

Sonority bias in Rugao di-syllabic syllable contraction

Chenchen Xu, Yen-Hwei Lin & Karthik Durvasula*

Abstract. Two analyses, vowel sonority and the linear order of pre-contraction vowels, have been proposed to account for the vowel selection between two competing vowels in Chinese syllable contraction. An experiment was run to test whether sonority and/or linear order bias the vowel selection in Rugao syllable contraction. Our results confirmed the role of vowel sonority, and did not present supporting evidence for the linear order analysis. Sonority hierarchies along the dimensions of both height and centrality exhibit the same consistent and robust pattern, providing a new perspective to look at competing vowels in vowel-related phonological processes.

Keywords. vowel sonority; syllable contraction; vowel height; vowel centrality

1. Introduction. Syllable contraction, also called syllable merger or syllable fusion, is the process by which consecutive syllables are merged into one syllable that has the segments from both; tone contraction happens along with the segment contraction process (Lin 2007). For example, in Standard Mandarin, the di-syllabic word [pu³⁵. jaw⁵¹]¹ ('do not') can be contracted into a single syllable [pjaw⁵¹] in casual speech. Syllable contraction is documented for many varieties of Chinese including Taiwan Mandarin (Chung 1997, Lin 2007), Taiwanese Southern Min (Chung 1996, Hsu 2003), Hakka (Chung 1997), Cantonese (Wong 2006), Tianjin (Wee et al. 2005) and Jianghuai (Xu 2014). This paper investigates syllable contraction data from speakers of Rugao, a dialect of the Lower Yangzi Mandarin group (also known as Jianghuai) (Li & Thompson 1989:4), spoken in Rugao, Jiangsu Province, China. The focus will be on segmental contraction of di-syllabic tokens with clear consonantal left and right edges. The tonal contraction patterns will be discussed in a separate paper.

Chinese syllable contraction happens mostly to di-syllabic words in casual, connected speech (Tseng 2005). The structures of the involved syllables and the realization of the contracted syllable may vary, and such variation carries even within speakers. The same speaker in different speech rates can contract the same word to different degrees, and the preserved segments in the contracted output may vary (Wong 2006, Cheng & Xu 2009). Lexical frequency (Myers & Li 2009) and morpheme boundaries (Chung 1997) also play a role in shaping the output of the contracted syllables. The forms of the contracted output also follow some general phonological rules such as glide formation, segmental coalescence, and phonotactic adjustments.

According to Chung's (1996) skeletal tier account, cases of di-syllabic contraction can be viewed as six slots merging into three slots (Figure 1), assuming that Chinese syllables have a basic three-slot template XXX. In such a template, the vowel nucleus takes the middle slot. In this templatic account, the association of segments to the template starts with the two edges: when two syllables are to be contracted, the leftmost and rightmost segments of the di-syllabic word must be preserved in the contracted output. This so-called *Edge-in Association* is based on the *Edge-in Effect* proposed by Yip (1988). *Edge-in Association* has been observed for syllable

* We would like to thank the MSU Phonology & Phonetics Group, and reviewers and audiences at the 2018 LSA meeting. Authors: Chenchen Xu, Michigan State University (xuchenc1@msu.edu), Yen-Hwei Lin, Michigan State University (liny@msu.edu), Karthik Durvasula, Michigan State University (durvasul@msu.edu).

¹ The numbers indicate tones on a 1-5 scale, with 1 for the lowest and 5 the highest pitch.

contraction in many other Chinese languages including Standard Mandarin (Cheng & Xu 2009), Cantonese (Wong 2006) and Jianghuai (Xu 2014).

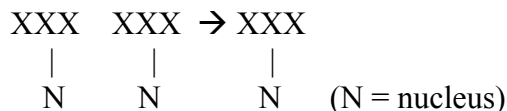


Figure 1. Syllable contraction (Chung 1996)

In di-syllabic contraction, two vowels compete for one nucleus position after the *Edge-in Association* process. With regards to which vowel is more likely to be preserved in the contracted output, Chung (1996) proposed the LR-scanning analysis, which suggests that the vowel on the left gets selected, e.g. [hɔ + laŋ] → [hɔŋ] (‘by someone’) in Taiwanese Southern Min. Sun (2014) provides more data from other dialects of Chinese that supports the order-based vowel selection, but adds that the retained vowel can be the vowel on the left or right, depending on the language. By contrast, Hsu (2003) proposed a sonority-based analysis, also based on a template and the *Edge-in Association*, for the contractions in Taiwanese Southern Min. She suggested that the vowel of higher sonority is preserved in the contracted surface form. She adopted the universal sonority hierarchy proposed by Kiparsky (1979) and argued that the sonority sequence in Taiwanese Southern Min is a > e > o > i > u. Taking [sjo + kaŋ → sjaŋ] (‘the same’) as an example, the more sonorous vowel [a] is chosen as the nucleus.

According to the word list of Taiwanese Southern Min syllable contraction² in Hsu (2003), the LR-scanning analysis is compatible with 61% of the data, while the sonority-based model can account for 86% of them. Meanwhile, even though the sonority-based model seems to be a better account, it faces its own challenges. For example, in [hɔ + laŋ] → [hɔŋ] (‘by someone’), the less sonorous vowel [ɔ] was preserved in the contracted output.

Assuming that Rugao has a basic C(G)V(X) structure³ (C = consonant, G = glide, V = vowel, X = C or G) and treating the pre-nucleus and post-nucleus vocalic segments in this language as glides, the *Edge-in Association* was a strong factor as long as there are consonantal segments on the edges. In (1)⁴ are some examples of segmental contraction in Rugao, in which the leftmost segments (marked by single underlines) and the rightmost segments (marked by double underlines) are consistently kept in the contracted forms.

(1) Rugao contraction data showing edge-in association

- a. [si+ hej] → [sej] ‘time’
- b. [nun⁵ + xow] → [now] ‘warm’
- c. [ci + xun] → [cun] ‘to like’
- d. [jəʔ + jæn] → [jæn] ‘same’

² There are a total of 41 words, 3 of which have two contraction forms. The percentages here were calculated assuming there are 44 different tokens.

³ There are different views on the syllable structure of Mandarin dialects and how to treat pre-nucleus and post-nucleus vocalic segments. See discussions in Duanmu (2007) and Lin (2007).

⁴ The syllable contraction data in this paper are taken from Xu (2014) and data collected by the author in downtown Rugao in August 2015 using sociolinguistic interviews. The interviewees were five native speakers who were working in Rugao at that point, speaking Rugao at work and home, and had not lived in other cities except for the 4 years of college. The interviewees came in groups of two or three close friends and talked in topics of daily life, such as an interesting/important event, daily life in Rugao and other topics of their interest.

⁵ The place of the nasal coda is not clear and varies among speakers. This paper uses [n] to represent the nasals for simplicity.

With regards to vowel nucleus selection in Rugao, the LR-scanning model accounts for about 50% of the data, which is just at chance; the sonority-based model accounts for a larger proportion of the data ($\approx 70\%$ ⁶). The examples in (2) show that some are compatible with both analyses, but some others pose crucial challenges to the order-based analysis.

(2) Rugao syllable contraction showing vowel nucleus selection

- a. [ɕjæ̃n + ɕjən] → [ɕjæ̃n] ‘believe’
- b. [tsā + kow] → [tsaw] ‘this’
- c. [si + xej] → [sej] ‘time’
- d. [jin + xej] → [ɹej] ‘then’
- e. [ɕjən + jɔ̃n] → [ɕjɔ̃n] ‘credit’

For (2a) and (2b), the selected vowel in the contracted output is both on the left side and more sonorous than the other vowel, so both the LR-scanning analysis and the sonority-based analysis can account for them. However, data points (2c) to (2e) select vowels on the right side, which cannot be explained by the LR-scanning. The sonority account can solve this problem, assuming a universal sonority hierarchy of [æ/a] > [e/ɛ, o/ɔ] > [i, u/ɯ] > [ə]. For example, high vowels are generally lower in sonority than mid vowels. This explains why [i] and [ɪ] both lose to [e] in (2c) and (2d). The mid central vowel [ə] is among the least sonorous vowels, so the more sonorous [ɔ] in (2e) is selected. Such sonority hierarchy can also explain the vowel selection in (2a) and (2b), as the low vowels [æ/a] are among the most sonorous vowels.

In order to better understand the generalizations made by the speakers and compare the two analyses for vowel nucleus selection via a purely phonological approach while controlling for other biases, an experiment testing contraction of nonce words was conducted. The questions to be addressed are: (1) If native speakers of Rugao are forced to contract nonce words, will they have a systematic preference of a particular type of vowel in the contracted output? (2) If participants do have a systematic bias, is such bias well accounted by the sonority-based analysis? And how? Native speakers of Rugao would hypothetically show a more consistent preference to the more sonorous vowels. The linear order of the two vowels in the pre-contraction forms should have a minor, if any, effect on their choices.

2. Experiment design. As seen in the previous section, many other factors such as morpheme boundary and word frequency can affect the form of contracted output in the contraction process. The nonce word method was adopted to exclude these other factors and solely test on the phonological aspect of this process.

2.1. THE BASICS OF RUGAO VOWELS. Table 1 summarizes Rugao vowels including allophones, assuming pre-nucleus and post-nucleus [i, u, y] are glides [j, w, ɥ]. Among these, [a, o, e, ɛ, i, ɯ, ɯ] can appear in open syllables, while [æ, ɔ, ɛ, ə, i, u] may be followed by a nasal or glottal stop.

⁶ The remaining 30% that do not fit the sonority model include some that fits the LR-scanning model, and some others that involve processes such as vowel coalescence and tone restrictions.

| | Front | Central | Back |
|-------------|----------------|---------|------|
| High | i | i, ɨ | ɯ, u |
| Mid | e, ε | ə | o, ɔ |
| Low | æ ⁷ | a | |

Table 1: Rugao vowels

2.2. STIMULI. All stimuli were phonologically well-formed in Rugao with a couple of exceptions (see below). As the experiment tests the vowel nucleus selection between competing vowels when the segments on the left and right edges are held constant, the syllable structure of the di-syllabic nonce word was kept as $C_1V_1(N_1)^8 + (C_2)V_2N_2$. The initial consonant C_1 was $[p^h]$, as it is frequent and can combine with most of the vowels. C_1 marks the left edge for the contraction process and the nasal coda N_2 on the right edge marks the right boundary, so that the contracted syllables all had a C_1VN_2 structure⁹. Only nasal and glottal stops are allowed in the coda position in Rugao, but glottal stops require specific short high checked or low checked tones, so $[n]$ was used for all the codas. C_2 is a velar fricative $[x]$, except for the stimuli with $[i]$ and $[ɨ]$, for which zero onset was used to keep phonotactic well-formedness. All syllables for the pre-contracted words and the contracted forms had a low falling tone, the most productive tone in Rugao.

Since this experiment tests contraction on tokens with clear left and right consonantal edges, only a subset of vowels that can be followed by a consonant can be tested. For consistency in codas, $[\æ, \varepsilon, \text{ɔ}, i, u]$ are used as they can be followed by a nasal coda. To test on centrality, $[\text{ɨ}]$ needs to be used. Since $[\text{ɨ}]$ cannot appear in closed syllables, $[\text{ɨ}]$ and two unattested syllables, $[\text{*p}^h\text{ɨ}i]$ and $[\text{*p}^h\text{ɨ}n]$, were introduced (see Appendix). Table 2 shows the vowels used for stimuli construction. V_1 and V_2 of different sonority were compared in two dimensions: (1) Vowel height, assuming the hierarchy low > mid > high, i.e., $[\æ > \varepsilon > i]$ for unrounded vowels, and $[\text{ɔ}] > [u]$ for rounded vowels respectively; (2) Centrality or peripherality, assuming the hierarchy $[\æ, \text{ɔ}, \varepsilon, i, u] > [\text{ə}] > [\text{ɨ}]$. See appendix for a complete list of stimuli.

| | Front | Central | Back |
|-------------|-------|---------|------|
| High | i | ɨ | u |
| Mid | ε | ə | ɔ |
| Low | æ | | |

Table 2: Rugao vowels used for stimuli construction

⁷ Some speakers from suburban and rural areas do not have the $[\æ]$ vowel; the corresponding words are pronounced with an $[a]$ instead.

⁸ The syllable $[\text{p}^h\text{ɨ}n]$ with a low-falling tone has very low frequency. Contracted syllables might tolerate syllables with such low frequency, but this might be problematic for the pre-contraction di-syllabic word. In such cases, a C_1V_1 syllable $[\text{p}^h\varepsilon]$ was used when $[\text{p}^h\text{ɨ}n]$ was intended. $[\text{p}^h\text{ɨ}n]$ was added in the stimuli but excluded from the analysis.

⁹ One exception is the contracted form for the high mid vowel, which is CVV due to phonotactic restrictions.

A native speaker of Rugao¹⁰ (female, aged 28) produced all the stimuli in a carrier sentence: [ŋow su? ____ sən p^hin] (‘I say ____ three times.’). The stimuli were not processed further, in order to reflect the naturalistic variation.

2.3. PARTICIPANTS. Participants were 21 native speakers of Rugao (7 males, 14 females, aged 19-32 years), born and raised in downtown Rugao City. They had not lived in other places except for 1 to 4 years at college. All of them reported speaking Rugao exclusively or both Rugao and Standard Mandarin at home.

2.4. PROCEDURE. The experiment was conducted using a MacBook and a Bose QC5 noise-cancelling headset in a quiet room. PsychoPy (Peirce 2007) was used for the presentation of the stimuli. Before the experiment, the first author made sure that the participants understood the task by explaining the concept of syllable contraction as “two syllables pronounced as one” and asking for their judgment of a short conversation of two natural sentences produced by the first author: [nəj sən sej lɛ a? ŋow mæn lɛ!] (‘When are you coming? I’m coming right now!’). Underlined words [sən] [sej] and [mæn] are contractions of [sən.dej] (‘what’), [si.xej] (‘time’) and [ma.sən] (‘right away’) respectively. The experiment proceeded only when the participant agreed that these sentences, as well as the three contracted words, sounded natural.

The experiment involved two phases: a training phase and a test phase. In the training phase, participants practiced with a short version of the test using the tokens that were not used for the test phase. In both phases, participants listened to groups of three words: a di-syllabic nonce word with two different vowel nuclei (Word), a contracted, single syllable whose nucleus is one of the two vowels from the original di-syllabic word (Option 1), and then another contracted syllable with the other vowel nucleus from the original di-syllabic word (Option 2). For example, after the word [p^hin²¹. xən²¹], participants heard two options [p^hin²¹] (Option 1) and [p^hɛn²¹] (Option 2). They then chose between Option 1 and Option 2 by identifying which one was a better contraction of the “word” they just heard. Participants were able to listen to the stimuli multiple times by pressing the keyboard. All test stimuli triplets were presented twice in a pseudo-random order. All participants received the same stimuli.

3. Results.

3.1. TWO MODELS COMPARED. Since the experiment was a forced choice task between two options and there were equal numbers of left vowel choices and right vowel choices, a hypothetical 50% would be expected if null effects were assumed. For all the responses, an average of 52% responses are choices of vowels on the left, and 48% are choices of vowels on the right, which suggests a near-even distribution. There might be a weak bias of left vowel responses for some participants, but the mean proportions of left and right vowels across all participants were almost identical (Figure 2). A one-way within-subjects ANOVA was conducted, and the result shows that the difference is not statistically significant [$F(1, 20) = 0.44, p = 0.52$]. This suggests that the left preference is not a consistent, or the only, bias; there must be factors other than the linear order of the pre-contraction vowels that have influenced participants’ vowel choices.

¹⁰ This was the first author.

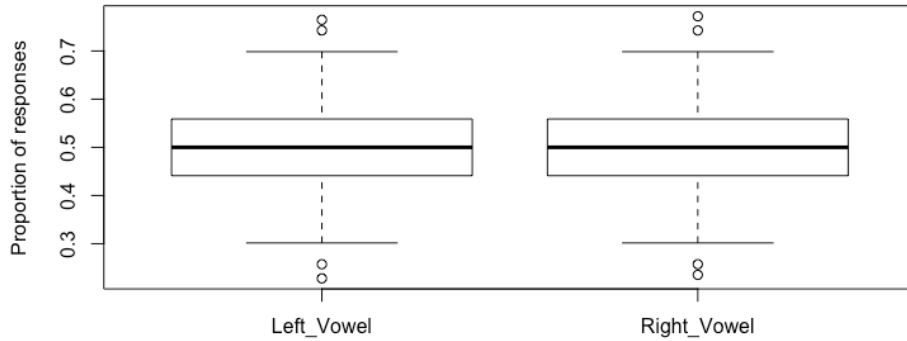


Figure 2. Proportion of left and right vowel responses

The sonority-based analysis accounts for the data better, as a general preference was found for the more sonorous vowels. Figure 3 shows the proportion of responses to the more sonorous vowel and the less sonorous vowel based on the sonority hierarchies in Section 2.2, taking all comparison groups. A one-way within-subjects ANOVA on proportions of responses suggests that, overall, the more sonorous one was more likely to be chosen [$F(1, 20) = 15.16, p < 0.01$]. This suggests that sonority in general plays a role in biasing participants' vowel choice for the contracted syllable.

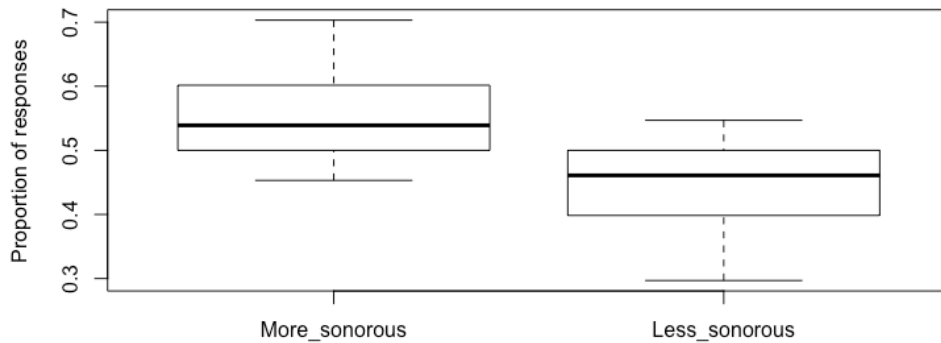


Figure 3. Overall proportions of choices of more sonorous and less sonorous vowels

3.2. VOWEL HEIGHT. Among the features that may have to do with sonority, height (or F1 value) is the predictor that most studies have argued for (Parker 2002, Gordon et al 2012). A ranking of low > mid > high is expected based on previous proposals for the universal hierarchy of vowel sonority. Figure 4 shows the proportions of choice of vowels, comparing vowel height. For example, for the Mid vs. Low comparison, the lower, more sonorous vowel is [æ] and the higher, less sonorous vowel is [ɛ]. Generally, a higher proportion of choices were made to the lower, more sonorous vowels than the higher, less sonorous vowels, assuming vowel height predicts sonority. Five vowels were involved for vowel height testing: low [æ], mid [ɛ], high [i], high rounded [u] and mid rounded [ɔ]. In Figure 4, whether a vowel choice is the more sonorous vowel is based on the proposed sonority ranking based on height separated into two sub-groups: [æ] > [ɛ] > [i] for the unrounded, front vowels, and [ɔ] > [u] for the rounded, back vowels. For these comparisons, a one-way within-subjects ANOVA was conducted to compare the effect of height-based vowel sonority and the proportion of responses in High_Low, Mid_Low, High_Mid, and High-back_Mid-back conditions. The result shows that there was a statistically significant effect of height-based vowel sonority on vowel selection in syllable contraction [$F(1, 20) = 22.56, p < 0.01$].

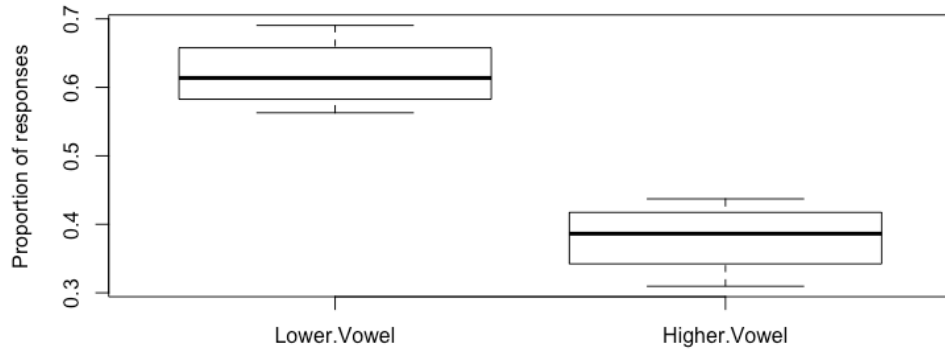


Figure 4. Proportion of choices of vowels of comparative height

There is not only a general effect of sonority on vowel selection, but also a consistent preference for more sonorous vowels across all comparison groups arranged by height, shown in Figure 5. The preference of vowels in each group points to the same direction as in Figure 4. Participants were consistently more likely to choose the comparatively lower vowels when other vowel features were the same. For the front, unrounded vowels [æ, ε, i], low vowel [æ] was chosen more than high vowel [i] on average, although a paired Wilcoxon test¹¹ shows that the difference between high vowel and low vowel is not statistically significant [W = 104.5, p = 0.06]. However, as can be seen in Figure 5, the difference is in the expected direction. Mid vowel [ε] was chosen more than high vowel [i] on average. A paired Wilcoxon test shows that the difference between the mid vowel and high vowel is statistically significant [W = 115.5, p < 0.05]. For low vowel [æ] and mid vowel [ε], [æ] received more responses than [ε] on average, although paired Wilcoxon test shows that the difference is not statistically significant [W = 76, p = 0.16]. The back, rounded vowels [ɔ, u] also exhibit the same height effect. Mid vowel [ɔ] was chosen more than the high vowel [u]. A paired Wilcoxon test shows that the difference between mid rounded vowel and high rounded vowel is statistically significant [W = 141, p < 0.01].

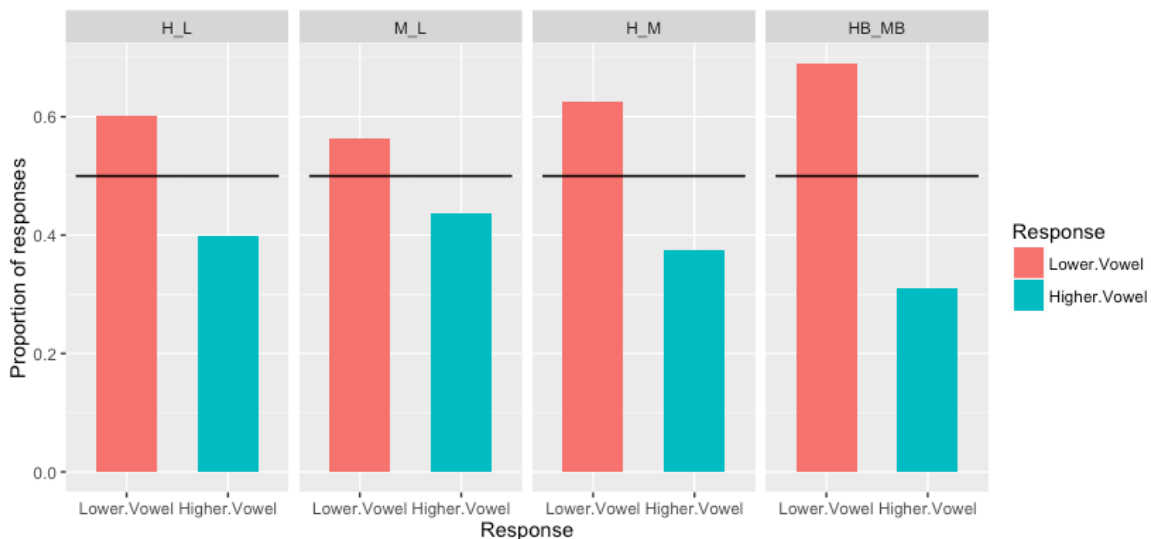


Figure 5. Proportion of responses of vowels of different height. Solid black line shows 0.5. (L: [æ], M: [ε], H: [i], HB: [u], MB: [ɔ])

¹¹ Wilcoxon test was run because it does not make a normality assumption about the errors.

3.3. CENTRALITY (PERIPHERALITY). Previous phonological and phonetic studies have suggested that non-low central vowels are ranked low in many sonority scales in many tested languages (Gordon et al 2012). According to the typically proposed sonority hierarchies, there is a hypothetical ranking of $[i, \epsilon, \text{æ}, u, \text{ɔ}] > [\text{ə}] > [i]$, with peripheral vowels being more sonorous than central vowels and within central vowels, high central vowel being even less sonorous. Figure 6 shows the proportion of choices between peripheral vowels and the mid central vowel [ə], grouping all peripheral vowels $[i, \epsilon, \text{æ}, u, \text{ɔ}]$ into a group. Generally, participants were more likely to choose peripheral vowels than the mid central vowel [ə]. A one-way within-subjects ANOVA was conducted to test the overall effect of centrality on participants' vowel selections in the 5 conditions of peripheral vs. central vowel comparisons. The result shows a statistically significant effect of centrality on vowel selection in syllable contraction [$F(1,20) = 13.77, p < 0.01$].

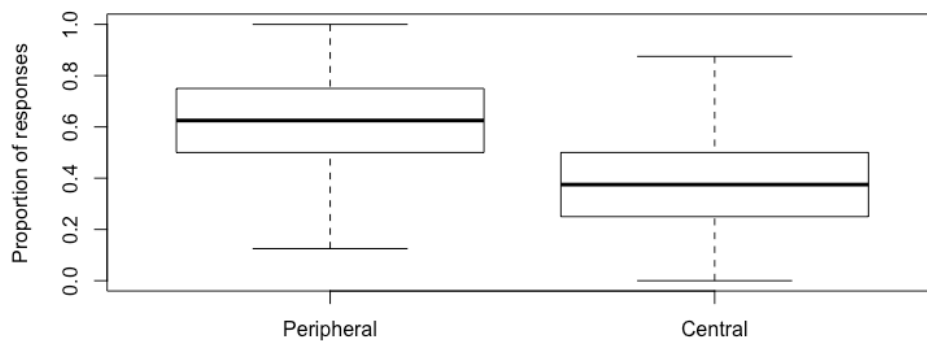


Figure 6. Proportion of choices of peripheral vowel and mid central vowel

In addition to the general effect of centrality, the preference for peripheral vowels is consistent across all comparison groups. Figure 7 shows the proportion of responses in each peripheral vs. [ə] comparison. A paired Wilcoxon test was conducted for each of the five comparison groups. The results show that the differences between central and non-central vowels is statistically significant for high-back [$W = 82, p < 0.01$] and mid-back [$W = 136, p < 0.01$]. Although mid-front [$W = 126, p = 0.21$], high-front [$W = 49.5, p = 0.8$], and low groups [$W = 55, p = 0.5$] did not produce significant results, these groups all had a difference in the same direction.

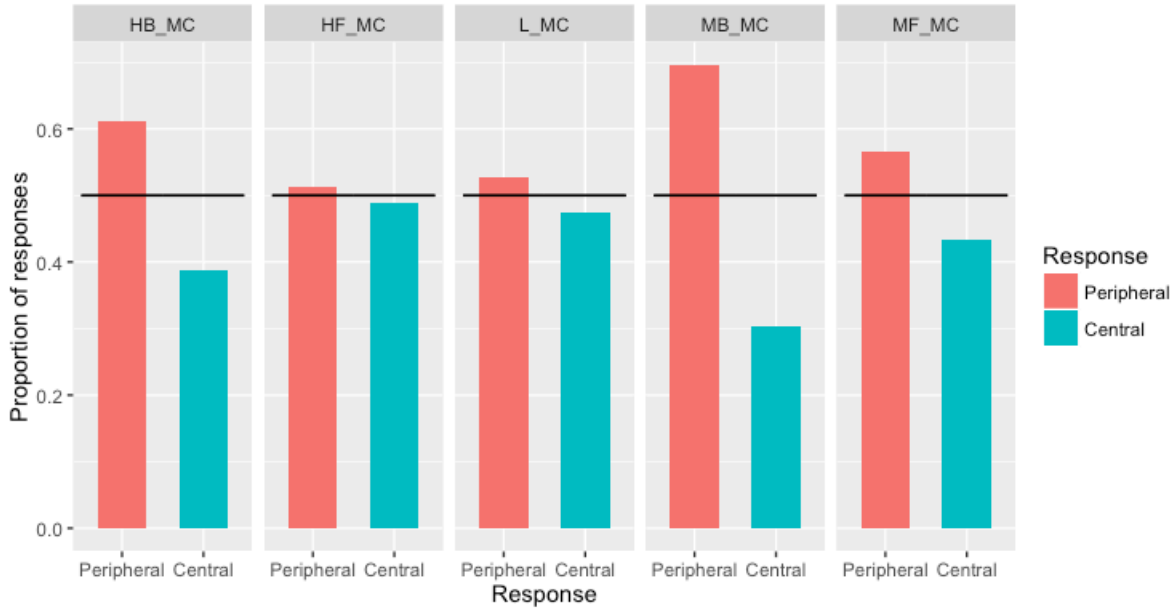


Figure 7. Proportion of responses of vowels of peripheral vowels and mid central vowel [ə] (HB: [u], HF: [i], L: [æ], MB: [ɔ], MF: [ɛ], MC: [ə])

Despite very limited evidence of a preference of [ə] over [i] in the corpus data, the mid central vowel [ə] is typically suggested to be more sonorous than the mid high vowel [i] according to the commonly accepted sonority hierarchy. However, the nonce word experimental data did not reveal any bias between the two central vowels, as shown in Figure 8. A paired Wilcoxon test on the high central and mid central groups shows that there is no statistically significant difference between choices of high central vowel and mid central vowel [$W = 71$, $p = 0.9$].

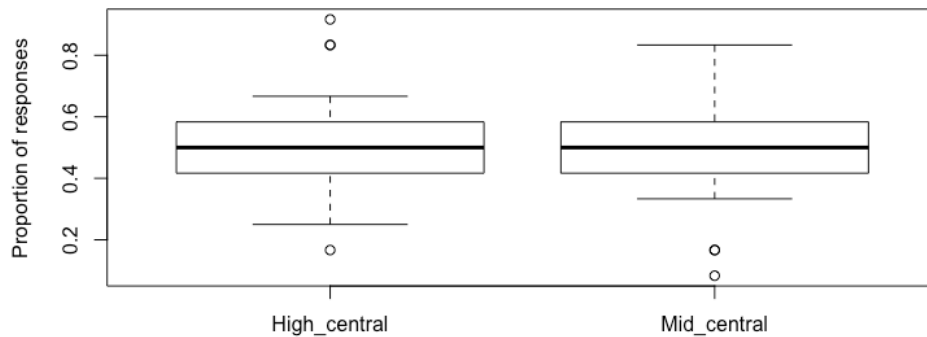


Figure 8. Proportion of choices of high central and mid central vowels

4. Discussion. This experiment examined Rugao speakers' vowel selection in nonce word syllable contraction and compared the two main proposed analyses in accounting for the vowel selection. In general, there was a strong and consistent preference of more sonorous vowels, but no observable preference for vowels on the left or any order-based preference. Although the difference between the more sonorous vowels and less sonorous vowels was about 60% vs. 40%, it is consistently so. More importantly, syllable contraction occurs in casual speech and mostly to high-frequency words and phrases. This could have made the experiment a relatively difficult task for some participants, and hence they may have been reluctant to contract nonce words,

especially in an experimental setting. Given the formality and artificiality of the experiment, the observed difference is sizable. Furthermore, the same trend is consistently found for all comparison groups and in both dimensions of vowel sonority.

Height and centrality are known dimensions of vowel sonority, and they both biased the vowel selection in the nonce word contractions. First, there was a general vowel preference order of low > mid > high. It is clear that high vowels were less preferable to mid vowels and low vowels. This is true for both rounded and unrounded vowels. Ranking of mid and low vowels was in the same direction, but more evidence is needed. A possible explanation for the statistics not being significant is that [ɛ] and [æ] are allophones of the same phoneme and that they share some phonetic similarities, which weakened the bias. Second, peripheral vowels were generally more preferable to central vowels. Non-high central vowel [ə] was consistently less preferable to peripheral vowels [æ, ɛ, ɔ, i, u]. The mid central vowel [ə] is typically ranked as more sonorous than the mid high vowel [i] in the sonority scale and it was observed to be more preferable in contractions in the corpus data, but the nonce word experiment results do not show a noticeable bias between the two vowels in vowel selection. This could suggest that some part of the sonority hierarchy may be language-specific, but there is very limited real word data to begin with. An alternative explanation pointed out by a participant at the 2018 LSA meeting is that Rugao [i] might be different phonetically, i.e. fricative and apical, so the universal sonority ranking may not be compatible with it. More investigation is needed to confirm this.

Data of all the comparisons that involve high vowels seem messier than those without (Figure 5, Figure 7), as the difference between the vowel predicted by sonority and the other one is small. This is probably because of the Glide Formation process that satisfies the Maximal Syllable constraint. According to Hsu (2003) and Xu (2014), the contracted forms follow Maximality (Prince 1985) and construct the largest possible syllable. One way to keep the maximal syllable in the contracted output is to form a glide whenever possible. High vowels [i, u] are possible candidates for the glides [j, w], e.g., [p^hin²¹ + ɛən²¹] → [p^hjən²¹] ('favor'). None of the stimuli with high vowels undergo glide formation as a sacrifice to keeping the three-X-template structure consistent. For example, stimulus [p^hi + xæn] has a more natural contraction [p^hjæn]¹², which was not one of the given options [p^hin] and [p^hæn]. As a consequence, participants might have been reluctant to choose between the given two options without glide formation, as neither of them was an optimal choice. This could explain the reason why the low vowel and high vowel comparison did not produce statistically significant results.

During the experiment, participants could press the "C" key if they wanted to hear a triplet again. Most participants did use this option. Responses of repetitions were excluded from the analysis; only their last choice was included in the shown results. A preliminary manual check suggests more repetitions occurred in response to tokens with illicit syllables. The presence of illicit syllables in contraction options [*p^həi] and [*p^hiɪn] might have confused participants and biased their choices. The corpus data reveals a noticeable preference for well-formed syllables in contracted syllables, which may be the reason for participants' hesitation when being forced to choose between a licit syllable with the less sonorous vowel and an illicit syllable with the more sonorous vowel. More research needs to be done to confirm whether well-formedness is ranked higher than the sonority bias in forming the contracted output.

¹² This is not an attested syllable in Rugao. However, it seems like a more natural contraction for the first author (a native speaker). The same applies to potential contracted forms such as [p^hjun] and [p^hwin].

5. Conclusion. This study was based on the phonological analysis of vowel selection in Rugao syllable contraction, and confirmed the sonority bias using experimental data from nonce word contractions. Rugao syllable contraction shows edge-in effects and the vowel nucleus selection is conditioned by sonority more than a linear-scanning bias. As predicted by the sonority-based analysis, more sonorous vowels are more likely to be selected as the vowel nucleus in the contracted syllables. Specifically, sonority along the dimensions of both height and centrality exhibit the same bias on the nucleus of the contracted syllable output. Sonority as a single property is thus a way to unify two different vowel features, height and centrality, in accounting for the corpus and experimental data. A more nuanced ranking of specific vowels could not be made from the current experiment, partially because of insufficient statistical power, but the general pattern is robust and consistent. The vowel that gets preserved in di-syllabic contractions is essentially the survivor of two vowels in competition for a single timing slot, which may go beyond syllable contraction and generalize to vowel competition in vowel-related phonological processes in a broader sense. This study of syllable contraction thus offers a new way to understand how sonority biases vowel selection in vowel competition.

References

- Cheng, Chierh, & Yi Xu. 2009. Extreme reductions: Contraction of disyllables into monosyllables in Taiwan Mandarin. *Proceedings of the Tenth Annual Conference of the International Speech Communication Association*, 456-459.
- Chung, Raung-Fu. 1996. *The segmental phonology of Southern Min in Taiwan*. Taipei: Crane Publishing.
- Chung, Raung-Fu. 1997. Syllable contraction in Chinese. *Proceedings of Chinese Languages and Linguistics III: Morphology and Lexicon*, 199-235.
- Duanmu, San. 2007. *The phonology of Standard Chinese*. 2nd edn. Oxford: Oxford University Press.
- Gordon, Matthew, Edita Ghushchyan, Bradley McDonnell, Daisy Rosenblum, and Patricia A. Shaw. 2012. Sonority and central vowels: A cross-linguistic phonetic study. In Steve Parker (ed.) *The sonority controversy*, 219-256. Berlin: Mouton de Gruyter.
- Hsu, Hui-Chuan. 2003. A sonority model of syllable contraction in Taiwanese Southern Min. *Journal of East Asian Linguistics*, 12(4), 349-377.
- Kiparsky, Paul. 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10(3), 421-441.
- Li, Charles. N., & Sandra A. Thompson. 1989. *Mandarin Chinese: A functional reference grammar*. Oakland, CA: University of California Press.
- Lin, Yen-Hwei. 2007. *The sounds of Chinese*. Cambridge & New York: Cambridge University Press.
- Myers, James, & Yingshing Li. 2009. Lexical frequency effects in Taiwan Southern Min syllable contraction. *Journal of Phonetics*, 37(2), 212-230.
- Parker, Stephen George. 2002. *Quantifying the sonority hierarchy*. Amherst, MD: University of Massachusetts dissertation.

- Peirce, Jonathan W. 2007. PsychoPy - psychophysics software in Python. *Journal of Neuroscience Methods*, 162(1), 8-13.
- Prince, Alan S. 1985. Improving tree theory. *Proceedings of Annual Meeting of the Berkeley Linguistics Society*, 11, 471-490.
- Sun, Hongju 孙红举. 2014. 论汉语合音现象的研究 [Analysis on Chinese merged sounds]. 西南大学学报 (社会科学版) [*Journal of Southwest University (Social Sciences Edition)*], 40(1), 115-125.
- Tseng, Shu-Chuan. 2005. Syllable contractions in a Mandarin conversational dialogue corpus. *International Journal of Corpus Linguistics*, 10(1), 63-83.
- Wee, Lian-Hee, Xiuhong Yan, & Jilun Lu 黄良喜, 严修鸿 & 路继伦. 2005. 天津方言的吞音现象 [Swallowing sounds in Tianjin]. *语言科学 [Linguistic Sciences]*, 4(4), 66-75.
- Wong, Wai Yi Peggy. 2006. *Syllable fusion in Hong Kong Cantonese connected speech*. Columbia, OH: The Ohio State University dissertation.
- Xu, Chenchen. 2014. An OT analysis on syllable contraction in Jianghuai Chinese. Paper presented at the 22nd Annual Conference of the IACL & the 26th North American Conference on Chinese Linguistics, University of Maryland-College Park, MD.
- Yip, Moria. 1988. Template morphology and the direction of association. *Natural Language & Linguistic Theory*, 6(4), 551-577.

Appendix: Stimuli (L = low, M = mid, H = high, MR = mid rounded, HR = high rounded, HF = high front, MC = mid central, HC = high central)

| Comparison group | Word | Option 1 | Option 2 |
|------------------|--|--------------------|--------------------|
| M + L | p ^h ε + xæn | p ^h εn | p ^h æ̃n |
| L + M | p ^h æ̃n + xεn ¹³ | p ^h æ̃n | p ^h εn |
| H + M | p ^h in + xεn | p ^h in | p ^h εn |
| M + H | p ^h ε+ in | p ^h εn | p ^h in |
| H + L | p ^h in + xæn | p ^h in | p ^h æ̃n |
| L + H | p ^h æ̃n + in | p ^h æ̃n | p ^h in |
| MR + HR | p ^h ɔ̃n + xun | p ^h ɔ̃n | p ^h un |
| HR + MR | p ^h un + xɔ̃n | p ^h un | p ^h ɔ̃n |
| M + MC | p ^h ε + xən | p ^h εn | p ^h ən |
| MC + M | p ^h ən + xεn | p ^h ən | p ^h εn |
| MR + MC | p ^h ɔ̃n + xən | p ^h ɔ̃n | p ^h ən |
| MC + MR | p ^h ən + xɔ̃n | p ^h ən | p ^h ɔ̃n |

¹³ Some speakers may produce a glide [j] before [ε].

| | | | |
|---------|-------------------------|-------------------|--------------------|
| HF + MC | p ^h in + xən | p ^h in | p ^h ən |
| MC + HF | p ^h ən + in | p ^h ən | p ^h in |
| HR + MC | p ^h un + xən | p ^h un | p ^h ən |
| MC + HR | p ^h ən + xun | p ^h ən | p ^h un |
| L + MC | p ^h æn + xən | p ^h æn | p ^h ən |
| MC + L | p ^h ən + xæn | p ^h ən | p ^h æn |
| MC + HC | p ^h ən + i | p ^h ii | *p ^h əi |
| HC + MC | p ^h i + xən | p ^h ən | *p ^h in |
| HC + MC | p ^h i + xən | p ^h ən | p ^h in |