋ}̃]$  serve to enhance the contrast in place of articulation of the nasal coda. But in the suffixed forms, the nasalization trigger is deleted for both rhymes. What

is left are two nasalized low vowels that still need to maintain a salient contrast between each other. However, for the listener, perceiving the difference between the two nasalized low vowels is a difficult task. There are two reasons for this.

First of all, the difference in degree of nasalization between \*[ã̃] and [ã̄] is simply not big enough to be perceptually salient. It is true that two degrees of nasalization are licensed in the stem forms, but they are a byproduct of a contrast elsewhere, namely the nasal coda. The degree of nasalization cannot manifest a contrast on its own, when the nasal coda is deleted. Therefore, only one degree of nasalization is allowed in a context without nasal codas.

A second problem with having \*[ã̃] and [ã̄] in the suffixed forms is that the nasalization would get in the way of the F2 contrast. With backing of the front vowel, the F2 distance between the two vowels is already reduced from 3 to 1 on the F2 axis. It is universally more difficult for a speaker to distinguish between two nasalized vowels than between two oral ones (see Bond 1976; Wright 1986; Krakow et al. 1988; Beddor 1993; Stanton 2018). Retaining the nasalization in both vowels makes it harder to maintain an F2 contrast between them.

To capture the objective of the language to maximize the contrast in degree of nasalization, I include a single-dimension MinDist constraint in my analysis. This is **MinDist = VND:2**, as defined in (22a).

- (22) (a) **MinDist = VND:2**: The Euclidean distance between two vowels without a nasalization trigger must be bigger than or equal to 2 along the vowel nasalization axis. (VND stands for Vowel Nasalization Distance.)
- (b) **Ident[Nasal]-BD**: The vowel nasalization value of the derived form should be the same as that of the base form. Assign one violation if the vowel nasalization value of the derived form is different from that of the base form by a degree of 1 on the vowel nasalization axis. Assign two violations if they are different by a degree of 2, etc.
- (c) **MinDist = VND:2 » Ident[Nasal]-BD**

The suffixed rhymes \*[ã̃-ä̃-ã̄] (/an-a-aŋ/) violate **MinDist = VND:2**, for their distance along the vowel nasalization axis is a mere 1 between the pair \*[ã̃-ä̃] and between the pair \*[ã̃-ã̄]. This is shown in the tableau for Beijing in (23a), as candidate (a).

A gradient BD faithfulness constraint for the nasalization axis (22b) is also included. **Ident[Nasal]-BD** works in a similar fashion as **Ident[F2]-BD**. It helps rule out candidate (c) of (23a), where instead of denasalizing the vowel in the /an-r/ rhyme, it denasalizes the vowel in the /aŋ-r/ rhyme and raises the degree of nasalization in /an-r/. While satisfying **MinDist = VND:2**, candidate (c) causes 2 more violations of **Ident[Nasal]-BD** than the winning candidate (b).

In (23a), I only include candidates with the optimal F2 values and rhotic suffixes, in order to put the limelight on nasal constraints. A diagram of the vowel nasalization dimension of each candidate is illustrated in **Figure 9** (also applicable to (23b)), so that the evaluation of the nasal constraints is more transparent to the reader. **Figure 8** shows the vowel nasalization dimension of the stem rhymes.

(23b) shows the tableau for Liaoning, where the distribution of /r/ allophony in each of the candidates is adjusted for the dialect. The open-syllable stem /a/ takes a non-retroflex [ɿ] as suffix. Liaoning also differs from Beijing in its ranking of the PreserveContrasts-BD constraint. In Liaoning, it is ranked as high as MinDist = VND:2 and MinDist = RD:2, ruling out any neutralization candidate. In addition, candidate (c) is ruled out by a different constraint in Liaoning than in Beijing. It incurs a fatal violation of MinDist = RD:2, due to the small rhyme distance of 1 between [äɿ]<sub>(0,0,1)</sub> (/a/) and \*[ɑɿ]<sub>(0,0,0)</sub> (/aŋ/).

(23) Nasal Constraints:

(a) Beijing:

[ān-ä-āŋ]	MinDist = VND:2	MinDist = RD:2	Id[Nasal]-BD	PresCont-BD
a. [ä̃ɿ-äɿ-āŋ]	*!*			✓✓✓
<sup>13</sup> b. [äɿ=äɿ-āŋ]			*	✓✓
c. [ä̃ɿ-äɿ-ɑɿ]			**!*	✓✓✓

(b) Liaoning:

[ān-ä-āŋ]	MinDist = VND:2	MinDist = RD:2	PresCont-BD	Id[Nasal]-BD
a. [ä̃ɿ-äɿ-āŋ]	*!*		✓✓✓	
<sup>13</sup> b. [äɿ-äɿ-āŋ]			✓✓✓	*
c. [ä̃ɿ-äɿ-ɑɿ]		*!	✓✓✓	***

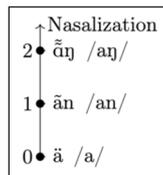


Figure 8: Stem rhyme nasalization values in tableaux (23a&b).

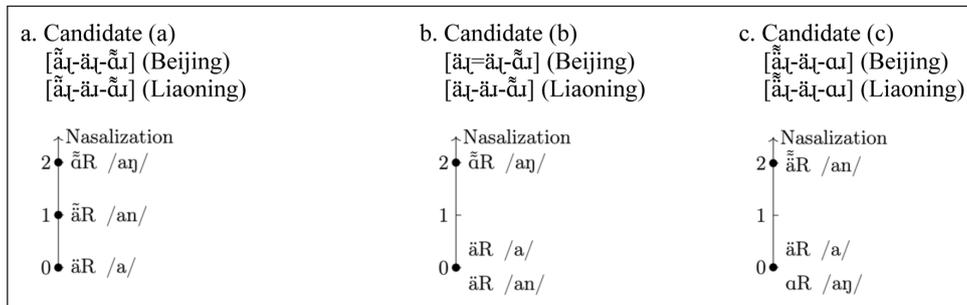


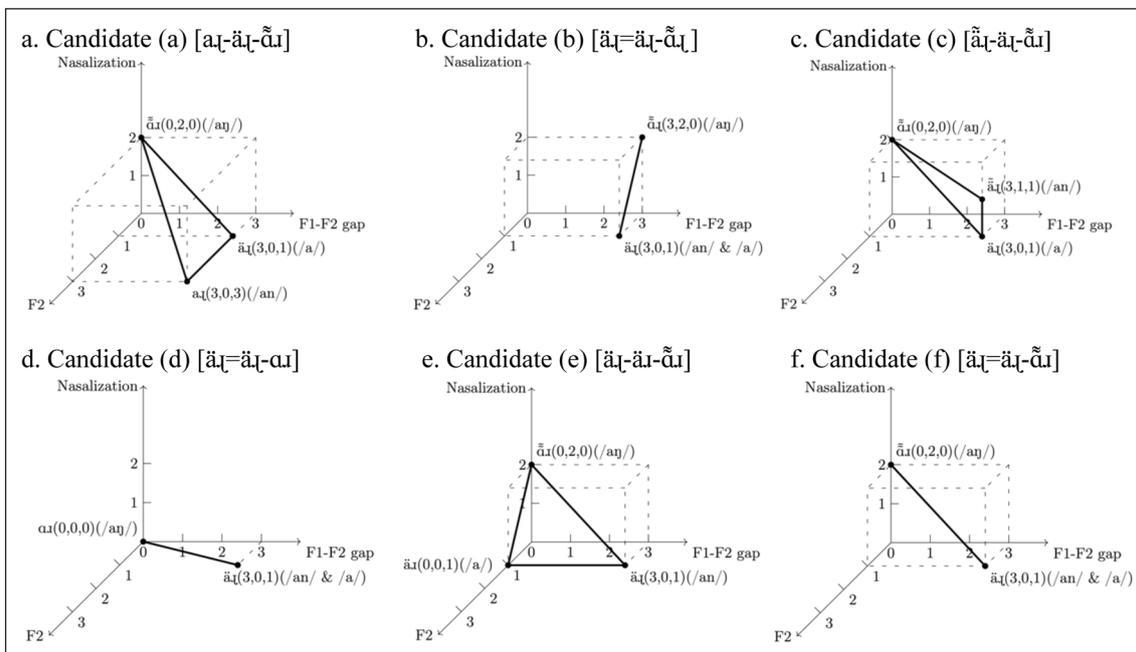
Figure 9: Candidate rhyme nasalization values in tableaux (23a&b).

Note that stem forms are not subject to  $\text{MinDist}=\text{VND}:2$ , due to the presence of nasalization triggers. Therefore, varying degrees of nasalization are licensed in the stem forms, allowing  $[\tilde{a}\tilde{n}]$  and  $[\tilde{a}\tilde{\eta}]$  to surface at the same time.

The full ranking of all constraints in Beijing is shown in (24), with candidate rhyme spaces illustrated in **Figure 10**. (25) displays the full ranking in Liaoning, accompanied by candidate rhyme spaces in **Figure 11**.

(24) Beijing full constraint ranking:

$[\tilde{a}\tilde{n}-\tilde{a}-\tilde{a}\tilde{\eta}]$	$\text{F2} < 2$ /_R	$\text{Id}[\text{Form}$ $\text{Trans}]-\text{BD}$	$\text{MinDist}$ $=\text{VND}:2$	$\text{MinDist}$ $=\text{RD}:2$	$\text{Id}[\text{Nas}]$ -BD	$\text{Id}[\text{Retro}]$ - IO	$\text{Id}[\text{F2}]$ -BD	$\text{PresCont}$ -BD
a. $[\text{a}_{\downarrow}-\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}]$	*!				*	*	*	✓✓✓✓
b. $[\tilde{\text{a}}_{\downarrow}=\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}]$		*!			*		***	✓✓
c. $[\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}]$			*!*	*		*	***	✓✓✓✓
d. $[\tilde{\text{a}}_{\downarrow}=\tilde{\text{a}}_{\downarrow}-\text{a}_{\downarrow}]$					***	*	**!* *	✓✓
e. $[\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}]$					*	**	***!	✓✓✓✓
f. $[\tilde{\text{a}}_{\downarrow}=\tilde{\text{a}}_{\downarrow}-\tilde{\text{a}}_{\downarrow}]$					*	*	***	✓✓



**Figure 10:** Candidate rhyme spaces in tableau (24).

(25) Liaoning full constraint ranking:

$[\tilde{a}n-\tilde{a}-\tilde{a}ŋ]$	F2 < 2 /_R	Id[Form Trans]-BD	MinDist = VND:2	MinDist = RD:2	PresCont -BD	Id[Nas] -BD	Id[Retro]- IO	Id[F2] -BD
a. $[a_{ɹ}-\tilde{a}ɹ-\tilde{a}ɹ]$	*!				✓✓✓	*	**	*
b. $[\tilde{a}_{ɹ}-\tilde{a}ɹ-\tilde{a}ɹ]$		*!			✓✓✓	*	*	***
c. $[\tilde{a}_{ɹ}-\tilde{a}ɹ-\tilde{a}ɹ]$			*!*		✓✓✓		**	***
d. $[\tilde{a}_{ɹ}-\tilde{a}ɹ-a_{ɹ}]$				*!	✓✓✓	***	**	***
e. $[\tilde{a}_{ɹ}-\tilde{a}ɹ-\tilde{a}ɹ]$					✓✓✓	*	**	***
f. $[\tilde{a}_{ɹ} = \tilde{a}_{ɹ}-\tilde{a}ɹ]$					✓✓!	*	*	***

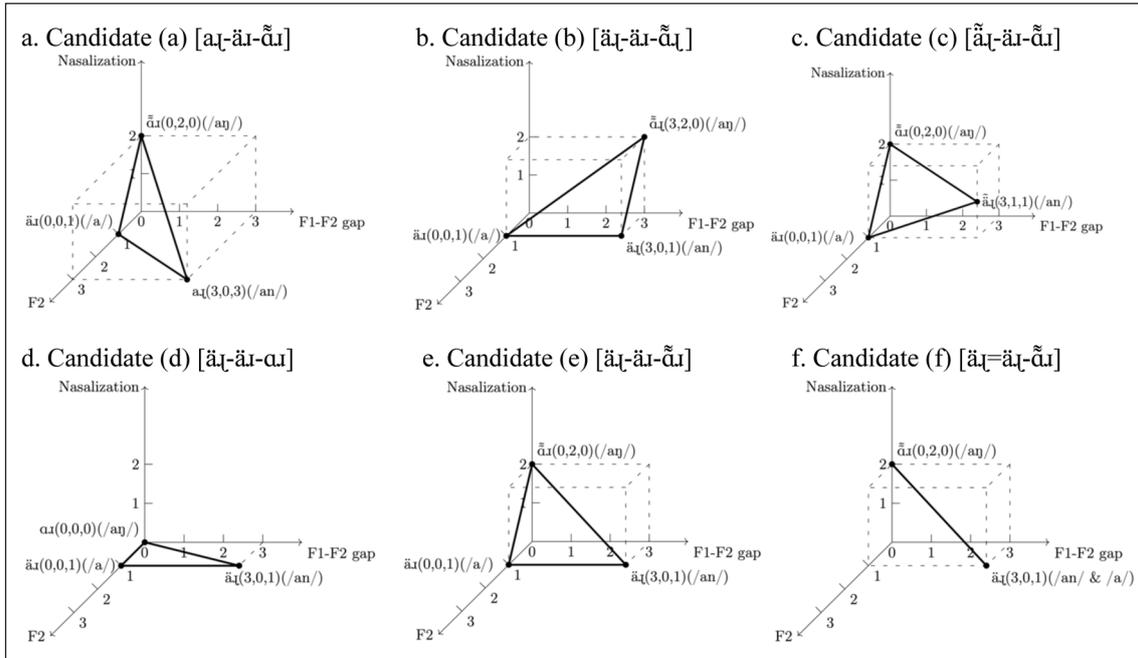


Figure 11: Candidate rhyme spaces in tableau (25).

## 5. Discussion

### 5.1 The role of BD faithfulness

In Flemming’s (1995) original description of Dispersion Theory, faithfulness constraints have no place in the framework. But in this study, it can be seen that BD faithfulness constraints play an integral part in deriving the distribution of the rhotic suffix allophones in the two

dialects of Mandarin. The BD faithfulness constraint `Ident[FormantTransition]-BD` works alongside the contrast preservation constraints, to dictate which rhotic allophone the nasal rhymes /an/ and /aŋ/ should use. BD faithfulness constraints are not necessarily incompatible with a contrast preservation analysis. In Liaoning, the suffixed /a/ rhyme is instructed by `MinDist=RD:2` and `PreserveContrasts-BD` to enhance its contrast from /an/ by resorting to a different rhotic suffix. But without a BD faithfulness constraint dictating which rhotic allophone /an/ should take, /a/, solely with the goal to be distinct from /an/, would have received no instruction on which rhotic allophone itself ought to use. In both Beijing and Liaoning, the BD faithfulness constraint `Ident[FormantTransition]` requires the /aŋ/ rhyme to use a non-retroflex rhotic suffix, so that the suffixed form is more faithful to the stem form. BD faithfulness constraints play an important role in deriving the allophonic distribution of the rhotic suffix, and are therefore not at odds with the overall objective of contrast preservation.

Interestingly, `PreserveContrasts-BD` is itself a BD faithfulness constraint. Unlike Flemming's `MaximizeContrasts`, it does not generate a set of contrastive phonemes to form an inventory, but draws upon the contrasts already existent in a set of base forms, making sure that the contrasts are still present in the derived forms. `PreserveContrast-BD` evaluates the mapping between base forms and derived forms, and can therefore be considered a BD faithfulness constraint. But unlike a regular BD faithfulness constraint, where the similarity between a specific base form and its derived form are evaluated, `PreserveContrasts-BD` is not concerned with the specific segments or features in the derived form. Instead, it requires that the number of elements in the set of derived forms match the number of elements in the set of base forms. In the example of low vowel rhymes in Mandarin, the set of base forms contains 3 elements: [ã], [ä], [ãŋ]. `PreserveContrasts-BD` requires that there be 3 elements in the set of suffixed forms as well. As to how the derived forms individually map onto the base forms, `PreserveContrasts-BD` has no say in the matter.

## 5.2 Alternative analyses

The two key ingredients of my contrast preservation analysis are contrast enhancement and output-output mapping. I argue that neither are dispensable. In this section, I consider three alternative analyses. The first alternative analysis is one without reference to either contrast enhancement or output-output mapping. The second alternative analysis considered has elements of contrast enhancement, but no output-output mapping. The third alternative analysis contains output-output mapping, but no contrast enhancement. I provide arguments against each of the three analyses, and show that both contrast enhancement and output-output mapping are crucial to the analysis of Mandarin r-suffixation.

### 5.2.1 Neither contrast enhancement nor output-output mapping

The first alternative analysis is one that derives the suffixed forms by phonotactics, or markedness constraints, by evaluating the suffixed forms alone, without any reference to the stem forms. In this analysis, the two neighboring segments, the vowel and the rhotic coda, have co-occurrence restrictions. The choice of rhotic coda should be predictable from the quality of the vowel. A back vowel [ɑ] dictates that it should be followed by a non-retroflex [ɹ]. The central vowel [ä] should have a preference for which rhotic allophone it can appear next to as well. But as we see in Liaoning, [ä] can either be followed by a retroflex coda in [äɻ] (/an/) or a non-retroflex coda in [äɹ] (/a/). Phonotactics of the surface forms alone cannot predict the pattern of Liaoning.

### 5.2.2 Contrast enhancement without output-output mapping

Another possibility to consider is that the suffixed forms can be accounted for by contrast enhancement on the suffixed forms alone, without any mapping to the stem forms. In this analysis, the suffixed forms are compared to the input forms /an-r/, /a-r/, and /aŋ-r/ only. The /-r/ suffix would only surface as [-ɹ] if it is required to enhance a contrast; otherwise it appears as the faithful [-ɻ]. This analysis can predict the behavior of /a-r/ in Liaoning. /a-r/ surfaces with [-ɹ] to enhance its rhyme distance from /an-r/. However, there would be no explanation as to why in both dialects, /aŋ-r/ would also surface with [-ɹ] as suffix. There is no incentive for /aŋ-r/ to enhance the contrast between itself and other rhymes, since by virtue of having a nasalized back vowel, it is already quite far from the other two rhymes. Without output-output mapping, the suffixed form [äɹ] (/aŋ-r/) is inexplicable.

### 5.2.3 Output-output mapping without contrast enhancement

A third alternative analysis is one that only makes reference to output-output mapping, but does so without the goal of contrast enhancement. This would mean that the suffixed form is predictable from its stem form, using a combination of IO and BD faithfulness constraints. For instance, the open-syllable suffixed rhyme /a-r/ ought to be predictable from its stem form. In both dialects, the stem form is [ä]. It should follow that the suffixed form is identical between the two dialects as well. But this is not what we observe. Beijing has [äɻ] while Liaoning shows [äɹ]. The divergent suffixed forms cannot be derived from the same set of IO and BD faithfulness constraints. Instead, their difference lies in the different ranking between contrast preservation constraints.

## 5.3 Diachronic sound change

Beijing and Liaoning are two concurrent dialects of Mandarin that display different r-suffixation patterns. There is reason to believe that they might represent two stages in diachronic sound

change as well. Wang & He (1985) reported a generational difference among Beijing speakers. According to them, older speakers tend to have the contrast between /an/ and /a/ after r-suffixation, whereas younger speakers at the time have neutralized the two suffixed rhymes. The observation was supported by a perceptual study conducted on 50 speakers under the age of 35, who overwhelmingly could not distinguish between /an-r/ and /a-r/, with a confusion rate at 96.4%. However, the same group of young speakers had less trouble distinguishing between the mid vowel rhymes /ən-r/ and /ə-r/, displaying varying degree of identification rate. Wang & He suspected the neutralization was a sound change in progress, that somehow affected low vowel rhymes more than mid vowel rhymes.

Wang & He (1985) did not provide any acoustic data for the older speakers, so it is hard to see how they realized the contrast between /an-r/ and /a-r/. It is possible that the older speakers of Beijing in the 1980s had a similar r-suffixation pattern as present-day Liaoning speakers. Exactly how the sound change took place is an interesting question for future research. One relevant study is the *Bidirectional Phonetics and Phonology* learning algorithm proposed by Boersma & Hamann (2008). The algorithm uses the same set of gradient constraints in the process of perception and production. It is capable of modeling the non-teleological sound changes across generations that eventually arrives at a stable state of dispersion. The learner can predict the sound changes of the three Polish sibilants across centuries. However, Boersma & Hamann admit that their learner cannot replicate the process of contrast neutralization across generations. Therefore, it is not capable of modeling the recent sound change in Beijing that results in contrast neutralization.

#### 5.4 High vowels and diphthongs

The high vowel rhymes suffixation patterns reported by Chao (1968), Wang & He (1985), Duanmu (2007), and Lin (2007) for Standard Mandarin are similar to those of the low vowel rhymes in Beijing. The alveolar nasal rhyme /Vn/ neutralizes with the open-syllable rhyme /V/, while the velar nasal rhyme /Vŋ/ maintains contrast with some degree of nasalization. The /i/ and /in/ rhymes neutralize as [iər], whereas /iŋ/ is realized as [iōr]. /y/ and /yn/ neutralize as [yər]. /u/ becomes [ur], while /uŋ/ becomes [ūr] (or [uŋ<sup>r</sup>] according to Duanmu 2007).

The low vowel diphthong rhymes: /aj/ and /aw/ also display interesting suffixation patterns. /aj/ is reported to have neutralized with /an/ and /a/ in Standard Mandarin, surfacing as [ar]. Suffixed /aw/, on the other hand, is variably transcribed as [aur] (Chao 1968; Wang & He 1985) or [au<sup>r</sup>] (Duanmu 2007; Lin 2007). The rounding feature of the offglide is retained in the suffixed form.

The neutralization facts reported about Standard Mandarin, or Beijing, is not surprising, given the dialect ranks MinDist=RD:2 above PreserveContrasts-BD. It remains to be seen how

Liaoning realizes the high vowel and diphthong rhymes in the suffixed forms. It is possible that /r/ allophony and other additional acoustic cues are recruited to preserve the contrasts after suffixation. More production and perception data from both Beijing and Liaoning speakers need to be collected, in order to extend the contrast preservation analysis to account for these rhymes. I leave this as an open question for future research.

## 6. Conclusion

This paper argues for a contrast preservation analysis of Mandarin r-suffixation. The low vowel rhymes, /an/, /a/, /aŋ/, are examined for two dialects: Liaoning and Beijing. The overarching objective for both dialects is to maintain salient contrasts between the suffixed rhymes. This is manifested in a constraint dictating the minimal rhyme distance between each pair of suffixed rhymes:  $\text{MinDist} = \text{RD}:2$ . The two dialects differ in whether they prioritize preserving the three-way contrasts from the stem forms in the suffixed forms. Liaoning places a top priority on preserving the three-way contrasts, ranking *PreserveContrasts-BD* alongside  $\text{MinDist} = \text{RD}:2$ , whereas Beijing ranks *PreserveContrasts-BD* below  $\text{MinDist} = \text{RD}:2$ . When they are confronted with the possibility of neutralization between /an-r/ and /a-r/ in the suffixed forms, the two dialects make different choices. Liaoning averts the crisis by resorting to /r/ allophony, using the retroflex [-ɻ] for /an-r/ and a non-retroflex [-ɹ] for /a-r/. Beijing, on the other hand, is happy to undergo /an-r/-/a-r/ neutralization, because preserving the three-way contrast is not a priority for the dialect.

---

## Ethics and consent

The experiment and data collection method in this paper was approved by MIT's Committee on the Use of Humans as Experimental Subjects (COUHES), in accordance with protocol 0902003098.

## Acknowledgements

Many thanks to Adam Albright, Edward Flemming, Michael Kenstowicz, and Donca Steriade for valuable discussions. This paper also benefited from audience feedback at the 27th Manchester Phonology Meeting and MIT Phonology Circle. I am equally grateful for the comments and suggestions made by the anonymous reviewers. All remaining mistakes are my own.

## Competing interests

The author has no competing interests to declare.

---

## References

- Beddor, Patrice. 1993. The perception of nasal vowels. In Huffman, Marie K. & Krakow, Rena A. (eds.), *Nasals, nasalization, and the velum*, 171–196. San Diego: Academic Press. DOI: <https://doi.org/10.1016/B978-0-12-360380-7.50011-9>
- Boersma, Paul & Hamann, Silke. 2008. The evolution of auditory dispersion in bidirectional constraint grammars. *Phonology* 25. 217–270. DOI: <https://doi.org/10.1017/S0952675708001474>
- Boersma, Paul & Weenink, David. 2017. *Praat: Doing phonetics by computer*. Version 6.0.24. (<http://www.praat.org/>) (Accessed 2017-01-24.)
- Bond, Dzintra S. 1976. Identification of vowels excerpted from neutral and nasal contexts. *Journal of the Acoustical Society of America* 59. 1229–1232. DOI: <https://doi.org/10.1121/1.380988>
- Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chen, Marilyn Y. 2000. Acoustic analysis of simple vowels preceding a nasal in Standard Chinese. *Journal of Phonetics* 28(1). 43–67. DOI: <https://doi.org/10.1006/jpho.2000.0106>
- Chen, Shuwen & Mok, Peggy Pik Ki & Tiede, Mark & Chen, Wei-rong & Walen, Douglas H. 2017. Investigating the production of Mandarin rhotics using ultrasound imaging. In *Book of Abstracts, Ultrafest VIII*, 17–18. Potsdam.
- Cheng, Ming-Chung. 2014. An optimality-theoretical exploration of retroflex diminutives in the Nanjing dialect. *Journal of National Taiwan Normal University* 59(2). 135–171.
- Chung, Hyunju & Pollock, Karen E. 2014. Acoustic characteristics of adults' rhotic monophthongs and diphthongs. *Communication Sciences & Disorders* 19(1). 113–119. DOI: <https://doi.org/10.12963/csd.13088>
- Da, Jun. 1996. A constraint-based approach to the chameleon /r/ in Mandarin dialects. *North East Linguistics Society* 26. 57–70.

- Delattre, Pierre & Freeman, Donald C. 1968. A dialect survey of American r's by X-ray motion picture. *Linguistics* 44(6). 29–68. DOI: <https://doi.org/10.1515/ling.1968.6.44.29>
- Dmitrieva, Olga. 2012. Contrast dispersion and the positional typology of geminates. *The 20th Manchester Phonology Meeting*. Manchester.
- Duanmu, San. 2007. *The phonology of Standard Chinese*. 2nd edition. Oxford: Oxford University Press.
- Faytak, Matthew & Liu, Suyuan & Sundara, Megha. 2020. Nasal coda neutralization in Shanghai Mandarin: Articulatory and perceptual evidence. *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 11(1). 23. DOI: <https://doi.org/10.5334/labphon.269>
- Flemming, Edward. 1995. *Auditory representations in phonology*. Los Angeles: University of California dissertation.
- Hamann, Silke. 2003. *The phonetics and phonology of retroflexes*. Utrecht: Utrecht University dissertation.
- Harding, Sue & Meyer, Georg. 2003. Changes in the perception of synthetic nasal consonants as a result of vowel formant manipulations. *Speech Communications* 39(3–4). 173–189. DOI: [https://doi.org/10.1016/S0167-6393\(02\)00014-6](https://doi.org/10.1016/S0167-6393(02)00014-6)
- Hsieh, Feng-fan & Jiang, Song & Chang, Yueh-chin. 2019. *A cross-dialectal comparison of er-suffixation in Beijing Mandarin and Northeastern Mandarin: An electromagnetic articulography study*.
- Jiang, Song & Chang, Yueh-chin & Hsieh, Feng-fan. 2019. An EMA study of er-suffixation in Northeastern Mandarin monophthongs. In Calhoun, Sasha & Escudero, Paola & Tabain, Marija & Warren, Paul (eds.), *Proceedings of the 19th International Congress of Phonetic Sciences*. Melbourne, Australia. 2149–2153. Canberra: Australian Speech Science and Technology Association Inc.
- Krakov, Rena A. & Beddor, Patrice S. & Goldstein, Louis M. & Fowler, Carol A. 1988. Coarticulatory influences on the perceived height of nasalized vowels. *Journal of the Acoustical Society of America* 83(3). 1146–1158. DOI: <https://doi.org/10.1121/1.396059>
- Lee, Wai-Sum. 2005. A phonetic study of the “er-hua” rimes in Beijing Mandarin. In *INTERSPEECH 2005*, 1093–1096. DOI: <https://doi.org/10.21437/Interspeech.2005-433>
- Li, Jian & Cheng, Le. 2014. The acoustics properties of the nasals and nasalization in Standard Chinese. *Information Technology Journal* 13(11). 1793–1799. DOI: <https://doi.org/10.3923/itj.2014.1793.1799>
- Lieberman, Alvin M. & Delattre, Pierre C. & Cooper, Franklin S. & Gerstman, Louis J. 1954. The role of consonant–vowel transitions in the perception of the stop and nasal consonants. *Psychological Monographs: General and Applied* 68(8). 1–13. DOI: <https://doi.org/10.1037/h0093673>
- Lilienthal, Janine. 2009. The articulatory and acoustic impact of Scottish English /r/ on the preceding vowel-onset. In *INTERSPEECH 2009*, 2819–2822. DOI: <https://doi.org/10.21437/Interspeech.2009-720>
- Lin, Yen-Hwei. 1993. Degenerative affixes and templatic constraints: Rhyme change in Chinese. *Language* 69(4). 649–682. DOI: <https://doi.org/10.2307/416882>

- Lin, Yen-Hwei. 2007. *The sounds of Chinese*. Cambridge: Cambridge University Press.
- Liu, Anqi. 2017. An acoustic study of Mandarin rhotic suffix. In Botinis, Antonis (ed.), *Proceedings of 8th Tutorial and Research Workshop on Experimental Linguistics*. Heraklion, Crete, Greece. 69–72.
- Luo, Mingqiong. 2016. A perceptually-based approach to Chinese syllable-tone patterning. In *Proceedings of Speech Prosody 2016*, 99–103. DOI: <https://doi.org/10.21437/SpeechProsody.2016-21>
- Magri, Giorgio & Storme, Benjamin. 2019. Constraint summation in phonological theory. In Baek, Hyunah & Takahashi, Chikako & Yeung, Alex Hong-Lun (eds.), *Proceedings of the 2019 Annual Meeting on Phonology*. DOI: <https://doi.org/10.3765/amp.v8i0.4673>
- Malécot, André. 1956. Acoustic cues for nasal consonants: An experimental study involving a tape-splicing technique. *Language* 32(2). 274–284. DOI: <https://doi.org/10.2307/411004>
- Petersen, Stacy. 2016. Vowel dispersion in English diphthongs: Evidence from adult production. In Hansson, Gunnar Ólafur & Farris-Trimble, Ashley & McMullin, Kevin & Pulleyblank, Douglas (eds.), *Proceedings of the 2015 Annual Meeting on Phonology*. DOI: <https://doi.org/10.3765/amp.v3i0.3680>
- R Core Team. 2018. *A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. (<http://www.R-project.org/>)
- Recasens, Daniel. 1983. Place cues for nasal consonants with special reference to Catalan. *Journal of the Acoustical Society of America* 73(4). 1346–1353. DOI: <https://doi.org/10.1121/1.389238>
- Stanton, Juliet. 2018. Environmental shielding is contrast preservation. *Phonology* 35(1). 39–78. DOI: <https://doi.org/10.1017/S0952675717000379>
- Tan, Song. 2018. La dynamique de la rhotacisation en mandarin du Nord-Est. [The dynamic of the rhotacization in northeastern Mandarin.] *La Linguistique* 54(2). 127–144. DOI: <https://doi.org/10.3917/ling.542.0127>
- Tian, Jun. 2009. An optimality theory analysis of diminutive suffixation of Beijing Chinese. In *Proceedings of the 23rd Northwest Linguistics Conference*, 217–231.
- Tunley, Alison. 1999. *Coarticulatory influences of liquids on vowels in English*. Cambridge: University of Cambridge dissertation.
- Varis, Erika. 2011. Vowel hiatus and dispersion theory. *Linguistic Society of America Annual Meeting Extended Abstracts* vol. 2. DOI: <https://doi.org/10.3765/exabs.v0i0.570>
- Wang, Lijia & He, Ningji. 1985. Beijinghua erhuayun de tingbian shiyan he shengxue fenxi. [The perceptual and acoustic studies of Beijing ‘er’ rhymes.] In Lin, Tao & Wang, Lijia (eds.), *Beijing yuyin shiyan lu*. [A collection of phonetic studies on Beijing Chinese.] Beijing: Beijing University Press. 27–71.
- Wang, Sheng-fu. 2017. A dispersion theoretic account of Taiwanese CV phonotactics. In Jesney, Karen & O’Hara, Charlie & Smith, Caitlin & Walker, Rachel (eds.), *Proceedings of the 2016 Annual Meeting on Phonology*. DOI: <https://doi.org/10.3765/amp.v4i0.3982>
- Wright, James. 1986. The behavior of nasalized vowels in the perceptual vowel space. In Ohala, John J. & Jaeger, Jeri J. (eds.), *Experimental phonology*, 45–67. Orlando: Academic Press.

Xu, Zhongshi. 2020. *Choosing rhotacization site in Beijing Mandarin: The role of perceptual similarity*. Los Angeles: University of California, Los Angeles MA thesis.

Zee, Eric & Lee, Wai-Sum. 2001. An acoustical analysis of the vowels in Beijing Mandarin. In *Seventh European Conference on Speech Communication and Technology*, 643–646. DOI: <https://doi.org/10.21437/Eurospeech.2001-169>

Zhang, Jie. 2000. Non-contrastive features and categorical patterning in Chinese diminutive suffixation: Max[F] or Ident[F]? *Phonology* 17(3). 427–478. DOI: <https://doi.org/10.1017/S0952675701003979>

