

THE PHONOLOGY AND PHONETICS OF RUGAO SYLLABLE
CONTRACTION: VOWEL SELECTION AND DELETION

By

Chenchen Xu

A DISSERTATION

Submitted to
Michigan State University
in partial fulfilment of the requirements
for the degree of

Linguistics — Doctor of Philosophy

2020

ABSTRACT

THE PHONOLOGY AND PHONETICS OF RUGAO SYLLABLE CONTRACTION: VOWEL SELECTION AND DELETION

By

Chenchen Xu

In Chinese languages, when two syllables merge into one that has the segments from both, the segments compete to survive in the limited time slots (Chung, 1996, 1997; Lin, 2007). The survival or deletion of segment(s) follows a series of rules, including the Edge-In Effect (Yip, 1988) and vowel selection (R.-F. Chung, 1996, 1997; Hsu, 2003), which decide on the outer edge segments and vowel nucleus, respectively. This dissertation is dedicated to investigating the phonological patterns and phonetic details of syllable contraction in Rugao, a dialect of Jianghuai Mandarin, with more focus on the vowel selection and deletion process. First, I explored the segment selecting mechanism, including the preservation or deletion of the consonantal and vocalic segments, respectively. Based on the phonological analyses, I further investigated two major questions: 1) what determines the winner of the two vowel candidates for the limited nucleus slot in the fully contracted syllable, the linearity of the vowels (R.-F. Chung, 1996, 1997) or the sonority of the vowels (Hsu, 2003), and 2) is a fully contracted syllable phonetically and/or phonologically neutralized to a non-contracted lexical syllable with seemingly identical segments with regards to syllable constituents, lengths, and vowel quality?

The corpus data suggest that, 1) the Edge-In Effect (Yip, 1988) is prevalent in Rugao syllable contraction in deciding the survival of the leftmost and rightmost segments in the pre-contraction form whether they are vocalic or not, unless the phonotactics of the language overwrite it. 2) In fully contracted syllables, the winner of the two vowel candidates is contingent upon the sonority of the vowels as well as the phonotactics of the language. Following such

patterns, a forced choice experiment focused on the selection of the vowel nucleus that controlled the syllable structure and used nonce words confirmed the influence of sonority in the vowel competition and ruled out the factor of relative linear order of the vowels. Generally, the vowel of higher sonority is more likely to survive than the competitor of lower sonority ranking, assuming a vowel sonority hierarchy based on height and centrality. The surviving vowels in the contracted syllable were then further examined with production experiments and acoustic measurements. The results suggest that the deletion of the losing vowel is in fact incomplete, and manifested in two ways, 1) the contracted vowel is longer than the lexical vowel in general, although the ratio of vowel in the duration of the syllable may or may not be different, 2) the contracted vowel has different F1 and/or F2 values than its lexical counterpart, suggesting the vowel quality has altered in the contracted syllable. The phonologically defined process is shown to be phonetically quite complex, suggesting that the lexical distinction is maintained in some ways even though the two words seem neutralized.

For my family and my hometown, Rugao

ACKNOWLEDGEMENTS

Having to complete, revise and defend a dissertation in a global pandemic is never the way I expected my last year of my graduate school to be. But this is finally the moment. Looking back to my graduate life, I realize that I have received much encouragement and support from many people, without whom I couldn't possibly have completed this endeavor.

First, I'd like to express my gratitude to my advisor and committee chair Dr. Yen-Hwei Lin. She is the reason that I chose to study phonology. Before taking her classes, I believed that phonology was not my thing. However, I fell in love with phonology the very first semester of being Dr. Lin's student. Lin Laoshi, you always help me expand my visions and offer detailed guidance for my next steps. Your wisdom and patience are invaluable for me. Thank you for being so encouraging when I feel hesitant; thank you for being so supportive when I feel lost; and thank you for being so understanding when I feel stressed. I am so indebted to you for your tremendous support.

I am also grateful for Dr. Karthik Durvasula, my committee member, who taught me so many skills and offered so much help in my research. He always pushes me to think deeper and more broadly and to challenge myself to be better. He is also a great co-author for the many papers and presentations we worked on together. Thank you, Karthik! Without your help, I couldn't have come this far in my research.

I have an amazing dissertation committee, the other two members being Dr. Suzanne Wagner and Dr. Anne Violin-Wigent. They have offered so much valuable feedback and input for my dissertation and my two comprehensive writing papers. Suzanne, thank you for your advice for my sociolinguistics comp paper, which turned into a published paper that I still enjoy

reading. And thank you for always being so, so supportive in other aspects of my graduate life. Anne, you have such a welcoming personality. I enjoyed it every time I talked to you.

I would like to extend my gratitude to my teaching supervisors, Dr. Alan Munn, Dr. Xiaoshi Li and Dr. Chunghong Teng for their patient guidance. Alan, you are such a great supervisor. I love teaching linguistics because of your guidance. Thank you! Li Laoshi and Teng Laoshi, thank you for your patience in supervising my Chinese teaching. Teng Laoshi, I hope you rest in peace. I miss you.

My cohort, Andrew Armstrong, Ni La Le, Ai Taniguchi, Sayako Uehara and Xiaomei Wang, my TA fellows Cara (Danny) Feldscher, Qian Luo and Alex Mason, and friends Ho-Hsin Huang, Monica Nesbit and Yingfei Chen, have filled my graduate life with laughter and warmth. With you, I could always find a sense of belonging when home is tens of thousands of miles away. I will never forget the moments we spent together studying, laughing and having fun.

I would also like to thank the many participants for my experiments, without whom I couldn't possibly have completed my research. Thank you for putting up with the nonsense words and long-ish experiments and giving me valuable data. My thank you also goes to my phonolab people. Thank you for your feedback for my research and helping me improve my presentations.

Last but not least, I would like to thank my family, my husband Pengfei Ding, my son Jeremy Youheng Ding and my parents Jie Xu and Jieqin She. Thank you for always being there for me. And thank you for all your support and unconditional love. I love you more than anything.

TABLE OF CONTENTS

LIST OF TABLES	X
LIST OF FIGURES	XI
1 INTRODUCTION.....	1
1.1 The concept of syllable contraction	1
1.2 Syllable contraction in Chinese: sociolinguistics, phonology and phonetics	4
1.3 A brief introduction to Rugao and Rugao syllable contraction	7
1.3.1 The dialect of Rugao	7
1.3.2 Syllable contraction in Rugao	10
1.3.3 An overview of the corpus	14
1.4 An overview of this dissertation	15
1.4.1 Research questions.....	15
1.4.2 Methodologies.....	17
1.4.3 A roadmap of the dissertation	18
2 PHONOLOGICAL PATTERNS.....	20
2.1 Introduction to the Rugao dialect.....	20
2.1.1 Consonants	20
2.1.2 Vowels	24
2.1.3 Tone	26
2.1.4 Syllable	27
2.2 Edge-In Association.....	28
2.2.1 What syllable contraction is and how it happens	28
2.2.2 The Edge-In Effect.....	33
2.3 Vowel selection.....	45
2.3.1 Two approaches	45
2.3.2 Sonority and vowel selection	52
2.4 Further discussion	56
2.4.1 More on vowel competition	56
2.4.2 Violable phonotactics.....	57
2.4.3 Problems with using corpus data	59
2.5 Conclusion	61
3 VOWEL COMPETITION IN SYLLABLE CONTRACTION.....	62
3.1 Introduction.....	62
3.1.1 Motivation for this study.....	62
3.1.2 Linear order or sonority?	63
3.1.3 Vowel sonority.....	66
3.1.4 A preview of the experiment.....	69
3.2 A forced-choice contraction task	71
3.2.1 Stimuli.....	71
3.2.2 Participants and procedure	77

3.2.3	Results and analysis	79
3.2.3.1	The two models compared	79
3.2.3.2	Height.....	83
3.2.3.3	Centrality (peripherality)	88
3.3	Discussion.....	92
3.3.1	Sonority in syllable contraction	92
3.3.2	Remaining issues	96
3.4	Conclusion and implications.....	99
4	THE PHONETICS OF VOWELS IN SYLLABLE CONTRACTION	102
4.1	Introduction.....	102
4.1.1	Motivation for this study.....	102
4.1.2	A preview of the experiments	109
4.2	Production Experiment I: natural sentence repetition task	111
4.2.1	Stimuli.....	111
4.2.1.1	Test items	111
4.2.1.2	Stimuli.....	115
4.2.2	Participants and procedure	119
4.2.3	Measurement.....	121
4.2.4	Results.....	123
4.2.4.1	Length and vowel ratio	123
4.2.4.2	Vowel quality and vowel space	128
4.2.4.2.1	F1.....	128
4.2.4.2.2	F2.....	130
4.2.5	Interim summary	131
4.3	Experiment II: Self-paced word contraction task	131
4.3.1	Stimuli.....	132
4.3.2	Participants and procedure.....	132
4.3.3	Measurements	134
4.3.4	Results.....	135
4.3.5	Vowel/Word ratio	135
4.3.5.1.1	Contracted vs. monosyllabic lexical.....	136
4.3.5.1.2	Contracted vs. disyllabic lexical.....	137
4.3.6	Vowel quality and vowel space	138
4.3.6.1.1	F1.....	138
4.3.6.1.2	F2.....	140
4.3.7	Interim summary	143
4.4	Discussion.....	143
4.4.1	Length and vowel ratio	143
4.4.2	Incomplete vowel deletion	145
4.4.3	Syllable contraction	149
4.4.4	Remaining issues	152
4.4.5	High vowels	152
4.4.6	How much difference can the experimental methodology make?	155
4.4.7	Back to the centrality issue	159
4.5	Conclusions and implications	160

5	CONCLUSION	164
5.1	Summary of findings and further discussions.....	164
5.2	Remaining issues and future projects	169
	APPENDICES	172
	Appendix A: Full list of questions used for corpus data collection.....	173
	Appendix B: Full list of stimuli for practice session–forced choice contraction.....	173
	Appendix C: Full list of practice sentences, annotated in standard <i>pinyin</i>	173
	Appendix D: Full list of <i>burn-in</i> sentences, annotated in standard <i>pinyin</i> . Contractible words bolded.....	174
	Appendix E: Full list of test sentences. Tested contractible words bolded. Tested monosyllabic word underlined.	174
	Appendix F: Script of audio introduction to syllable contraction and task	178
	Appendix G: Full list of words in the Rugao Syllable Contraction Corpus	179
	REFERENCES.....	184

LIST OF TABLES

Table 1. Rugao initials (Wu 2006:58).....	21
Table 2. Rugao consonants	22
Table 3. Rugao vowels—phoneme level	25
Table 4. Rugao vowels—including all allophones	26
Table 5. Distribution of Rugao vowels	73
Table 6. Vowels used for stimuli construction	73
Table 7. Stimuli for non-central vowels, compared by height.....	76
Table 8. Stimuli for vowels of same height, compared by roundedness or backness.....	76
Table 9. Stimuli for central vowel and peripheral vowels, compared by centrality	76
Table 10. Contractible words, numbers indicating tones.....	113
Table 11. Word list—monosyllabic lexical words. [Numbers indicate tones. Tokens with an asterisk are dialect words that do not have standard written forms.].....	115
Table 12. Number of tokens analyzed for each experiment	157

LIST OF FIGURES

Figure 1. Example of non-contracted disyllabic word [zəŋ.ka]	13
Figure 2. Example of fully contracted word [zɑ]	13
Figure 3. Tianjin syllable elision (adopted from Wee (2014))	30
Figure 4. Universal vowel sonority hierarchy based on height and centrality (adopted from Gordon et al 2012)	53
Figure 5. PsychoPy interface for training and test. [Note: The English translations were removed from the interface that the participants used.]	78
Figure 6. Proportion of responses for vowels on the left and vowels on the right. [Each line refers to a single participant. Each dot represents a single data point.]	81
Figure 7. Proportion of responses of more sonorous vowels and less sonorous vowels. [Each line refers to a single participant. Each dot represents a single data point.]	82
Figure 8. Proportion of responses of vowels of different height. [Each line refers to a participant; each dot represents a single data point. Some lines and dots are overlapped.]	84
Figure 9. Proportion of responses arranged by height group. [Each line refers to a participant; each dot represents a single data point.]	86
Figure 10. Proportion of responses of rounded/back and unrounded/front vowels. [Each line refers to a single participant. Each dot represents a single data point.]	87
Figure 11. Proportion of responses arranged by roundedness/backness. [Each line refers to a participant; each dot represents a single data point.]	88
Figure 12. Proportion of responses to peripheral and central vowels. [Each line refers to a single participant. Each dot represents a single data point.]	89
Figure 13. Proportion of responses to vowels arranged by peripherality/centrality. [Each line refers to a participant; each dot represents a single data point.]	91
Figure 14. Praat screenshot for [pjən.ɛən] ('fridge'). [Note: the Praat annotation may not be accurate for ease of annotation.]	96
Figure 15. Example of fully contracted token, [tsju?], contracted from [tsən.ju?]	122
Figure 16. Word duration—Experiment I. [Red boxes present contracted tokens. Green boxes represent lexical tokens.]	124

Figure 17. Vowel duration—Experiment I. [Asterisks indicate significance levels.]	126
Figure 18. Vowel ratio—Experiment I. [Asterisks indicate significance levels.]	127
Figure 19. F1 (Hz) at midpoint—Experiment I. [Asterisks indicate significant levels.].....	129
Figure 20. F2(Hz) at midpoint—Experiment I. [Asterisks indicate significant levels.].....	131
Figure 21. Vowel/Word ratio—Experiment II. [No.con.mono = non-contracted monosyllabic, no.con.disy = non-contracted disyllabic. Asterisks indicate significant levels.]	136
Figure 22. F1 (Hz) at midpoint—Experiment II. [No.con.mono = non-contracted monosyllabic, no.con.disy = non-contracted disyllabic. Asterisks indicate significant levels.]	140
Figure 23. F2 values (Hz) at midpoint—Experiment II. [No.con.mono = non-contracted monosyllabic, no.con.disy = non-contracted disyllabic. Asterisks indicate significant levels.].	142
Figure 24. F1(Hz) for words with high vowels [i, u]. [Experiment I data on top, and Experiment II data below]	154
Figure 25. F2(Hz) of words with high vowels [i, u]. [Experiment I data on top and Experiment II data below].....	155
Figure 26. Screen shot of Praat interface on [pəj], contraction of [pən.jow] ('friend')	160

1 INTRODUCTION

1.1 The concept of syllable contraction

In many languages, syllable contraction can be a process whereby a syllable is omitted. In contrast, in Chinese, syllable contraction is the phenomenon by which the components of multiple syllables merge, resulting in a decrease in syllable number (R.-F. Chung, 1996, 1997; Hsu, 2003; Y.-H. Lin, 2007). It is most commonly the process by which two consecutive syllables are merged into one syllable that has segments from both (Y.-H. Lin, 2007). One major difference between the Chinese-type of syllable contraction and that of languages such as English (e.g., *I am, that's*) is that the segments in Chinese syllable contraction do not necessarily undergo complete deletion, making Chinese syllable contraction phonetically and phonologically more complicated (Kuo, 2010; Myers & Li, 2009; Wong, 2006). With regards to segment deletion, there are at least two degrees of contraction: full contraction and partial contraction, the former having only one surviving vocalic nucleus while the latter may have more than one. Chinese syllable contraction is also heavily conditioned by phonology and possibly phonetics, allowing only very limited variation for a certain word. For these reasons, some researchers make a distinction between syllable contraction and syllable fusion (Wong, 2006). However, for the simplicity of the concept, I will call the general phenomenon of syllable count reduction and segment elision as syllable contraction.

Most contracted forms are not lexicalized, nor do they have standard written forms. However, some contracted forms have been lexicalized in the longer history of use, and of these, some have been assigned conventional written forms. For example, in the Northern Mandarin

dialect, the syllables *san*₅₅ *ge*₀ (三个, “three-CLASSIFIER₁”, [sæn₅₅kə₀]₂) merge into a single syllable *sa*₅₅ [sa₅₅], represented by its own written form 仨 (Chao, 1927). Some lexicalizations have even entered Standard Mandarin, such as the contraction of *ni*₂₁₄*men*₀ (你们, “you-plural”, [ni₂₁₄mən₀]) into *nin*₁₅ (您, [nin₁₅], “you (honorific)”) (Lü, 1984). With the ever-increasing use of online text, more and more new written words are being invented to represent some of the most popular contracted words in Mandarin. For example, the contraction of 女朋友 (“girlfriend”, *nu*₂₁₄.*peng*₁₅.*you*₀, [ny₂₁₄.phəŋ₁₅.jow₀]), contracted as [ny₂₁₄.phew₅₁], can be written as 女票 (*nu*₂₁₄.*piao*₅₁, [ny₂₁₄.phjaw₅₁]) in online texts (Xu & Mao, 2017). As seen from the phonetic transcription, 女票 does not truthfully represent the pronunciation of the contraction, and speaker may actually pronounce [ny₂₁₄.phjaw₅₁]. In fact, the lexicalized contraction may often be influenced by the orthography throughout the history of its development. This dissertation only concerns contracted words that do not have conventional written forms for this reason.

Syllable contraction has been documented and analyzed for many varieties of Chinese including but not limited to Taiwan Mandarin (Cheng & Xu, 2009; R.-F. Chung, 1997; Hsiao, 1995; Y.-H. Lin, 2007; Tseng, 2005b, 2005a), Taiwan Southern Min (R.-F. Chung, 1996, 1997; Hsu, 2003), Hakka (R.-F. Chung, 1997), Cantonese (Wong, 2006), Tianjin (Wee, 2014) and Rugao (Wu, 2006; Xu et al., 2018), among many others. Most cases of the available Chinese data involve two syllables merging into one syllable, although contraction of more than two syllables does happen (Sun, 2014; Tseng, 2005b; Wee, 2014). In the Mandarin Conversational

¹ Classifiers are the obligatory words that come between numerals and nouns.

² Numbers on the syllables indicate tones on a 1-5 scale, with 1 for the lowest and 5 the highest pitch (Chao, 1968).

Dialogue Corpus, for example, 74% of the contracted tokens are from disyllabic sequences, and the rest involve three or even more syllables (Tseng, 2005b). As disyllabic contraction is the most common type of contraction, I will limit the scope of this dissertation to cases where two syllables contract into one, more specifically, syllable contraction of the originally disyllabic word into a single syllable, or a tri-syllabic word with a contractible disyllabic part. A few examples of such cases are listed below:

Standard Mandarin, data obtained from (Y.-H. Lin, 2007)

- (1) pi₃₅ + tɛjau₂₁₄ → pjau₃₅₂ ‘comparatively’
- (2) na₅₁ + ji₅₅ → nei₅₃ ‘that one’
- (3) ta₅₅ + mən₀ → tam₅₃ ‘they’

Tianjin, data obtained from (Wee, 2014); tones not marked in the original text

- (4) tɕhi + ɕaŋ + thai → tɕhiaŋ + thai ‘weather station’
- (5) mien + faŋ + tɕhaŋ → mian + tɕhaŋ ‘cotton factory’
- (6) tɕi + kuan + tɕhiaŋ → tɕyan + tɕhiaŋ ‘machine gun’

The data above demonstrates some examples of typical syllable contraction, in which the initial consonant usually appears in the contracted syllable while the rimes of both syllables may contribute to the contracted form (Y.-H. Lin, 2007). The contracted forms are often not well-formed syllables in the language, such as [tam₅₃] in (3) and [mian] in (5). In fact, the violation of the phonotactic constraints in syllable contraction is commonly seen (Hsiao, 1995; Hsu, 2003; Y.-H. Lin, 2007). This is possibly due to the compromise that has to be made to maintain the meaning of the original disyllabic form to avoid confusion and the violable phonetic process in fast speech (Y.-H. Lin, 2007).

Tone contraction happens along with the segment contraction process, as seen in the data presented in (1)—(3) above. Compared to the segment combinations in the syllables, the tones are even more likely to be irregular, which is likely due to the similar reasons for irregular syllables and some OCP motivations (Y.-H. Lin, 2007; Wee, 2014). In this dissertation, I approach this phenomenon starting from the segmental contraction and focus on how the segments behave in this process. The tonal patterns will be studied in a different project. For this reason, tones are not discussed in the main body of this dissertation unless directly relevant to the segmental analyses.

1.2 Syllable contraction in Chinese: sociolinguistics, phonology and phonetics

Syllable contraction in Chinese has attracted research in a wide variety of Chinese dialects/languages and in various aspects of linguistics, including the sociolinguistics, phonology and phonetics of this phenomenon. I will briefly introduce some representative work in these three subfields in this subsection.

First, syllable contraction happens mostly in casual, connected speech; the latter in turn promotes the likelihood of contraction occurrence (Tseng, 2005a; Wee, 2014; Wong, 2006). Interspeaker and intraspeaker variations are observed, with different degrees of contraction and slightly variant contraction forms for the same words (Cheng & Xu, 2009; Chung, 2006; Kuo, 2010; Wong, 2006). While lay speakers generally are unaware of the use of syllable contraction, the use of contracted forms varies in different speech registers, the social class and age of the speakers, personal speech habits, and emotional involvement in the topic (Chung, 2006). As a product of speech style, Chinese syllable contraction has complex sociolinguistic meanings, according to a socio-perceptual study: lay listeners from mainland China tend to associate

contraction-featured speech with a southern accent and have different attitudes towards contraction users of different genders (Xu & Mao, 2017). In particular, public attitudes are more uniform towards the stereotyped contraction forms, i.e., those online-popular contractions with newly invented written forms such as 酱紫 (*jiangzi*, ‘this way’) and 表 (*biao*, ‘don’t’) (Xu & Mao, 2017), which were perceived as southern and feminine.

Speech-related factors, including high lexical frequency and high speech rate, have been found to promote the occurrence of contraction (Cheng & Xu, 2009; Myers & Li, 2009). Lexical frequency is reported to be an influencing factor in the use of Taiwan Southern Min syllable contraction: more frequently used lexical items are more likely to be contracted in natural speech, and both segment reduction and tonal merger correlate with lexical frequency (Myers & Li, 2009). There is also a strong correlation between speech tempo and the frequency of syllable fusion in Cantonese (Wong, 2006), the faster the tempo, the more cases of contraction. A nonce word production experiment also reported a positive correlation between speech rate and degree of contraction, i.e., the faster the speech is, the more likely contraction happens for Standard Mandarin speakers (Cheng & Xu, 2009).

Available phonetic studies of syllable contraction focus mostly on the likelihood and degree of contraction. First, phonetic studies offer many details about the phonetics of words that are prone to contraction. More specifically, di-syllabic words that undergo full contraction (i.e., contraction to a true single syllable with one vocalic nucleus) are characterized by continuous formants, with no interruption of nasals, intervocalic plosives, fricatives or affricates³ in between the two vowels in the pre-contraction form; di-syllabic words with a clear interruption of

³ Cheng and Xu (2009) did not test approximants.

formants by nasals⁴, intervocalic consonants between the vowels or a lowered F1 are least likely to contract in Standard Mandarin syllables (Cheng & Xu, 2009). Syllable contraction is not a categorical shift, but a gradient undershot of articulatory targets as a function of time pressure; for this reason, different contracted outputs with different degrees of contraction could surface for the same pre-contraction form. In some cases, one or two consonantal segments are lost, and in others, vowel coalescence occurs in addition to the consonant elision (Cheng & Xu, 2009). The structures of the involved syllables and the realization of the actual segments vary to a large extent, and such variation carries even within speakers. With regards to the phonetic details of the contracted output, the serial experiments by Kuo (2010) find that, first, syllable contraction is optional and gradient, which leads to the implication that syllable contraction is not a complete neutralization. The fully contracted words have longer durations, larger ratios of vowel duration and wider F0 range. Male speakers use a larger vowel space for the fully contracted tokens, with a higher /i/. In addition, speakers can utilize the differences in durations and F0 range to distinguish between the fully contracted tokens and their lexical counterparts.

A large proportion of the existing research on syllable contraction involves phonological analysis of this process. With regards to the contracted form, despite the variation in the actual contracted outputs, the contraction is definitely not random; there are many phonologically driven factors that determine the contracted form, such as prosody (Wong, 2006), Edge-In Association (R.-F. Chung, 1996, 1997; Hsu, 2003), and vowel sonority (Hsu, 2003). In particular, the forms of the contracted output follow some general rules. Assuming that Chinese

⁴ This experiment was based on Taiwan Mandarin nonce words. However, the Rugao data includes many cases of contraction of words with intervocalic nasals.

syllables have a basic three-slot template XXX, di-syllabic contraction can be viewed as six slots merging into three slots, the latter keeping some segments of the six-slot form while losing some others (Chung, 1997). In this templatic account, the association of segments to the templates starts with the two edges: when two syllables are to be contracted, the leftmost and rightmost segments of the di-syllabic word need to be preserved in the contracted output, the so-called *Edge-In Association*. With regards to what vowel survives in the contracted output, Chung (1996, 1997) proposed the LR-scanning model, while Hsu (2003) employs vowel sonority to account for the vowel selection patterns.

Other linguistic factors that may influence the use of contraction are: morpheme boundaries (Chung 1997), function words, adverbial modification, the repeated part of reduplicated forms and non-final location of the word (K. S. Chung, 2006), and vowel coupling (D. C.-H. Li, 2011). I will not discuss these factors here as they are not directly related to the research questions of this dissertation.

1.3 A brief introduction to Rugao and Rugao syllable contraction

1.3.1 The dialect of Rugao

The Rugao dialect, or *Rugaohua*, is the general term that I use for the dialect that is spoken within the City of Rugao. Rugao, located in the Yangtze River Delta on the northern bank of the Yangtze River, is a county-level city under the administration of Nantong, Jiangsu Province, P. R. China. According to data from the official website of Rugao government⁵, the city of Rugao currently has a population of about 1.2 million. Rugao is usually considered a

⁵ The address of the website is: <http://www.rugao.gov.cn/rgsrmzf/sj kf/sj kf.html>. Accessed on January 10, 2020.

dialect of Mandarin within the sub-area of Jianghuai Mandarin; Jianghuai is a dialect of the Lower Yangzi Mandarin group (Li & Thompson, 1989; Li, 1989). More precisely, within the Jianghuai dialect area, Rugao belongs to the Tong-Tai (or Tai-Ru) Jianghuai that generally consists of Tongzhou and part of Nantong including Rugao. Historically a Wu dialect region, with the massive immigration from the northern areas after the Han Dynasty (500 AC), Rugao gradually shifted to a Mandarin dialect region⁶ (Wu, 2006). Phonologically, the dialect of Rugao features complex *n/l* variation, a simple affricate system, heavy vowel nasalization, rhotacization, a six-tone system with two checked tones and so on (Wu, 2006). I will provide a more detailed description of the phonological system of Rugao in Chapter 2.

As the City of Rugao had been an immigrant city in its history with different settlements from different periods of time and people from different regions of origin, the language spoken in Rugao exhibits great regional variation with its own city. Based on Wu⁷ (2006) and the author's observation, there are at least three smaller regions⁸ within Rugao: Rucheng, West (locally called *Xixiang*) and South (*Nanxiang*); each region has their unique vocabularies, vowels and tones that can be recognized by layman speakers to distinguish them from the other two

⁶ The language contact, mingling and evolving make the language spoken in Rugao extremely complicated and even unintelligible for speakers of other Jianghuai varieties. Some researchers would argue that Rugao is a Wu dialect (Ting, 1966). This dissertation takes the stance of treating Rugao as a Jianghuai Mandarin dialect as stated in most dialect studies.

⁷ Wu (2006) further separates South and West and proposes five sub-regions.

⁸ The Rucheng dialect area includes the townships of downtown Rugao and its surrounding towns. The South area is on the bank of the Yangtse River and receives more influences from the Wu dialect area across the River. The West is influenced by the Taizhou dialect, as the region sits on the border between Rugao and the City of Taizhou.

regions. As this dissertation is not a dialect study on Rugao, I will not discuss the details of dialectal differences for the simplicity of this introduction. I choose the variety in Rucheng as the representative of the Rugao dialect, the “downtown accent”, as the main source of data for these reasons. All the transcriptions presented in this dissertation thus reflect the speech in downtown and surrounding townships and may be slightly different from the pronunciations in the other two regions.

It is also worth noting that there are great generational variations and an ongoing sound change in Rugao. According to the author’s observations, the young speakers are diverging from the “traditionally standard Rugaohua” in the pronunciations of all aspects of the language. In particular, some of the consonants and the vowels of the younger generation, the author included, are already observably different from the descriptions in relatively recent work (Huang, 2011; Wu, 2006), as I will present in 20. Furthermore, the author also observed that the young generation, especially those in the North area, are more likely to be contraction users and to contract more than the older contraction users. It is in line with the general trend in the broader Chinese speaking area in China, as young women are more likely to contract in general (Xu & Mao, 2017). As the focus of this dissertation is on syllable contraction, I will not discuss the sociolinguistic variation in the language itself in detail, but I would like to clarify that the transcribed data presented in this dissertation represent the young generation’s speech and may be different from the traditionally defined “Rugaohua”. The phonological descriptions in Chapter 2:1 also reflect the current young generation’s speech.

1.3.2 Syllable contraction in Rugao

Although syllable contraction in Chinese has been explored in much depth and other dialects, such a phenomenon in Rugao has rarely been analyzed. Other than the work of the author herself, the general introductory book on the Rugao dialect (Wu, 2006) is the only work that very briefly mentioned syllable contraction, in which syllable contraction is called *heyin* (合音, “merged sounds”), a term that is commonly used in Chinese linguistics to refer to the phenomenon of reducing syllable numbers and deleting segments.

“如皋话中有些词语由于习惯性的快读和连读音变而合成一个新的音节，常被误认为单音字，但却很难找到这种本字，……” (Wu, 2006)

[There are some words in Rugao that, due to the habitual fast speech and the sound alterations in connected speech, are merged to a new syllable. (These words) are often mistakenly considered as monosyllabic words, but (it is) hard to find the original characters, ...]

From the description above, one can see that Wu (2006) justifies the reason for contracting words as the fast speech and sound alterations in fast speech, similar to other studies of Chinese syllable contraction. He also considers the lexicalized contraction as a result of the same. Below are some example contracted words from this book (Wu, 2006). Note that the original text only includes IPA transcription of the contracted form. The numbers on the syllables indicate tone categories instead of the Chao-style five-degree tones (Chao, 1968). More than one number on the syllable means more than one possible tone. The IPA transcription for the precontraction forms is provided by the author of the dissertation. In (6)—(8), contractible parts are underlined.

- | | | |
|--------|-------------------|-------------|
| (1) 要么 | jɔ5.mə5 → iɔŋ5 | ‘otherwise’ |
| (2) 错啊 | tʃhɔw5.a5 → tʃhɑ5 | ‘wrong’ |
| (3) 需要 | ɕY1.jɔ5 → ɕio5 | ‘need’ |

- | | | |
|---------|---------------------------|-------------|
| (4) 你家 | nei3.ka5 → nia5/3 | ‘your’ |
| (5) 我家 | ŋo3.ka5 → ŋua5/3/ua3/ŋa3 | ‘my’ |
| (6) 儿子啊 | a12.tsɿ5.a5 → a12.tsa5 | ‘your’ |
| (7) 好的哟 | xo3.tej5.jɔ5 → xo3.tiɔ3/0 | ‘OK’ |
| (8) 小伙啊 | ɕjo3.xo3.a5 → ɕjo3. xua5 | ‘young man’ |

This word list above mainly includes disyllabic words or trisyllabic words with a disyllabic contractible part, and most are cases where the contractible part ends with a vowel. There are already some patterns that emerge from this short word list. First, it seems that the first segment of the contractible part is always preserved in the output, with only one exception in one of the variants in (5). Second, all the non-schwa vowels at the very end of the contractible part are also kept in the contracted form, while the schwa in (1) is lost. Third, the medial high vowels (or glides, although the original text uses high vowels exclusively) may survive as much as allowed by the phonotactics of Rugao syllable, as in (3) (7) and (8), and in some cases may even violate the syllable constraint and create illicit syllables, as the *nia5/3 in (4) and the *ŋua5/3 in (5). Fourth, even without the intervening of glides or high vowels, the contracted syllable may still be phonologically ill-formed. For example, although the syllable tsha in (2) does exist in Rugao, the combination of *tsha5 with a high level tone constitutes a lexical gap.

Despite its simplicity and the lack of analysis, the examples provided by Wu (2006) in the word list above provide a valuable first peek into the phenomenon of Rugao syllable contraction, especially with the observable patterns that can be summarized from it. However, the complexity of syllable contraction in Rugao is far beyond this. Many words of other types, for example, contraction of consonant-ending words, e.g., ɕjəŋ55 + jɔŋ55 → ɕjɔŋ55 ‘credit’, complete contraction of trisyllabic words, e.g., kow213 + si21 + a21 → ka2 ‘right?’, and contraction that involves vowel coalescence instead of deletion, e.g., now55 + ɕja55 → nɛa55 ‘those’, are often

observed in Rugao syllable contraction but are simply missing from Wu’s word list. Utilizing a wider variety of data, the patterns of segment preserving as well as the vowel selection process were analyzed preliminarily in previous work of the author based on a smaller set of Rugao syllable contraction data and experimental data (Xu et al., 2018). According to this work, Rugao real word contraction has exhibited some patterns that are also found in other Chinese languages/dialects, including preserving the first and last segments and the seemingly sonority-driven vowel selection. Xu et al. (2018) laid out a foundation for the analyses in this dissertation. The experimental results for the sonority bias in vowel selection in the most recent work (Xu et al., 2018) seems to be inconsistent across vowel groups despite the general trend that is predicted by the vowel sonority hierarchy. This leaves potential for more studies on this topic.

In addition, all previous studies are phonological analysis or phonology-inspired experimental studies, but no work has tackled the phonetic details of Rugao syllable contraction. Phonetic studies on this matter for other Chinese languages such as Taiwan Mandarin (Kuo, 2010) and Cantonese (Wong, 2006) have revealed that the phonetics of contracted syllables can be very informative in helping us further understand the details of this process. A preliminary check on the non-contracted word and its contracted form already suggests that there is much to explore in the phonetics of this phenomenon. For example, as shown in Figure 1 and Figure 2, the non-contracted disyllabic word [zəŋ₂₁₃.ka₅₅] (‘other’) and the contracted word [zɑ₂₁₃] have distinct visualizations. Although the consonant [z] looks quite similar in the two forms, the vowel [a] is visually different. First, the contracted [a] (in Figure 2) is much longer than the lexical [a] (in Figure 1). Is it a real “one syllable” word the same as a monosyllabic lexical word? Second, there seems to be some quality change throughout the duration of the vowel in the contracted [a] but not much in the lexical [a], suggesting that the contracted [a] may not have the

same vowel quality as the lexical [a], either. Finally, is the seemingly deleted vowel [ə] absolutely absent in the contracted form, although the contracted syllable sounds just like a [za] with an elongated vowel? Moreover, all the analyses of Rugao syllable contraction, whether phonological analysis or experimental study, have assumed that the process of syllable contraction is one with segment deletion. But the literature on incomplete neutralization (Port & O'Dell, 1985; Warner et al, 2004) suggests that what seems to be a complete process may in fact be incomplete. More and more questions arise when one delves deeper into the details.

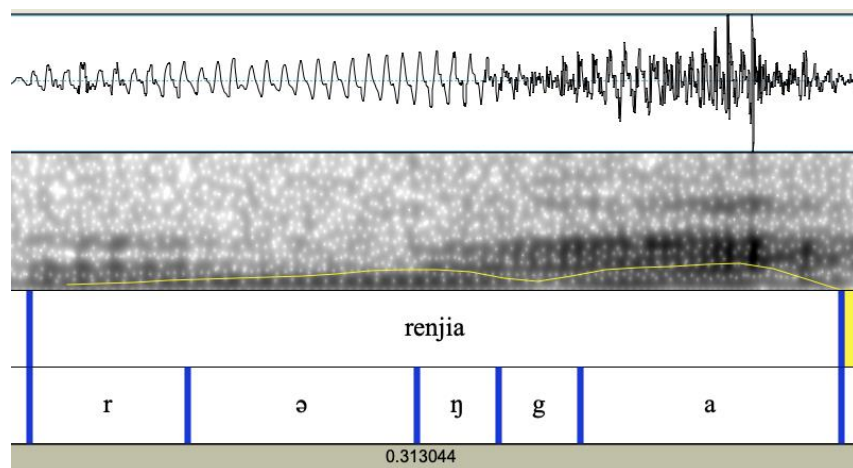


Figure 1. Example of non-contracted disyllabic word [zəŋ.ka]

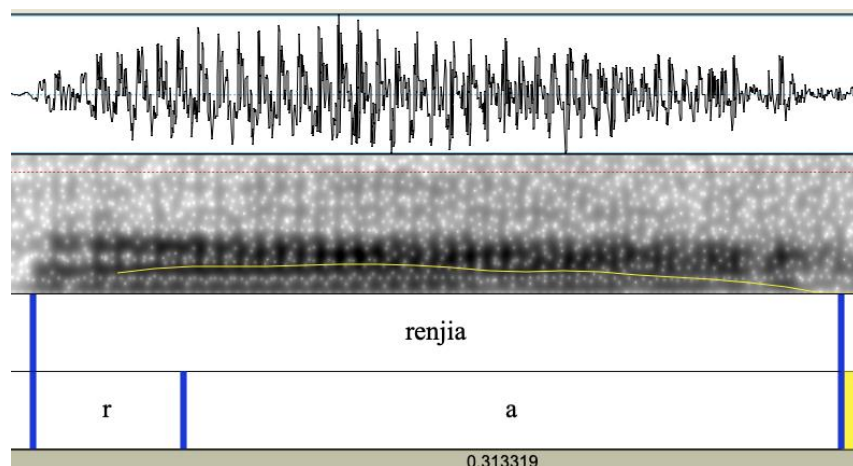


Figure 2. Example of fully contracted word [za]

1.3.3 An overview of the corpus

To have a broader view of the process of syllable contraction, this dissertation utilizes a variety of data, including a corpus that was built specifically for the study of Rugao syllable contraction, the Rugao Syllable Contraction Corpus. All the real-word Rugao contraction data presented in this dissertation are from this corpus, the most of which was collected by the author via field work in Rugao in 2015. There are also some words that are scribed by the author in everyday conversation with peer Rugao speakers from 2014-2018.

A total of six Rugao speakers (aged 28-30 at time of recording) were recorded in groups of three while they casually chatted about various topics of everyday life⁹. All speakers were born and raised in downtown Rugao and were living in downtown Rugao at the time of recording. Among the six speakers, two of them are primary school teachers, three of them work for local banks, and one is self-employed. Although Standard Mandarin is generally required for their conversations at work, all of them reported to speak Rugao frequently in their daily life with their friends and family and whenever and wherever they are not required to speak Standard Mandarin. Two sessions of interview were conducted, each with three speakers who are close friends with each other. This was to ensure they were relaxed enough for casual speech. In each session, the author first explained the task to them as being “to record some casual conversations in Rugao”, and then lead chats about multiple topics of everyday life, including food, clothing, touring, etc. (see full list of questions in [Appendix A](#)). Rugao dialect was used from the very beginning to the end, including introductions and explanations of the task.

⁹ The author is a fluent speaker of Rugao who was born and raised in Rugao and this made it convenient for her to lead the interviews.

A total of approximately 60 minutes of conversation were recorded during the interview sessions, from which 140 instances of contraction were elicited, with different contraction forms of the same word listed as separate entries. The author transcribed all the data using conventional IPA and checked all the transcriptions one month after the initial work; the accuracy rate was reasonably high, with minor changes to only several tokens out of 140. The full list of words can be found in [Appendix G](#).

1.4 An overview of this dissertation

1.4.1 Research questions

Generally speaking, this dissertation is dedicated to investigating the phonological patterns and phonetic details in Rugao syllable contraction. Based on the backgrounds of both syllable contraction in other Chinese languages and my own previous study of Rugao syllable contraction, I would like to expand the previous analyses and dig deeper into the segment selecting mechanisms and phonetic details of this phenomenon while leaving the tonal contraction for future projects.

For the general phonological analysis, I will find out how the syllables are contracted and how each of the surviving segments are selected, including the edge segments and the vowel nucleus. I am most concerned with disyllabic contraction, including contraction of disyllabic words and contraction of the disyllabic part of the trisyllabic word. In particular, how the contraction of two syllables into one happens is the major concern, as the disyllabic contraction is most common and most compatible with the existing models.

Among the segments, I am focused on how vowels behave in the process of contraction when two vowels are in a competition for survival. Although the consonants also undergo

contraction, the deletion or survival of consonantal segments is more predictable from the linear order. Vowels, however, undergo phonological processes such as vowel competition/selection, glide formation, and vowel deletion. The outputs of contracted vowels may also be different from the pre-contraction lexical forms. For these reasons, once the general phonological patterns have been established, I will look deeper into the vowels from different perspectives. To study the vowels in the contraction process, I will first examine the details of the vowel selection in competition for survival and find out whether the linear order of the two vowels (R.-F. Chung, 1996, 1997) or the sonority of them (Hsu, 2003) play a (more important) role in shaping the contracted output. A deep investigation on the vowel nucleus selection process will be conducted with an emphasis on the mechanisms behind the outputs of the vowels in syllables contracted from di-syllabic lexical words. The analyses and discussion of the relationship between vowel nucleus selection and vowel sonority will then facilitate a discussion of the general concepts of vowel sonority and syllables (Parker, 2011, 2012). After the vowel selection mechanism is explored in detail, I will shift the focus to the phonetic details of the contracted output and compare the contracted vowels and their corresponding lexical counterparts with regards to the phonetic cues, including syllable and vowels durations and the formant information.

Alongside the study of syllable contraction, some broader-issue questions will be raised and explored, including how vowel sonority may shape the outcome of vowel competition and how phonological knowledge is represented. With the exploration of these deeper issues, I will connect the different parts of the dissertation and provide my answers.

As a summary, the research questions of this dissertation are as follows:

- What are the phonological patterns of Rugao syllable contraction? In particular, what determines whether a segment survives or deletes in the contraction?

- In contraction cases where two vowels compete for survival, what determines the winner and the loser?
- Is the winning vowel of the competition in the contraction form still the same vowel? More precisely, is the vowel in the contraction form the same as a lexical vowel?
- Finally, what do the vowel competition and vowel deletion mean for phonological representation?

1.4.2 Methodologies

To answer the research questions, I will employ various methodologies for this study, including the classic phonological analysis, templatic analysis and experimental techniques. For the phonological analysis, I will explore the Rugao data, summarize the observable patterns and fit the patterns to the existing models that have already been proposed for the syllable contraction in other Chinese languages. Following Chung (1996, 1997) and Hsu (2003), I will first examine the Rugao data from the skeletal tier perspective using the templatic analysis in order to find the basic patterns. In particular, I will modify Chung and Hsu's models and adopt the Edge-In Association (Yip, 1988) to account for the "edge" segment preservation in Rugao. For the analysis of the non-edge segments, I will explore whether there is a linear-order preference (R.-F. Chung, 1996, 1997) or a sonority preference (Hsu, 2003) in the corpus data.

As the corpus only affords limited data with limited vowel combinations, full studies that investigate more combinations of vowels will be achieved in experimental settings. First, I will introduce a forced-choice task experiment, in which phonologically well-formed nonsense words with different vowel combinations are used to test speakers' vowel choice. The stimuli are

specifically arranged so that two hypotheses, the linear-order bias and the sonority bias, are tested. Within the sonority realm, I will explore in more depth along the different dimensions of vowel sonority, namely height and centrality, in order to test whether the sonority bias is consistently confirmed. Then I will introduce another two experiments, both production experiments of similar settings. These two experiments are conducted following the discussion of what the survivor of vowel competition really is. The contracted vowels will be elicited in strictly controlled phonological environments. I will then measure the duration and formants (F1, F2) of the contracted vowel and its corresponding lexical vowel and compare the results, aiming to find similarities and differences between the two.

1.4.3 A roadmap of the dissertation

As already seen in this chapter, this dissertation starts with an overview of the phenomenon of syllable contraction in Chinese languages and a brief discussion of contraction data in different languages. The relevant factors affecting the likelihood of contraction and the output forms, including the sociolinguistic, phonological, phonetic influences, have been discussed.

The main body of this dissertation will consist of two major parts. In the phonology part, I will analyze the phonological patterns and in particular, the vowel selection process and how it is related to vowel sonority and/or the linear order of pre-contraction forms. The phonetics part will include phonetic measurements of different acoustic aspects of contracted syllables and vowels.

Beginning with an introduction to the phonological system of Rugao, the phonological patterns of Rugao syllable contraction in the corpus data will be explored and analyzed in Chapter 2. These analyses will foreshadow the two major research questions for Chapter 3 and

Chapter 4, respectively, i.e., 1) which vowel survives in the vowel competition, and 2) whether Rugao syllable contraction is a case of complete or incomplete vowel deletion. In Chapter 3, the nonce word experiment on vowel selection as well as the role that vowel sonority plays will be fully discussed with the data. Chapter 4 will be an investigation on the phonetic properties of syllable contraction, mainly contracted vowels and their corresponding non-contracted forms and comparisons of the two. Finally, in Chapter 5, I will bring together the phonological and experimental analyses, the results, along with some discussions that are more theoretically motivated. In particular, I will explore the theoretical implications on various topics, including vowel sonority, phonological representation, and the relationship between phonetics and phonology. The shortcomings of the analyses in the dissertation and remaining issues will be also fully discussed in this chapter, with or without offering immediate solutions. In the end, a short summary of this dissertation and an overview of possible future projects stemming from this study will conclude the dissertation.

2 PHONOLOGICAL PATTERNS

2.1 Introduction to the Rugao dialect

In this sub-section, I will briefly introduce the phonological system of Rugao based on the descriptions in Wu (2006)¹⁰. Note that due to the areal and generational variation, I will update some details according to the speech of the current young generation (roughly 18-35 years of age) living in the Rucheng dialect division. After the basics of the Rugao language are laid out, I will present the phonological analysis of Rugao syllable contraction based on the corpus data.

2.1.1 Consonants

According to Wu (2006), Rugao has 21 *shengmu*, syllable initials, summarized in Table 1 below. Note that “ø” marks “zero initial”, the name for the initials of syllables without a pronounced Onset. All stops and affricates in Rugao are voiceless, and have an aspiration contrast. The aspirated consonants are marked with the aspiration superscript ^h. It is obvious that the initials listed in Table 1 are based on the traditional belief in Chinese linguistics that Chinese syllables are a combination of an initial and a rhyme, treating glides as part of the rhyme (R. Li, 1989; Ting, 1966). Accordingly, this system allows zero initials and treats all other segments as a

¹⁰ Wu (2006) is a general dialect book for general audience reading rather than for academic purposes. Even if the descriptions in the book were accurate, the information is outdated, as noted by the author of this dissertation.

single unit, rhyme. As rhymes with a different glide or consonant ending are treated as different, there are a total of 48 different rhymes¹¹ in Rugao, according to Wu (2006).

p	p _h	m	f	v
t	t _h	n		l
ts	ts _h		s	ʒ
tɕ	tɕ _h		ɕ	
k	k _h	ŋ	x	
∅				

Table 1. Rugao initials (Wu 2006:58)

This dissertation adopts the broader view of syllable and treats Rugao syllables as having a basic C(G)VC syllable structure as in other Chinese languages/dialects (Duanmu, 1990; Y.-H. Lin, 2007). In this structure, the first Consonant(s), whether a glide or non-glide consonant, is the Onset; whether the syllable has an Onset is optional, allowing zero Onset syllables such as /ɔŋ/. The Vowel + Consonant constitute the Rhyme, with the Vowel in the Nucleus and the Consonant in the Coda. On the basis of Wu's (2006) work and the C(G)VC syllable structure, I propose a 24-consonant system for Rugao, including all 20 non-zero initials already listed in Table 1. All the consonants are summarized in Table 2. Note that in Wu's (2006) system, only /ŋ/ is considered licit post-vocalically¹²; I added other consonants that are allowed in the Coda to the system, including:

¹¹ I will not list all 48 rhymes here as they're not directly related to the projects of this dissertation. Interested readers are referred to Wu (2006), pages 58-59.

¹² According to the author's observation, the velar nasal coda is being replaced by alveolar nasal codas unless preceded by a back vowel. There is much phonologically unconditioned variation between [n] and [ŋ] observed within and across speakers. The nasal merger is also reported in other Jianghuai varieties, such as in Yancheng (Cai, 2011). As this process is incomplete, in this dissertation, I use /n/ or /ŋ/ to transcribe the nasal codas according to the speaker's particular place of articulation for these nasals.

- I. Glottal stop /ʔ/, which is pronounced in the Coda for checked syllables¹³, e.g., /tiʔ/ and optionally emerges in the Onset for the “zero initial” syllables.
- II. Retroflex /ɭ/, e.g., /aɭ/.
- III. Glides /j, w, ɥ¹⁴/. According to the C(G)VC syllable structure, I also treat pre-nucleus and post-nucleus [+high] vocoids /j, w, ɥ/ as glides, i.e., consonants (Duanmu, 1990).

p	p ^h	m	f	v		w	ɥ
t	t ^h	n		l			
ts	tʂ ^h		s ¹⁵	ʐ	ɿ		
tɕ	tɕ ^h		ɕ			j	
k	k ^h	ŋ	x				
ʔ							

Table 2. Rugao consonants

There is not much allophonic variation observed for Rugao consonants, as most of the variation is not phonologically conditioned. But there are three notable rules, as shown below. Note that speakers of different sociolinguistic backgrounds may be vastly different in having these rules or not. I leave the details for future projects.

¹³ Checked syllables are those with checked tones, which are usually shorter than non-checked correspondence and with a stop in the coda.

¹⁴ The distribution suggests that [ɥ] may be an allophone of /w/. I leave them as two separate phonemes as traditionally done for now, but I am open to revise it given a full analysis.

¹⁵ For some speakers, /s/ is sometimes pronounced as a retroflex [ʂ] for words that are pronounced with a retroflex initial in Mandarin, possibly due to the influence of Standard Mandarin, e.g. /swej/ → [ʂwej]. I do not list /ʂ/ as a phoneme in Rugao for such inconsistency.

- For /ɹ/: [ɹ] is only pronounced in the Coda, e.g., [aɹ, əɹ, ɔɹ], while a fricative [ʒ] is only pronounced in the Onset, e.g., [ʒiʔ], [ʒæɪn].
- For some speakers, the alveolar sibilants /ts/, /tʃ/, /s/ may undergo palatalization¹⁶ when preceded by a [+front] vocoid, whether a high vowel or a glide, and be realized as [tɕ], [tɕʰ] and [ɕ], respectively, e.g., /tswen/ → [tɕwen], /tʃhiʔ/ → [tɕhiʔ], /sjen/ → [ɕjen].
- For some speakers, alveolar stops /t/ and /tʃ/ can palatalize before /i/, e.g., /ti/ → [tɕi] /tʃi/ → [tɕʰi].

On top of those, there are phonotactics that specify the possible combinations of glides and vowels. Glide /j/ is freer in combining with the vowels in the pre-nucleus position, in which it can proceed a variety of non-high vowels /ja, jæ, je, jɛ, jo, jɔ, jə/. Similarly, glide /w/ can combine with non-back vowels in the pre-nucleus position /wa, we, wɛ, wə/. The location of /ɥ/ is more restricted than the other two glides. Pre-vocalically, /ɥ/ can be followed by non-front vowels, such as /ɥa, ɥu, ɥə, ɥɔ/. /ɥ/ can only follow non-sibilant alveolar consonants and post-alveolar sibilants. Glides are generally not allowed after the vowel nucleus with two exceptions: /ej/ and /ow/. Whether these off glides are just gestures needs further study, but I treat these combinations as /ej/ and /ow/ for this dissertation.

¹⁶ The palatalization of the alveolar consonants is only for a portion of the speakers, and I am not sure how great the proportion is or what the sociolinguistic distribution is other than the rough age group. It is also a possible ongoing sound change. Further studies can provide evidence for these phonological alterations, but I would like to be conservative and list the traditionally used initials in the consonant tables.

***jV[high]**

/j/ can only precede non-high vowels.

***wV[front]**

/w/ can only precede back vowels.

***ɥV[front]**

/ɥ/ can only precede non-front vowels and cannot appear post-vocalically.

2.1.2 Vowels

The vowel system proposed here is also based on Wu (2006), but different from it in many ways. First, as Wu (2006) only lists the velar nasal post-vocalically, other cases of nasality are treated as nasal vowels, e.g., /ã/, /ẽ/. I suggest that these two nasal vowels are redundant for the vowel system as they are in complementary distribution with the oral counterparts /a/ and /e/. I will unify /a, ã/ and /e, ẽ/ respectively, and this way, nasal vowels only surface when there is an underlying nasal coda. Second, as sound change is evident especially for the vowels, I updated the vowel system to reflect the current way of pronunciation. Finally, the pre-nucleus and post-nucleus /i, u, y/ are treated as glides and are not all included in the vowel system. I adopt the view that all the traditionally regarded diphthongs are a vowel (monophthong) followed or preceded by a glide (Duanmu, 1990), so no diphthong is listed here. Table 3 lists Rugao vowels at the phoneme level. Most of the vowels here are similar to those in other Mandarin dialects except the lax high-front rounded /ʏ/, which is usually tense /y/ in other Mandarin varieties such as Standard Mandarin (R. Li, 1989).

	Front	Central	Back
High	i ʏ		u ¹⁷
Mid	e ɛ	ə	o ɔ
Low		a	

Table 3. Rugao vowels—phoneme level

Based on this 9-vowel system, the allophonic variations are described below, and all vowels are summarized in Table 4.

- /a/: /a/ is now pronounced with two variants in Rucheng¹⁸: [a] and [æ]. [æ] only proceeds nasal /n/ and glottal stop /ʔ/, e.g., [pæn], [læʔ], while [a] appears in open syllables and before /ɿ/, e.g., [pa], [na], [saɿ].
- /i/: /i/ has three variants: [i], [ɨ], [ɿ]. The distribution is: [i] is only in closed syllables, followed by either a nasal or glottal coda, e.g., [tin], [phiʔ]; [ɨ]/[ɿ] are only in open syllables, for which apical [ɿ] only follows the alveolar sibilants, e.g., [sɿ], [tsɿ], while [ɨ] only follows consonants other than alveolar sibilants, e.g., [mi], [tɕi].
- /u/: /u/ has three variants: [u], [ʊ] and [ʊ]. Similar to many other allophonic cases, [u] appears in open syllables, e.g., [pu], [nu], while [ʊ] must be followed by a nasal or glottal coda, e.g., [tsun], [phuʔ]. The [ʊ] is pronounced following labial-dental consonants /f, v/.
- Traditionally /ʊ/: According to the author's observation, the young generation is replacing /ʊ/ with the following two options:

¹⁷ /u/ used to have two variants [u] and [ʊ], but [ʊ] is lost in the younger generation's speech according to the author's observation. Older speakers still pronounce [u].

¹⁸ Note that there is no such variation in West or South.

- Merge with the existing phoneme /u/ in closed syllables or after /ɥ/, such as in /ɥʊ/ → /ɥu/, /tɕyʊn/ → /tɕyɯn/, /ɥʊn/ → /ɥɯn/.
- Replace with the new mid vowel /o/, sometimes realized with an off glide /w/, e.g., /pho(w)/, /lo(w)/, /kho(w)/.

	Front	Central	Back
High	i ɿ ʏ	ɨ	u ʊ ɤ
Mid	e ɛ	ə	o ɔ
Low	æ	a	

Table 4. Rugao vowels—including all allophones

2.1.3 Tone

There are four or six tones in Rugao depending on how the status of the checked tones are treated. The majority of literature agrees on the six types of tones in Rugao (Huang, 2011; Ting, 1966; Wu, 2006), including four non-checked tones and two checked tones, as listed below:

- *Yin Ping* (falling), HL, marked as X₂₁
- *Yang Ping* (rising), LH, marked as X₃₅
- *Yin Shang* (low), LL, marked as X₂₁₃
- *Yin Qu* (high), HH, marked as X₅₅
- *Yin Ru* (short high), hh, marked as X₅
- *Yang Ru* (short rising), lh, marked as X₃₅

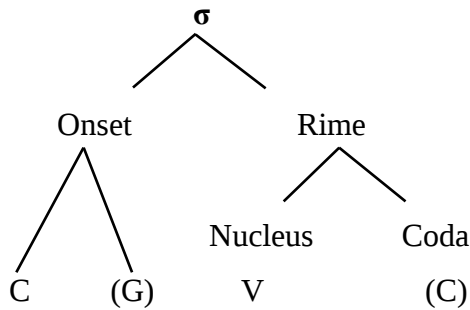
Note that the italicized words are traditionally used terms for marking tones in Chinese linguistics, and that H and L are used to represent the high end and low end of the pitch range of

the onset and offset of the tone. The lower-case h and l are also for this function and indicates a checked tone. I also listed the five-scale annotation in Chao-style (Chao, 1968) in numbers, which is used for all tone annotations in the dissertation, for the ease of reading of the data. “1” indicates the lowest on the pitch scale, and “5” indicates the highest; “X” means any tone bearing unit, namely, any syllable. The checked tones are all short compared to the non-checked tones and must appear in syllables with short vowels and a /-ʔ/ coda. This dissertation follows the six-tone system and distinguishes all the tones, marking them using different combinations of numbers. Note that as tone is out of the scope of this dissertation, I may not provide tone information if it is not directly related to the issue or data being discussed in the following analysis.

2.1.4 Syllable

As a Chinese dialect, Rugao is similar to other Chinese languages in imposing strict limits on syllable structures. Generally, a C(G)VC type syllable is most commonly seen in Rugao. The maximum size of the syllable consists of C-G-V-C, while all the segments are optional except the Vowel Nucleus. Possible syllable structures include V, VC, CV, CVC, GV, GVC, CGV, and CGVC. The constituents of syllables are shown below in 1).

1) Syllable structure of Rugao



As I do not treat vowels with quality change as diphthongs, all syllables have a Nucleus that is a monophthong from Table 3 or Table 4. Diphthongs with an off-glide are treated as rhymes with a monophthong in the Nucleus and a glide in the Coda. Those with a pre-Nucleus glide are treated differently: I put the Glide as part of the Onset, based on some evidence from syllable contraction: crucially, the pre-nucleus glides do not participate in the vowel selection process, but instead directly get preserved or deleted along with the preceding consonant, depending on the position. See more details in 2.2 and 2.3.

2.2 Edge-In Association

2.2.1 What syllable contraction is and how it happens

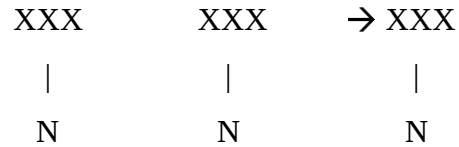
Now that the phonological system of Rugao is established, I will devote the rest of this chapter to the analysis of Rugao syllable contraction using the data from the Rugao Syllable Contraction Corpus (hence, the corpus). Note that, unless specified, only full contraction forms are presented below while the partially contracted versions are not listed. More partial contraction data points can be found in the full list of words from the corpus data in [Appendix G](#).

Before delving into the details of syllable contraction, it is a necessary first step to discuss what the process of syllable contraction means and how this process happens. As briefly

discussed in the Introduction, different researchers may have different views on syllable contraction, and many factors can contribute to this process. As the focus of this dissertation is on the “clearer” cases of contraction, it is to the benefit of the analyses to adopt a more clear-cut analysis of syllable contraction. Below I will introduce the two analyses by Chung (1996) and Wee (2014), both of whom have full analyses on the process of syllable contraction.

Chung (1996) proposed the skeletal tier account for Taiwanese Southern Min syllable contraction, which is essentially a templatic analysis. This analysis first assumes that Taiwanese Southern Min syllables have a basic three-slot template, XXX. Each X represents a time slot, with the Nucleus right in the middle. Contraction of a disyllabic word to a single syllable can thus be viewed as six X-slots merging into three slots, as illustrated in 2). The reduced number of slots means that some segments in the six-slot form, i.e., the pre-contraction form, must lose or delete, while the others get to be preserved and take the remaining three slots. Chung (1996) further proposes that in the template, the association of melodies to the templates starts with the two edges, which means the first segment of the first syllable and the last segment of the second syllable must survive in the contracted syllable. This type of association also predicts that the inner segments, namely, the last segment(s) of the first syllable and the first segment(s) of the second syllable, are most vulnerable for deletion. This so called Edge-In Association (Yip, 1988) will be discussed in more detail in the next sub-section. Hsu (2003) further expanded Chung’s templatic account (R.-F. Chung, 1996, 1997), but pointed out the problems of the three-X template analysis, especially placing all the vocoids in the middle X-slot, which makes many of the Southern Min data unexplainable. While proposing a sonority-based model, she also modified the template analysis and added that limiting the Edge-In to only consonants will solve some of the problems of the initial proposal.

2) Syllable contraction of Taiwanese Southern Min ($N = \text{Nucleus}$) (Chung 1996)



Wee (2014), on the other hand, adopted *Trunc(ate)* (Kager, 1999) to account for some of the patterns in Tianjin trisyllabic casual speech elision, a similar term to syllable contraction. Basically, *Truncate* means to merge into a single syllable. He calls the adjacent three syllables a string with two Windows, each Window being where the two syllables connect. As illustrated in Figure 3, the phonological material, the medial consonants at the syllable edges in particular, is elided. Without the intervening consonants, the initial and medial syllables, the Nuclei in particular, merge to form a new syllable. According to Wee (2014), the pattern that involves deletion of the consonants in the Coda of the first syllable and the Onset of the second syllable, as well as the coalescence of the two vowels, is most productive in Tianjin.

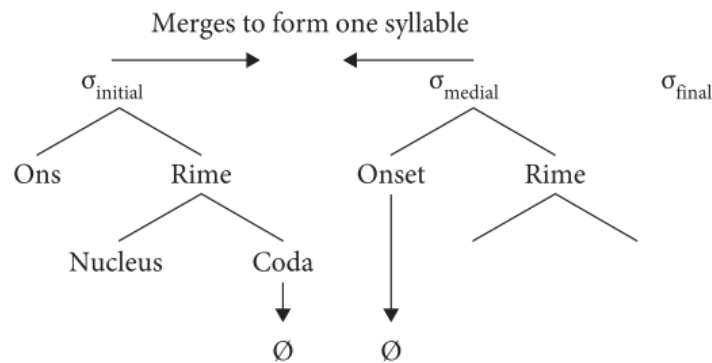


Figure 3. Tianjin syllable elision (adopted from Wee (2014))

What the skeletal tier account and the elision account have in common is that both analyses can successfully predict the decreased number of syllables and the loss of the inner segments, especially the inner consonants. Both analyses agree that the deletion of medial consonantal segments facilitates the contraction, but the skeletal tier model (R.-F. Chung, 1996,

1997) works better for full contraction that involves vowel deletion, while Wee's (2014) model is more efficient in allowing for partial contraction in which two vowels merge and produce another vowel. Based on these, I will combine the analyses (R.-F. Chung, 1996, 1997; Hsu, 2003; Wee, 2014) and propose the analysis for the Rugao syllable contraction based on the basic templatic analysis.

Before I show the X-slot structures, I first propose that the contracted syllables should be treated differently from a lexical syllable because the contracted syllable may have a different syllable structure. This proposal is based on the fact that the contracted syllables are different from lexical syllables in many ways including segment qualities and durations and allow phonotactically illicit syllables to surface. If the phonotactics have such freedom to be violated in the contracted syllables, the syllable structures of them should as well.

- The lexical syllables must have a pre-specified syllable structure as seen in 3);
- The contracted syllables are free to be fitted into a wider range of templates/structures, as described in more detail in 4).

First, a Rugao syllable has a basic C(G)V(C) structure (C=consonant, G=glide, V=vowel), with the glide part of the Onset, as shown in 1) in the previous subsection. I propose that *lexical* Rugao syllables have a basic three-X structure on the skeletal tier with the Nucleus in the middle, as shown in 3). The X-slot, essentially the time tier, can be taken by more than one segment and one segment may take more than one slot. This is based on the CGVC syllable structure in Rugao and the fact that the duration of segments varies depending on their constituent in the syllable. For syllables that have Onset, the pre-nucleus glides are very short while the post-nucleus glides are relatively longer. However, when a syllable starts with a glide, this glide is almost as long as a vowel. In addition, the same vowel can be longer in open

syllables than in closed syllables. The longer segments take two slots, while two shorter segments may share one slot.

3) *Rugao lexical syllables*

n	u	ŋ	khw	ε	p	a	xw	e	j	te	i	ʔ	j	o	w
			∨	∧		∧	∧								
X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X
	N			N		N		N			N			N	

- Onset consisting of one single non-glide consonant takes the first X-slot. Examples of this type are shown in /nuŋ/ and /teiʔ/ in 3).
- Pre-nucleus glide is part of the Onset, and takes the first X-slot, such as in /jow/.
- Complex Onset takes a single X-slot, such as in /khwε/ and /xwej/.
- Vowels in the open syllables may take two X-slots, e.g., /khwε/ and /pa/.

Crucially, the contracted syllables are different from the lexical syllables in that the time slots are freer with regards to what type of segment can take it. 4) are two examples of two different types of contraction.

4) *Rugao syllable contraction*

j ə ʔ	+ j æ n	→ j æ n
X X X	X X X	X X X
N	N	N

l o w	+ k _h w ε	→ l w ε
	∨ ∧	
X X X	X X X	→ X X X
N	N	N

- The contracted syllable can have the same structure as a lexical syllable when the contracted form is close to a regular syllable. For example, for $jəʔ + jæn \rightarrow jæn$, when contraction happens, the remaining three segments, j , ϵ , n each takes a slot, in which the vowel ϵ takes the middle Nucleus.
- The contracted syllable can have the Nucleus in the final X-slot, allowing for the pre-Nucleus glide to take a separate slot. For $low + khw\epsilon \rightarrow lw\epsilon$, the contracted form is different from a lexical syllable in that each segment takes an X-slot, with the Nucleus taken by the surviving vowel $[\epsilon]$.

2.2.2 The Edge-In Effect

As briefly discussed in the Introduction, the Edge-In effect is the most consistently found pattern in the contraction data of Chinese dialects that have been studied. For example, in Southern Min, the data below (obtained from Hsu 2003) shows that the leftmost and rightmost segments must be preserved in the contraction whether they are consonants or vowels. In the first two examples, both the initial and final segments are consonants. In the last two examples, the first segment is a consonant, but the second syllable ends with a vowel. In all cases, the first and last segments of the disyllabic word are preserved in the contracted output.

- | | |
|--|------------|
| (1) ${}_{19}sio{}_{20} + kan \rightarrow sian$ | ‘the same’ |
| (2) $ho + guan \rightarrow huan$ | ‘by us’ |
| (3) $ho + gua \rightarrow hua$ | ‘to me’ |

¹⁹ Note that I use (1), (2), (3), ... to number data, and 1), 2), 3), ... to number analysis-related captions.

²⁰ Tone information was not supplied in the original text of Hsu (2003). The same holds for the following cited data.

(4) toʔ + ui → toi ‘where’ Taiwanese Southern Min, from Hsu (2003)

In other Chinese dialects, such as Cantonese (Wong, 2006), Tianjin (Wee, 2014), Taiwan Mandarin (Tseng, 2005a, 2005b), and other Mandarin dialects (Sun, 2014), although the authors may or may not specifically mention the term Edge-In, the outmost segments are preserved, whether consonants or vowels. A few examples below illustrate this point.

(5) kheɪ + sət → khɛt ‘in fact’ Hong Kong Cantonese, from Wong (2006)

(6) tsi: + tou → tsi:u ‘know’ Hong Kong Cantonese, from Wong (2006)

(7) tsow + ʂa → tʂua ‘do what’ Lunan, Shandong Province, from Sun (2014)

(8) ken + tɕian → kan ‘front’ Pengzhou, Sichuan Province, from Sun (2014)

(9) mien + faŋ. tʂhaŋ → mian.tʂhaŋ ‘cotton factory’ Tianjin, from Wee (2014)

(10) tɕhi + ɕaŋ. thai → tɕhiaŋ. thai ‘weather station’ Tianjin, from Wee (2014)

In the three-X template analysis, when the six-slot disyllabic word contracts to a three-slot monosyllabic word, the association of melodies to the X-slot begins with the two edges (R.-F. Chung, 1996, 1997; Hsu, 2003). The Edge-In Association is initially proposed in Yip (1988) for the discussion of the direction of association in the affixation processes. The most crucial part of Edge-In is the *Anchoring*, which specifies that the outermost melodic elements should be associated to the outermost skeletal slots, in a one-to-one fashion, and this applies the same way to both stems and affixes. For the remaining elements, Yip (1988) proposed *Filling*, which means the association of the remaining melodies and the remaining slots should undergo the same process as *Anchoring*. Both *Anchoring* and *Filling* allow *Template Satisfaction* (McCarthy & Prince, 1994) which allows for language-specific templatic requirements. Chung (1996) first brought this idea into the analyses of syllable contraction in Taiwanese Southern Min to account for the fact that most of the contracted words maintain the leftmost and rightmost segments of

the precontraction form, a pattern that is commonly documented and observed in the Chinese dialects. Meanwhile, there is almost no evidence for a simple LR Association or RL Association, which essentially predicts that, for a disyllabic word, segments from either the left or the right syllable have the priority to survive in the contracted form.

The corpus data suggests that Rugao syllable contraction is quite consistent in keeping the segments on both the left and the right edges. Below in (11)–(16) 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. Note again that tones are not the focus of this dissertation even though the data points are marked with tone.

- | | | |
|------|---|-----------|
| (11) | <u>si</u> ₃₅ + hej ₅₅ → sej ₃₅ | ‘time’ |
| (12) | nun ₂₁₃ + xow ₅₅ → now ₂₁₃ | ‘warm’ |
| (13) | tsow ₅₅ + sən ₂₁₃ → tsun ₅₁₂ | ‘do what’ |
| (14) | jəʔ ₅ + jæn ₅₅ → jæn ₅₅ | ‘same’ |
| (15) | pəʔ ₅ + jɔ̃ ₅₅ → pjɔ̃ ₅₅ | ‘do not’ |
| (16) | ka ₂₁ + tɕhY ₂₁ → kaq ₂₁ | ‘go home’ |

On the right side, as seen in the examples above, all the final consonants in the disyllabic form are kept, whether a nasal or a glide. Generally, for Rugao disyllabic words, if a consonantal segment is present on the right edge, whether a glide, a nasal or a glottal stop, such a segment most likely survives, unless other phonotactics come into play and alter the final output.²¹ In the

²¹ These rules include some post-contraction process such as final -ʔ deletion, nasal POA alternation, etc.

data presented in (11)—(16), all the right-edge segments in the pre-contraction form (double-underlined) are faithfully maintained as the right-edge segments in the contracted form.

When the right edge is a vowel, the vowel must survive and gets the last X-slot, while allowing other vowels as well as consonants to delete or form glides. This type of contraction can be as clean as in (15) $pəʔ + jə \rightarrow pjə$, where the right-edge vowel /o/ survives as the rightmost segment in the contracted form. Because of the presence of a glide /j/, the Glide Formation rule applies, which satisfies Maximal Syllable and Maximal Segment.

5) *Glide Formation*

Search for glides or [+high] vowels and form a glide as much as allowed by the syllable.

6) *Maximal Syllable*

In the contracted syllable, keep the maximal syllable structure.

7) *Maximal Segments*

In the contracted syllable, keep as many segments as possible as allowed by Maximal Syllable and the phonotactics of the language.

In (16) $ka + tɕhɪ \rightarrow kaɪ$, the disyllabic word ends with a vowel /ɪ/. At the right edge, /ɪ/ must survive. /ɪ/ as a high vowel is thus realized as a glide [ɪ] which is shorter and weaker than vowel [ɪ]. In this way, the Maximal Segments rule is also obeyed by preserving both vowels in the contracted output. The syllable is also kept maximal (Maximal Syllable). However, as glide /ɪ/ is typically not allowed in the Coda, the contracted output creates an illicit syllable * $kaɪ$. In fact, the contracted output is often an illicit syllable, which anchors my previous claim that the contracted syllables are not regular syllables. I will discuss this issue later in this chapter in 2.4.2.

On the left side, the data points above show that the left-most consonantal segment must be preserved. All initial consonants in the words above, /s, n, ts, j, p, k/ (marked as single-

underlined), survive in the contracted syllables. Rugao does not have many Onset-free syllables to begin with. Many that are traditionally viewed as “initial free” syllables actually begin with glides /j, w, ɥ/. Such words are rarely found in the Rugao syllable contraction corpus, nor do they appear in any other data sources. Even for the words that begin with vowels, i.e., words with the V or VC syllables, the vowel still survives as it is or alters to a glide. The following pair of words in (17) and (18) are a CVC+CVC word and a VC+CVC word with similar segments. Like many other words, the word that means ‘same’ has two pronunciations, the everyday, casual speech accent [jəʔ.jæn], and the reading accent²², [iʔ.jæn]. The biggest difference between the two is that [i] is a syllable that starts with a vowel. The two pronunciations have very similar contraction forms, [jæn] and [iæn]. I use [i] for the latter as the first segment may be slightly longer.

(17) [jəʔ₅₅ + jæn₅₅] → [jæn₅₅] (‘same’, *non-reading accent*)

(18) [iʔ₅ + jæn₂₁] → [iæn₅₁]/[jæn₅₁] (‘same’, *reading accent*)

A take-away from the Edge-In analysis is that both the pre-nucleus and post-nucleus glides participate in the Edge-In process and can easily break away from the vowel nucleus, providing the motivation for treating the combination of a monophthong and a high vocalic segment as having a CG or GC structure instead of a diphthong (see descriptions in 2.2). The contraction process essentially breaks the CG or GC combination. For example, in *ɛjæŋ* + *xə* → *ɛja* (‘rural’), the glide [j] in the contracted form [ɛja] only possibly comes from the glide in the first syllable [ɛjæŋ]. The fact that /j/ can stand alone with the consonant /ɛ/ but without the vowel /æ/ suggest that the traditionally regarded diphthong /ia/ or even /iã/ (Wu, 2006) is easily

²² The reading accent, *wendu* in Chinese, is the pronunciation that is used for the formal reading of literature. Each dialect has its own system of *wendu* pronunciation, but all are heavily influenced by Standard Mandarin.

breakable and not a real diphthong. Treating the high vocoids as glides successfully avoids any complication that stems from diphthongs: as the Edge-In starts, the consonantal segments at the edges will associate to the template first. The vowel selection follows immediately, and then Glide Formation applies and keeps the glide /j/. Different from the previous analysis of Edge-In Association (R.-F. Chung, 1996, 1997; Hsu, 2003), the Nucleus does not have to be right in the middle of the three-X time slot for the contracted syllable. Each time slot can take more than one segment, and a segment can take more than one time slot. As shown in 8), the contracted syllable, although still taking three slots, has the Nucleus placed in the third slot.

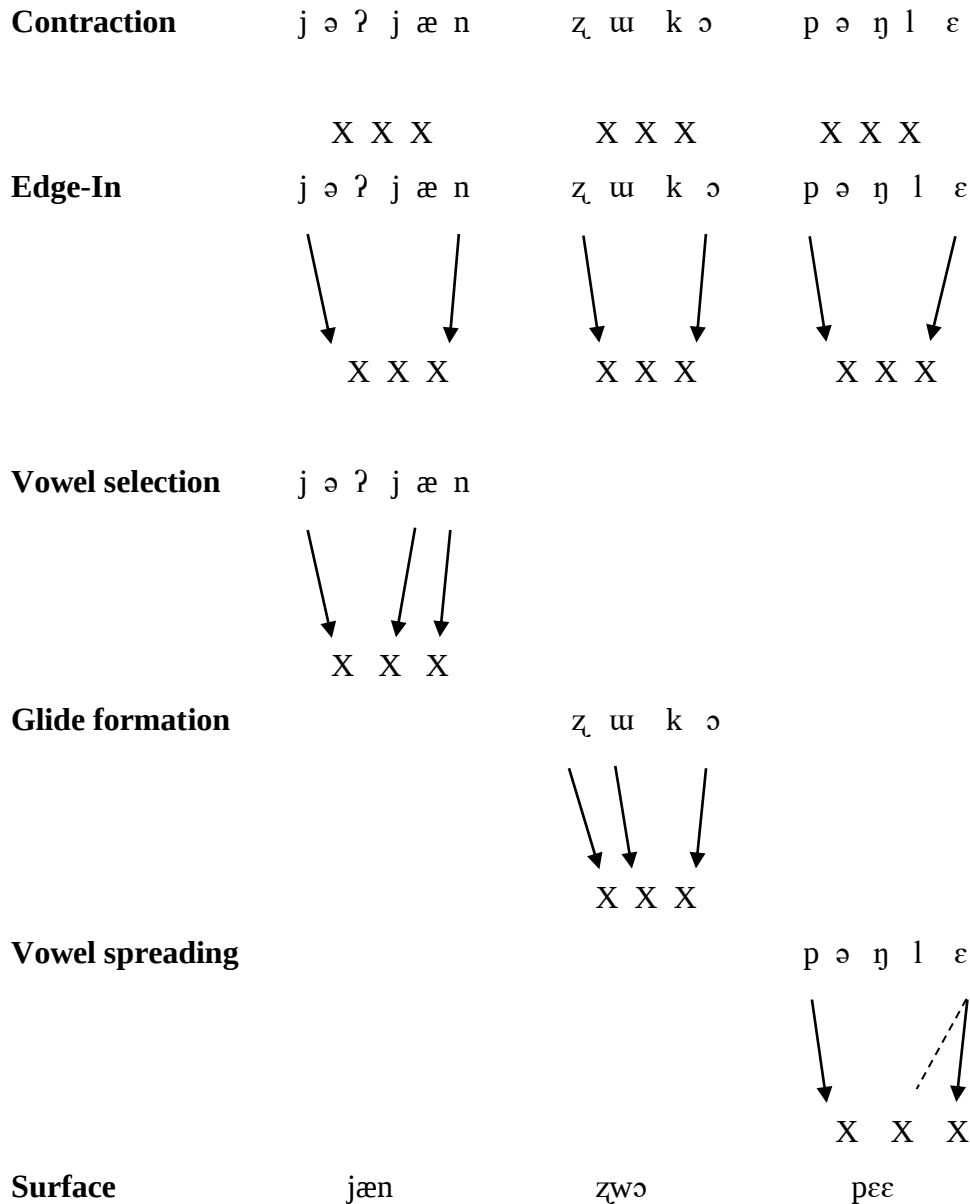
8) *Edge-In Association in Rugao*

ɛj æ n	x a	→	ɛ j a
∨	^		
X X X	X X X		X X X
N	N		N

As an interim summary, the Edge-In Association in Rugao is a strong predictor for the outer segments of the contracted syllable. Detailed derivations for three words of three different structural types are given in 9), aligned side by side. The differences in the Edge-In processes then make the derivation slightly different. Note that if a rule does not apply in a specific step for a word, that step is left blank.

9) *Derivations for jəʔ + jæn → jæn, zɯ + kɔ → zɯɔ, and pəŋ + lɛ → pɛɛ*

UR	j ə ʔ j æ n	z ɯ k ɔ	p ə ŋ l ɛ
		^	^
	X X X X X X	X X X X X X	X X X X X X



For the word /jəʔ.jæn/ (‘same’), it starts with six slots, each slot taken by each segment. When contraction happens, the six slots reduce to three. Edge-In specifies that [j] takes the first X-slot, while [n] takes the third X-slot. Then the Vowel Selection²³ rule determines that [æ] takes

²³ I will discuss Vowel Selection in more detail in the next subsection.

the middle slot, which yields [jæn]. Now that all three slots are taken, the Glide Formation rule and the Vowel Spreading rule are blocked.

For the word /ʒu.kɔ/ ('Rugao'), although this disyllabic word also takes six slots, it only has four segments. The vowels [u] and [ɔ] both take two time slots. When contraction happens, the leftmost segment [ʒ] gets the first X-slot. The rightmost segment is a vowel [ɔ], which gets the third X-slot. As the vowel is already associated to a slot, the Vowel Selection process is blocked. Instead, the next rule, Glide Formation applies and takes the high vowel [u] and forms a corresponding glide [w], filling the middle slot. The application of Glide Formation then blocks the Vowel Spreading and finally yields [ʒwɔ].

For the word /pəŋ.lɛ/ ('originally'), as the second syllable is an open syllable, the vowel takes both the second and third slots. When contraction happens, Edge-In starts with the two edges, [p] and [ɛ], which take the first and third slot respectively. The vowel has already associated to a slot, blocking the Vowel Selection. Next, Glide Formation searches for any [+high] vocoid, but fails to find such candidate. At the last step, the vowel [ɛ] that has associated to the third slot *spread* to the second slot, associating to two slots. Finally, the output [pɛɛ] is derived.

10) Vowel Spreading

When there is no candidate for the middle slot, spread the vowel from the third slot.

With the observed Edge-In having been discussed, it is worthwhile to point out here that the Edge-In is not totally non-violable. It is particularly true when the disyllabic word is followed by another one or two syllables or the contractible disyllabic part is within a trisyllabic or quadrisyllabic word, but also applies for some disyllabic words in the stand-alone form. For example,

the contracted forms in (19)—(22) do not preserve the right edges, although all of them keep the left edge:

- (19) $\epsilon jo_{213} + \epsilon j\text{æ}\text{ʔ}_{33} \rightarrow \epsilon y\text{æ}_{213}$ ‘primary school’
 (20) $t\text{ɛ}h\text{j}\text{ɔ}\eta_{35} + \text{ɹ}u\text{u}_{55} \rightarrow t\text{ɛ}h\text{j}\text{ɔ}\text{ɹ}_{35}$ ‘as if’
 (21) $\text{ɕ}u\text{j}_{55} + p\text{ɔ}\text{ʔ}_{55}. t\text{ɛ}h\text{æ}\text{ʔ}_{35} \rightarrow \text{ɕ}w\text{ib}_{55}. t\text{ɛ}h\text{æ}\text{ʔ}_{35}$ ‘can’t fall asleep’
 (22) $f\text{ɔ}\text{ʔ}_{35} + v\text{ɔ}\text{ʔ}_{55}. \text{q}u\eta_{35} \rightarrow f\text{ɔ}v_{35}. \text{q}u\eta_{35}$ ‘servant’

As briefly discussed before, the contracted form can be influenced by many factors including not only the contraction rules, but also the phonotactics of the language. To allow the irregularity, some other contraction-related rules need to apply before the phonotactics interfere. Some phonotactic rules that are relevant for the presented data include:

11) **Identical syllable* (derived from the No-Identity constraint (R.-F. Chung, 1996))

The contracted syllable must not be identical to any one of the syllables in the precontraction form.

12) *-ʔ deletion*

Delete the word-final /ʔ/.

13) *-ɹ suffixation*

Suffix -ɹ/ for diminutive morpheme -/aɹ/.

For (19) $\epsilon jo + \epsilon j\text{æ}\text{ʔ} \rightarrow \epsilon y\text{æ}$, the Edge-In applies first, followed by Vowel Selection. For Glide Formation, although there is a glide candidate /j/, the glide in the contracted form is a rounded [ɥ]. This is because **Identical syllable*, which disallows the contracted syllable to be identical to $\epsilon j\text{æ}\text{ʔ}$ in the pre-contraction form. Instead, glide /j/ takes the rounding feature of the following /o/ and gets preserved as [ɥ]. Additionally, the glottal stop [ʔ] is lost in the contracted output due to the ʔ-deletion process that is common in casual speech Rugao. In other words, the

contraction process, starting with the Edge-In and ending with Glide Formation, is followed by another step of derivation. The post-contraction ? -deletion rule finally reshapes the contracted output to $\epsilon y\text{æ}$.

For (20) $t_{\epsilon h j \text{?} \eta} + \text{?} \text{?} \text{?}$, according to the Edge-In, the hypothetical contraction $t_{\epsilon h j \text{?} \text{?}}$ is phonotactically inferior to $t_{\epsilon h j \text{?} \text{?}}$ in two respects. First, $*j\text{?}$ is an illicit combination, which violates the constraint that disfavors the combination of the glide /j/ and high vowels (i.e. $*j[\text{high}]$). Second, the -r suffixation (i.e., the diminutive morpheme) is quite common in Rugao and can apply to almost all vowels. For example, /tsha₂₁/ + /a₁₂₁/ $\rightarrow$ [tsha₁₂₁] ('little car, cart'), /ey₂₁₃/ + /a₁₃₃/ $\rightarrow$ [ey₁₂₁₃] ('little Xu (last name)'), /vɛn₂₁/ + /a₁₂₁/ $\rightarrow$ [vɛ₁₂₁] ('little turn'). The syllable $t_{\epsilon h j \text{?} \text{?}}$ can be derived from /tɕh_jɔ₂₁₃/ + /a₁₃₃/ $\rightarrow$ [tɕh_jɔ₁₂₁₃] ('little Qiao (name)'), and thus is a completely natural syllable. The Edge-In obeying option $t_{\epsilon h j \text{?} \text{?}}$ is not the best output, thus losing to the phonotactically well-formed $t_{\epsilon h j \text{?} \text{?}}$.

For (21), the case of $\text{?} \text{?} \text{?} + p\text{?} \text{?}$. $t_{\epsilon h \text{æ} \text{?}} \rightarrow \text{?} \text{?} \text{?} \text{?}$. $t_{\epsilon h \text{æ} \text{?} 24}$ may be more complicated. Although only the first two syllables are contractible, the second syllable, $p\text{?} \text{?}$, which means “not”, is in a morphologically weak position. A similar word $\text{?} \text{?} \text{?}$. $t_{\epsilon h \text{æ} \text{?}}$ ('can fall asleep'), if contracted, also produces a similar output that violates the Edge-In: $\text{?} \text{?} \text{?}$. $t_{\epsilon h \text{æ} \text{?}} \rightarrow \text{?} \text{?} \text{?} \text{?}$. The contracted output in both cases, $\text{?} \text{?} \text{?}$ and $\text{?} \text{?} \text{?}$, are illicit syllables with $*wi$ and non-glottal stops in the coda.

Next, I will show how phonotactics and the rules of contraction work together in shaping the final contraction form with a derivation of (19) $\text{?} \text{?} \text{?} + \text{?} \text{?} \text{æ} \text{?} \rightarrow \text{?} \text{?} \text{æ}$.

²⁴ Note that “+” is only used for syllable boundaries for the contractible part, while “.” is used for the non-contractible part.

14) Derivations for $\epsilon jo + \epsilon j \alpha ? \rightarrow \epsilon \eta \alpha$

UR $\epsilon j \ o \ \epsilon j \ \alpha \ ?$

$\vee \ \wedge \ \vee \ | \ |$

X X X X X X

Contraction $\epsilon j \ o \ \epsilon j \ \alpha \ ?$

X X X

Edge-In $\epsilon j \ o \ \epsilon j \ \alpha \ ?$

$\swarrow \quad \searrow$
X X X

Vowel selection $\epsilon j \ o \ \epsilon j \ \alpha \ ?$

$\swarrow \quad \searrow \quad \searrow$
X X X

***Identical Syllable**

Glide formation $\epsilon j \ o \ \epsilon j \ \alpha \ ?$

$\swarrow \swarrow \swarrow \quad \searrow \searrow$
X X X

-? deletion $\epsilon j \ o \ \epsilon j \ \alpha \ ?$

$\swarrow \swarrow \swarrow \quad \searrow$

Surface $\epsilon \ \eta \ \alpha$

The violation of Edge-In is found in many other Chinese dialects. The frequency and freedom of such violation may depend on the language. In the following Tianjin and Taiwan Mandarin examples, the rightmost edge is not preserved in the contracted output either. In the

two Tianjin examples, $zən + min$ is contracted as $zəm$ instead of $zən$ for two possible reasons: first, it is likely that $zən + min$ is contracted as $zən$ first, but is realized as $zəm$ because of coarticulation, as the bilabial stop /b/ follows immediately. The other possible explanation is the No-Identity constraint (R.-F. Chung, 1996; Hsu, 2003) that bans the contracted syllable that is identical to one of the syllables in the precontraction form. Similarly, the other words are not cases of simple truncation of the final segments either. The case of $thai + phiŋ \rightarrow thaim$ can be the result of a post-contraction process on the nasal that receives the [bilabial] feature from the preceding consonant [p_h].

(23) $zən + min. bi \rightarrow zəm. bi$ ‘Renminbi’ Tianjin, from Wee (2014)

(24) $thai + phiŋ. tɕie \rightarrow thaim. tɕie$ ‘Peace Street’ Tianjin, from Wee (2014)

(25) $in + wei. wo \rightarrow iu. wo$ ‘because I’ Taiwan Mandarin, from Tseng (2006)

(26) $wo + mən \rightarrow wom$ ‘we’ Taiwan Mandarin, from Tseng (2006)

In summary, the slightly modified Edge-In Association can successfully account for the pattern for the outmost segments in Rugao syllable contraction. The corpus data of Rugao syllable contraction has shown categorical left edge preserving, keeping the first segment of the contractible disyllabic part in the contracted syllable. Although right edge preserving can sometimes be overridden by other rules related to syllable contraction and/or the phonotactics of the language, the right edge segment is still most likely kept. One important difference of the Edge-In that is proposed in this dissertation and the previous models (R.-F. Chung, 1996, 1997; Hsu, 2003) is that I do not specify the location of the Nucleus in the XXX template for the contracted syllables, although the lexical syllables do place the Nucleus in the middle. The freedom of the location of the Nucleus in the contracted syllables makes it easier to account for the patterns in the great variations in the syllable structures of the contracted syllable.

2.3 Vowel selection

2.3.1 Two approaches

From the previous subsection, I have discussed how *Edge-In Association* selects the leftmost and rightmost segments for the contracted output. As shown in the derivations above, for contractible words that end with a vowel, the vowel can directly get the right edge for being the edge element; it does not need or allow any other process to determine the vowel in the contraction. For many others, however, the Edge-In itself can set only the left and right edges but is not sufficient to predict the vowel realization in the surface output. In these cases, the Vowel Selection process plays a crucial role in determining the final output. Below are examples of this type:

- (27) $\text{pjəŋ}_{21} + \text{ɛjæn}_{21} \rightarrow \text{pjæn}_{213}$ ‘fridge’
(28) $\text{ɛjæn}_{21} + \text{ɛjən}_{55} \rightarrow \text{ɛjæn}_{21}$ ‘believe’
(29) $\text{fɛn}_{213} + \text{tsən}_{55} \rightarrow \text{fɛn}_{215}$ ‘anyway’
(30) $\text{tʃa}_{21} + \text{xow}_{21} \rightarrow \text{tʃaw}_{21}$ ‘car accident’
(31) $\text{fɛn}_{213} + \text{aI}_{33} \rightarrow \text{fɛI}_{213}$ ‘on the contrary’

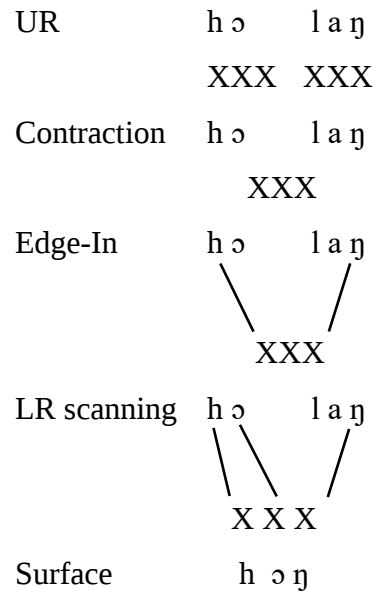
The contraction of both words in (27) and (28) show clear Edge-In and preserves the initial and final consonants. To satisfy the requirements for the biggest syllable possible, one glide /j/ is also kept in both words so that the contracted syllable has the biggest possible CGVC structure. However, at this point, no constraint or rule has specifically answered the question of which of the two vowels should survive given only one slot in the contracted syllable. Why is [æ] the survivor in both (27) and (28)? This subsection is mainly devoted to investigating the vowel selection process in Rugao, more specifically, how the vowel nucleus is determined when the left and right edges are both fixated by consonantal segments. As this dissertation adopts the

X-slot position (instead of strict CV or mora) for the time tier, I will introduce the two X-slot based models for the vowel selection process, discuss the advantages and shortcomings of both models, and then fit the models to the Rugao contraction data. Along this discussion, I will present my analysis for Rugao syllable contraction and vowel nucleus selection in more detail.

In cases with clear consonantal edges on both sides, after the Edge-In process, the two vowels in the pre-contraction disyllabic contraction form are both candidates for the Nucleus position in the contracted form. The two vowels are thus in a competition for survival in the limited slot. With regards to which vowel is more likely to be preserved in the contracted output, Chung (1996) proposed the LR-scanning model for the analysis of Taiwanese Southern Min syllable contraction. This model is based on the relative linear position of the segments in the pre-contraction disyllabic form. The segment that is pronounced earlier is on the Left and the other segment is on the Right. With a basic XXX syllable structure with the Nucleus right in the middle, this linear order basically means that the candidates for the contracted Nucleus are a) the Nucleus of the first syllable, on the left and b) the Nucleus of the second syllable, on the right. LR scanning suggests that when two vowels compete for a single Nucleus position in the contracted syllable, in a left to right order, scan [+syllabic] segments from the precontraction form and associate this segment to the Nucleus, i.e., the second X-slot among the three slots. If there is no such candidate, then scan [-consonantal] segments and associate to the middle X-slot. Below in 15) is a complete derivation of an example where the vowel on the left side of the pre-contraction word gets selected. For the word *hɔ + laŋ*, when contraction happens, the time slots are reduced, leaving an XXX template for the contracted syllable. The consonantal edges of the disyllabic word, [h] and [ŋ], are first selected as leftmost and rightmost edges respectively according to Edge-In Association. LR-scanning comes in where [ɔ] and [a] compete for the

single X-slot in the middle; [ɔ] is chosen over [a] in the surface form, as [ɔ] is on the left side of the pre-contraction word. The surface form is thus [hɔŋ].

15) LR-scanning analysis for [hɔ + laŋ] → [hɔŋ] (Chung 1996)



This linear order based LR-scanning model successfully predicts the nucleus selection in the Southern Min example above as well as in many other cases of contraction from Taiwanese Southern Min, such as in (32)—(34). In addition, LR-scanning allows for Glide Formation, which means the high vowels or mid vowels can become glides and skip the scanning, creating the GV type of Nucleus.

(32) tsit + tsun → tsin ‘this moment’

(33) li + tsap → liap ‘twenty’

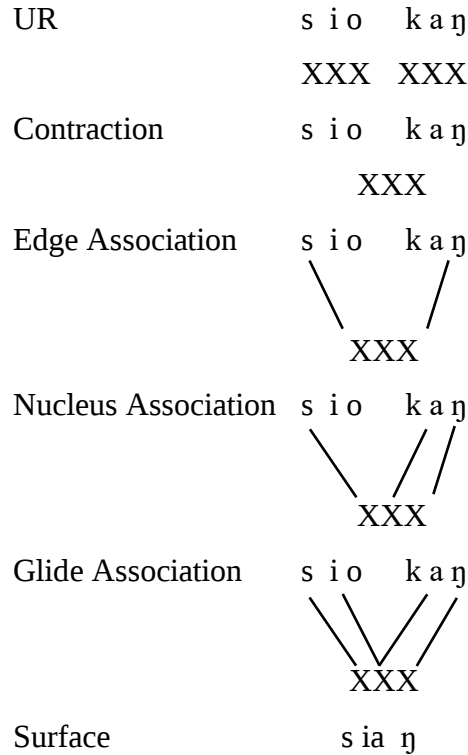
(34) tsa + hŋ → tsan ‘yesterday’

LR-scanning essentially predicts that the vowel on the *left* is more likely to get selected for the Nucleus given the pre-set edges and other segments are accounted for. In Hsu’s (2003) list of Taiwanese Southern Min words, although LR scanning is compatible for a good proportion of data (about 61%), it faces critical challenges from this language. For example, in

ho + guan → *huan* ('by us'), LR-scanning would have predicted [hon], as the leftmost [+vocalic] segment is [o]. Meanwhile, other researchers, such as Lin (1995) also expressed concerns with the LR-scanning. For example, *si + tsun* → *sun* ('time') and *tsit + tsun* → *tsin* ('this moment') have the same vowel candidates in the same left-right positions but contrastive outputs, with one that has the left vowel [i] while the other has the right vowel [u]. Unless one of the pair of words is contracted in a special way that is influenced by non-phonological factors, such vowel selection is impossible with either LR scanning or any other linear-based model.

To solve the problems that arise from the LR-scanning model, which includes the Edge-In Association and the X-slot, different researchers offer different alternatives. Hsu (2003) modified the basic *Edge-In Association* and proposed the sonority-based model. The major argument of the sonority-based analysis is that the vowel of *higher sonority* is preserved in the contracted surface. Meanwhile, a universal sonority hierarchy of $a > e > o > i > u$ (Kiparsky, 1979; Zec, 1995) is adopted and argued for based on contraction data in Taiwanese Southern Min. Taking the contraction of *sio + kan* → *sian* for an example, the more sonorant vowel [a] is chosen as the Nucleus in the surface. Glide Association follows immediately after the Nucleus Association, which links any possible glide to the same middle X-slot. A detailed derivation is shown below in 16).

16) Sonority-based analysis for [sio + kaŋ] → [siaŋ] (Taiwanese Southern Min, adopted from Hsu (2003))



A manual check suggests that the sonority-based analyses work for a greater proportion of data (87%) in the word list (Hsu, 2003). However, the sonority-based analysis also has its own limits. Crucially, the data in 15) shows a crucial counter-example to the sonority-based analysis for selecting the less sonorant vowel [ɔ] over supposedly more sonorant [a], although 16) is a counter-example to the LR-scanning analysis for choosing the vowel on the right, which can be better accounted for by sonority. Based on the findings for the Taiwanese Southern Min contraction data, I fed the two models to the Rugao contraction data in the corpus and found that about 55% of the data are compatible with the LR-scanning, which means that a good proportion of the vowel selections actually select the vowel on the right, but at least the right-vowel choices seem to be slightly less likely. The sonority-based model accounts for more Rugao data (≈70%).

The remaining 30% that do not fit the sonority model include some that fit the LR-scanning model, and some others that involve vowel coalescence, tone restriction, etc. I will leave these for the future discussions. The examples in (35)—(40) show that some cases can be accounted for using both approaches, but some others pose crucial challenges to any order-based analysis, whether LR-scanning or RL-scanning.

- | | | |
|------|---|-------------------|
| (35) | $\text{ɛjæn}_{21} + \text{ɛjən}_{55} \rightarrow \text{ɛjæn}_{215}$ | ‘believe’ |
| (36) | $\text{tsa}_{55} + \text{kow}_{55} \rightarrow \text{tsaw}_{55}$ | ‘this’ |
| (37) | $\text{fɛn}_{213} + \text{a}_{133} \rightarrow \text{fɛ}_{1215}$ | ‘on the contrary’ |
| (38) | $\text{si}_{55} + \text{xɛj}_{55} \rightarrow \text{sej}_{55}$ | ‘time’ |
| (39) | $\text{zɿn}_{35+} + \text{xɛj}_{21} \rightarrow \text{zɛj}_{351}$ | ‘then’ |
| (40) | $\text{ɛjən}_{55} + \text{jɔŋ}_{55} \rightarrow \text{ɛjɔŋ}_{55}$ | ‘credit’ |

For data points (35)—(37), the selected vowel in the contracted output is both on the left side and more sonorant than the other vowel, so both the LR-scanning analysis and the sonority-based analysis can account for them. However, data points (38)—(40) select vowels on the right side, which either requires an RL-scanning or some other explanation. However, even if RL-scanning can be employed, it would predict that any language is free in choosing LR or RL scanning for different words within the same language and this is undesirable. The sonority account can solve this problem, assuming a simple universal sonority hierarchy based on vowel height and centrality: $\text{æ/a} > \text{e/ɛ}, \text{o/ɔ} > \text{i}, \text{u/u} > \text{ə}, \text{ɨ}$ (Kiparsky, 1979; Parker, 2011; Zec, 1995). For example, high vowels are generally lower in sonority than mid vowels. This explains why [ɨ] and [i] both lose to [e] in (38) and (39). The mid-central vowels are among the least sonorant vowels, so the more sonorant [ɔ] is selected. Such sonority hierarchy can also explain the vowel selection in (35) and (36), as the low vowels [æ/a] are generally ranked high in the sonority scale, they are

the most preferred candidates for the vowel selection process. Below in 17), I show two derivations, one for *tsa + kow* → *tsaw* ('this'), which selects [a] over [o], and one for *ɛjən + jən* → *ɛjən* ('credit'), which preserves [ɔ] and loses [ə]. The case of *tsa + kow* → *tsaw* is a simple derivation within the current frame. With the Edge-In, an intermediate [tsXw] is formed. Missing a vowel nucleus, the two vowels [a] and [o] are in a competition for this X-slot. The more sonorant [a] wins and gets selected, resulting in [tsaw]. For *ɛjən + jən* → *ɛjən*, similar Edge-In and Vowel Selection are applied, deriving [ɛXŋ] and then [ɛɔŋ]. Because of the presence of the glide, it requires Glide Formation, which searches for [+high] vocoids and associates it to the first X-slot. Note that I treat glides as sharing the first X-slot with the initial consonant in this dissertation, which is different from other analyses that associate glides to the middle X-slot.

17) Derivation for *tsa + kow* → *tsaw* and *ɛjən + jən* → *ɛjən*

UR	ts a k o w	ɛj ə n j ɔ ŋ
	XXX XXX	XXX XXX
Contraction	ts a k o w	ɛj ə n j ɔ ŋ
	XXX	XXX
Edge-In	ts a k o w	ɛj ə n j ɔ ŋ
	$\begin{array}{c} \diagdown \quad \diagup \\ \text{XXX} \end{array}$	$\begin{array}{c} \diagdown \quad \diagup \\ \text{XXX} \end{array}$
Vowel Selection	ts a k o w	ɛj ə n j ɔ ŋ
	$\begin{array}{c} \diagdown \quad \diagup \\ \text{XXX} \end{array}$	$\begin{array}{c} \diagdown \quad \diagup \quad \diagup \\ \text{X X X} \end{array}$
Glide Formation		$\begin{array}{c} \diagdown \quad \diagup \quad \diagup \\ \text{X X X} \end{array}$
Surface	tsaw	ɛjɔŋ

Note that di-syllabic words that end with a vowel (i.e., when the second syllable is an open syllable) are exempt from this additional process for vowel selection, as briefly mentioned earlier. Instead of a consonantal edge, the vowel on the right edge gets the last nucleus slot. Since this vowel already serves as the Nucleus for the contracted syllable, the selecting of another vowel is blocked while the lengthening of the last vowel or glide formation occurs instead, as seen in the data below. In (41) and (42), the vowels are lengthened. In (43), with the presence of a glide [j], the output can be either [nɛɛ] with the lengthened vowel, or [njɛ] with a glide. (44) shows clear glide formation. This variation in the glide preservation suggests the Vowel Spreading and Glide Formation processes (see previous section) can be exclusive and optional at least in some cases. Such cases provide more details into how the Edge-In association process works but are not directly related to the Vowel Selection. Readers are referred to 9) for detailed derivations for this type of contraction.

- | | | |
|------|--|----------------|
| (41) | tsəʔ ₅ + xa ₅₅ → tsaa ₅₅ | ‘these’ |
| (42) | si ₅₅ + tɕi ₅₅ → ši ₅₅ | ‘century’ |
| (43) | nej ₂₁₃ + lɛ ₅₅ → nɛɛ ₂₁₅ /njɛ ₂₁₅ | ‘you (plural)’ |
| (44) | nej ₂₁₃ + ka ₅₅ → nja ₂₁₅ | ‘your’ |

2.3.2 Sonority and vowel selection

From the analysis presented above, it is clear that the sonority-based analysis is a good tool for accounting for the patterns in Rugao syllable contraction especially after modifications are made to the details of the Edge-In. The vowel nucleus selection process and the sonority hierarchy are closely related in that vowels of higher sonority are more likely to be selected in the contracted output. The next questions are, first, how exactly does sonority affect the vowel

selection assuming the role of vowel sonority? And second, what is the vowel hierarchy of Rugao?

With regards to vowel sonority, there is a well-accepted universal sonority hierarchy based on vowel *height* and *centrality*, shown in Figure 4. This hierarchy basically ranks the lower vowels on the higher end of the sonority scale, while the higher vowels on the lower end of it, rendering a basic order of low > mid > high. There is no commonly accepted ranking of vowels of the same height, so those with the same height are not ranked. Central vowels such as /ə/ and /ɪ/ rank even lower than the non-central high vowels on this and many other vowel sonority scales, and non-low central vowels are generally regarded as less sonorant than the peripheral vowel qualities (Gordon et al., 2012; Parker, 2008, 2011, 2012). In particular, compared to the commonly accepted correlation of sonority and vowel height, centrality is a more consistent predictor of sonority across the languages being studied, based on many acoustic measurements and perceptual experiments (Gordon et al., 2012).

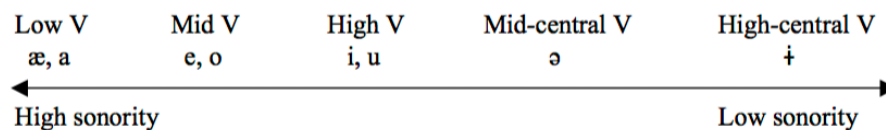


Figure 4. Universal vowel sonority hierarchy based on height and centrality (adopted from Gordon et al 2012)

Hsu (2003) adopted the universal sonority hierarchy, but expanded it and provided a more detailed ranking of $a > e > o > i > u$, for the simple five vowel inventory of Taiwanese Southern Min. Such hierarchy basically follows low > mid > high, but further ranks the mid vowels and high vowel as front > back, or unrounded > rounded. Although the general ranking for different languages may be similar, the detailed ranking could be slightly different in different languages. One can do a similar vowel preference ranking and then infer the vowel

sonority hierarchy for the case of Rugao. Although due to the small size of the corpus, it is impossible to find data for comparing each pair of vowels, the examples in (45)—(59) present some crucial data for vowel preference in Rugao contracted syllables. For example, data points (45)—(47) show the low vowels are more preferable than mid vowels. (48) and (49) are cases where mid vowels are chosen over high vowels. Based on the low > mid and mid > high ranking in (45)—(49), we can infer that low vowels may be more preferable than high vowels as well although direct evidence is lacking here. There are abundant data points that support peripheral > central. As shown in (50)—(59), it seems that central vowels ([ə] or [ɪ]) never survive when competing with any peripheral vowel. Within the central vowels, (58) and (59) offer evidence for a ranking of ə > ɪ.

Low > Mid

- (45) æ > o tow₂₁ + tshæn₃₅ → twæn₂₁₅ ‘how long’
 (46) æ > e phøj₃₅ + jæn₂₁₃ → phjæn₅₁₃ ‘foster’
 (47) a > o tsha₂₁ + xow₂₁ → tshaw₂₁ ‘car accident’

Mid > High

- (48) o > u nuŋ₂₁₃ + xow₅₅ → now₂₁₅ ‘warm’
 (49) e > i zɪn₃₅ + xej₂₁ → zøj₃₅₁ ‘then’

Peripheral > Mid central

- (50) æ > ə pjən₂₁ + ejæn₂₁ → pjæn₂₁ ‘fridge’
 (51) ɔ > ə ejən₅₅ + jɔŋ₅₅ → ejɔŋ₅₅ ‘credit’
 (52) ɛ > ə fən₂₁₃ + tsən₅₅ → fən₂₁₅ ‘anyway’
 (53) u > ə pəʔ₅₅ + huj₂₁ → puj₅₂₁ ‘cannot’
 (54) i > ə tən₂₁ + in₂₁ → tein₂₁ ‘experience’

Peripheral > High central

- (55) $e > i$ $si_{35} + xej_{55} \rightarrow sej_{35}$ ‘time/timing’
(56) $\varepsilon > i$ $si_{35} + t\epsilon j\epsilon n_{21} \rightarrow sj\epsilon n_{351}$ ‘time’
(57) $u > i$ $\epsilon i_{213} + xu\eta_{55} \rightarrow \epsilon u\eta_{215}$ ‘like’

Mid central > High central

- (58) $\partial > i$ $t\epsilon h i_{55} + s\partial \eta_5 \rightarrow t\epsilon h \partial \eta_{55}$ ‘actually’
(59) $\partial > i$ $si_{213} + \eta \partial n_{35} \rightarrow s\partial n_{213}$ ‘damn’

Summarizing from the data above, the vowel preference in Rugao syllable contraction has a basic ranking as shown below in 18). I leave vowels of the same height unranked for now due to the lack of supporting data but remain open in this matter. This ranking can be revised if future studies find direct evidence that support detailed ranking of vowels of the same height.

18) Vowel preference in Rugao syllable contraction based on corpus data

$a/\text{æ} > o, \text{ɔ}, e, \varepsilon > i, u/\text{u} > \partial > i$

Assuming the correlation between the vowel survival and sonority, the sonorities of Rugao vowels should rank in the same or similar way as in 18), as shown in 19). Due to the lack of data that supports a ranking of [ɤ], I will assume according to the general trend that it is placed with other high vowels and not ranked for now within the high vowel group.

19) Vowel sonority hierarchy in Rugao

$a, \text{æ} > o, \text{ɔ}, e, \varepsilon > i, u, \text{u}, \gamma > \partial > i.$

Such sonority ranking above in 19) is essentially in line with the universal sonority hierarchy based on vowel height and centrality as seen in Figure 4, with the low vowels ranking highest, followed by mid vowels and then high vowels, and the central vowels. It seems that allophones of the same phoneme in general do not compete, assuming from the lack of data that

clearly supports a ranking for pairs such as [a] vs. [æ]. However, problematic data are found for the phoneme /i/. [i] and [ɪ] are considered belonging to the same phoneme /i/, but seem to be ranked, with [i] above [ə] and [ɪ] below [ə]. One possible explanation for this is that the high likelihood of [i] to form a glide makes it more likely to survive in contraction, while [ɪ] as a central vowel is much less likely to form a glide. This contrast makes [ɪ] more vulnerable than [i]. [ɪ] is even more prone to deletion than mid-central [ə], which can be accounted for given that mid > high is established for vowels of all backness values.

2.4 Further discussion

2.4.1 More on vowel competition

In the analysis presented above, I demonstrated how syllable contraction happens based on some data in the corpus. The left and right edges of the precontraction form, whether consonant or vowel, are preserved as much as possible, with only a few exceptions without the right edges. For those that have clear Edge-In, sonority plays a crucial role in shaping the surface form, with the more sonorant vowel more likely to survive in the vowel competition. Until this point, I only presented analysis for data that shows a clearly “winning” vowel survives i.e., cases where either one of the two vowel survives while the other one deletes. However, the real word data is more complicated.

First, although I only investigate full contraction in this dissertation, cases of partial contraction are commonly seen in the corpus such as (60)—(62). Do these represent a separate issue from fully contracted words, in which the contraction-specific rules or constraints simply do not work? Or does this mean that vowels can tie up in the competition? There needs to be ways to account for these data. Second, there are limited data that suggest vowel coalescence

instead of deletion, such as in (61), which is again not compatible with the sonority-based analysis. Particularly, vowels that undergo coalescence are not necessarily sonority ties, so this cannot be explained by unranked sonority. Vowel coalescence is especially common in cases of extreme contraction, which reduces three syllables into one. Third, the current vowel competition theories only deal with two vowels in competition but does not provide analysis for contraction of more than two vowels, such as in (61)—(63).

- (60) $f\epsilon n_{55} + ts\epsilon \eta_{55} \rightarrow f\epsilon \epsilon \eta_{215}$ ‘anyway’
 (61) $ts\epsilon \eta_{55} + j\epsilon \eta_{55} + ts i_{21} \rightarrow ts\epsilon ndz_{521}$ ‘this way’
 (62) $p\epsilon \eta_{55} + j\epsilon \eta_{55} + j\epsilon n_{55} \rightarrow p\epsilon \epsilon n_{55}/pj\epsilon n_{55}$ ‘not the same’
 (63) $kow_{213} + si_{21} + a_{21} \rightarrow ka_{21}$ ‘is it?’

As the partial contraction and full contraction as well as full deletion and vowel coalescence may co-exist for the same words, there should be a way to allow for this coexistence. I propose that the different types of contraction should be treated differently by ordering the rules. For example, violating the rule *VV in particular is crucial in allowing two vowels in the same contracted syllable, but *VV has to apply for the full contraction in order to render the right output. For the cases with vowel coalescence and extreme trisyllabic contraction, the sonority-based Vowel Selection is ranked lower to allow for another option for the vowel competition.

2.4.2 Violable phonotactics

Comparing the contracted syllables and the lexical syllable, it seems that the contracted syllables have much more freedom in violating the phonotactics of the language (Hsiao, 1995; Hsu, 2003). It is a similar case for Rugao. First, according to the three X-slot analysis, the

contracted syllable may place the Nucleus in the final or medial X-slot depending on the form.

The partially contracted syllable may even have two vowels in the Nucleus.

The rules that directly affect the process of contraction are usually obeyed, although there is an order of application. To facilitate contraction, Maximal Segment has to be violated to lose segments in the process of contraction. Maximal Syllable is violated less frequently, meaning that although the syllables are maintained as complex as possible that is allowed by the language, speakers have the freedom to use a less perfect syllable. This also means that rules such as Glide Formation can be optional, which is supported by the data.

The surface contracted syllables are often phonotactically ill-formed syllables in the language, which means that many of the phonotactics of the language can be violated. In fact, the data seems to suggest that it is likely that most of the phonotactics regarding the syllable structure and segment sequence can be violated. Below in (64)–(72) are some examples that show such violation. First, Rugao in general does not permit complex Onset unless a glide appears as the second segment, for example, *ɛjæŋ* vs. **ɛɹæŋ*, *kɛ*, *kwe* vs. **knɛ*. Note again that the asterisk *** in this dissertation means phonotactically illicit. However, complex Onsets are quite common for contracted syllables. For example, the contracted syllables in (64) and (65) both have complex Onsets that do not contain glides **pɹ*, **pŋ*. Second, the Codas of contracted syllables are also much freer in allowing segments that are not found in lexical syllables, such as the *-v* and *-m* codas in (66) and (67).

(64) $pəʔ_5 + zɿn_{55} \rightarrow *pɿn_{55}$ ‘otherwise’

(65) $pəʔ_5 + ŋɛ_{21} \rightarrow *pŋɛ_{521}$ ‘alright’

(66) $fəʔ_{35} + vu_{55} \rightarrow *fəv_{35}$ ‘serve’

(67) $tɕhɯŋ_{35} + phu_{21} \rightarrow *tɕhum_{351}$ ‘all’

In addition, the restrictions on the GV or VG sequences are also easily breakable. As discussed previously, there are strict rules for the glides with regards to what vowels or consonants each of them can follow or precede. But these rules are readily breakable in a contracted syllable to accommodate for Glide Formation and Maximal Syllable. This type of syllable is commonly seen in the corpus. Below in (68)—(72) are a few examples.

- (68) ka₂₁ + tɕhY₂₁ → *kaɥ₂₁ ‘go home’
- (69) ŋow₂₁₃ + ka₅₅ → *ŋwa₂₁₅ ‘my’
- (70) phi₅₅ + kow₂₁₃ → *phiw_{55/513} ‘bottom’
- (71) vej₂₁ + səŋ₂₁₃. mɐI₃₅ → *vem₂₁. mər₃₅ ‘why’
- (72) tsha₂₁ + xow₂₁ → *tshaw₂₁ ‘car accident’

2.4.3 Problems with using corpus data

There are some other problems with the analysis above that stem from the nature of corpus data itself. Notably from the data presented above, both the vowel on the left and the right can survive in the contraction, but the likelihood is not half/half. It seems that the vowels on the *right* are actually more likely to survive due to the right-edge preservation. Does this mean that Rugao may have a low-ranking right vowel preference instead of the left? Or alternatively, are words with the more sonorant vowel on the right more likely to contract? The seemingly right vowel preference is observed in a subset of the data, but the evidence is far from enough for one to confirm it. As the speakers for the corpus use only a limited number of words in the limited timeframe and provided topics, the corpus lacks the variety of words necessary to make a generalization. It is still possible if both left and right vowels are equally likely to survive phonologically, the right vowel is chosen for other reasons. In addition, there is hardly any

potential for meaningful analysis on some other aspects of the data, such as all possible combinations of vowels in competition, any statistics of what kind of vowel combinations most easily trigger contraction, and so on. This is also due to the limitations of the corpus. However, an experimental setting will make it easier to test on matters that cannot be done with corpus data, offering all different combinations of vowels in well-controlled phonetic environments. Phonologically well-formed words in particular can help eliminate the other non-phonological influences.

Another issue of using only the corpus data is that all the data transcription is done impressionistically by the author using standard IPA. From such transcription, syllable contraction may seem as if a vowel is completely deleted if it loses in the vowel competition on the one side. On the other side, the surviving segments also look like they are identical to their lexical counterparts, although the syllable itself may not be a lexically well-formed one. The categoricity of using impressionistic IPA transcriptions has the undesirable consequence of potential loss of phonetic details, which at this stage seems missing from both the data and the analyses. However, such details are important for us to fully understand the process of syllable contraction, as they offer deeper insights to processes such as vowel deletion/surviving and ultimately vowel competition. Clearly from the analysis, the contracted syllables are not the “regular” syllables. They are distinct from lexical syllables in that most of the phonotactics can be violated. Such freedom of a contracted syllable sparks many questions and suspicions about what is really transcribed by the human linguist. Does a losing vowel have to be completely *removed* from the surface output? Or does it actually *exist* in some other way? The irregular nature of the contracted syllables makes it more difficult for the transcriptions to faithfully

present the actual representation. These questions cannot be answered with the current analyses and call for investigations using other methodologies.

2.5 Conclusion

This chapter started with an overview of the language of Rugao and an introduction of its phonological system. As the most recent work on Rugao, I recaptured and updated consonant, vowel, and tone systems in the former work, mostly Wu (2006) and Huang (2011). On the basis of the description of the phonotactics of Rugao, the major observable patterns in syllable contraction data were investigated within the framework of templatic analysis. Rugao syllable contraction exhibits similar patterns to many other Chinese languages, including the Edge-In and sonority-based Vowel Selection. They work with the phonotactics of the language and shape the surface output of the contracted syllable.

The phonological analyses lay the basic foundation for the investigation of Rugao syllable contraction using corpus data. Due to the limitations of the corpus and real word data in general, some issues have to be solved using experimental techniques. In the next two chapters, I will present two experiments that are designed to explore the two major questions that are discussed above. The first experiment is a nonce word contracted task targeting the vowel selection process with regards to the sonority bias as well as the linear position bias using different combinations of vowels. The second experiment is a production experiment that again utilizes different combinations of vowels to elicit disyllabic word contraction in a better controlled setting, trying to look deeper into the vowels in the contracted syllables. Both experiments will provide insights for further understanding the surviving vowel and the deleting vowel in the process of syllable contraction.

3 VOWEL COMPETITION IN SYLLABLE CONTRACTION

3.1 Introduction

3.1.1 Motivation for this study

As discussed in the previous chapters, the mechanisms behind the selection of vowel in the contracted form is essentially one of vowel competition in a broader view. Such mechanism is not only key to understanding the process of syllable contraction, but also generally how the winner of the vowel competition is determined. Two opposing types of analyses are under discussion here with regards to the vowel selection process in Chinese languages, the linearity-based analysis (Chung, 1996, 1997; Sun, 2014) and the sonority-based analysis (Hsu, 2003), both of which have support from real word contraction data from Taiwanese Southern Min and other Chinese languages such as Taiwan Mandarin (Cheng & Xu, 2009; Kuo, 2010). From the Rugao data presented in the phonological analysis, Rugao seems to support the sonority-based model better than the linearity-based model. It is possible that languages may vary in their choices of the vowel nucleus, but the fact that many languages provide contradicting data for both analyses might have been an artifact of real word data because of its great complexity. There needs to be some way to testify that one of the analyses, or both analyses, are in fact employed consistently regardless of the factors that are independent of vowels.

Due to the complexity of syllable contraction data, I will utilize experimental techniques for such inquiry. As the first study of this kind using an experimental approach, I will start this investigation with disyllabic words that have clear consonantal edges and those that can provide absolutely contrastive environments for the comparison of the two opposing analyses. On the basis of phonological analysis (Cheng & Xu, 2009; Chung, 1996, 1997; Hsu, 2003; Kuo, 2010;

Xu, Lin, & Durvasula, 2018) and theories of sonority (de Lacy, 2010; Gordon et al. 2012; Parker, 2008, 2011), a forced-choice experiment was designed to probe into the vowel selection using disyllabic nonce word stimuli that were constructed based on the Rugao contraction pattern. The focus of the experiment will be on the positional preference and the sonority bias of two vowels in competition. As the sonority-effect is partially attested in the real word data, I will explore the effect of vowel sonority and examine the different dimensions and aspects of the sonority hierarchy in order to find whether each of them contributes to the vowel preference alone. In this way, the vowels and vowel sonority are under closer examination alongside the exploration of the role that sonority plays in the syllable contraction and vowel selection.

This chapter is dedicated to the investigation of the vowel selection process with a focus on the comparison of the two existing models. The effect of vowel sonority will also be explored in more depth provided that the relationship between sonority and vowel selection in syllable contraction is established. The layout of this chapter is the following: I will first briefly recapture these two analyses of the vowel selection, discuss the advantages and disadvantages using the Rugao contraction data, and then introduce the experimental methodology and the results. After that, I will discuss the findings, compare the two analyses in how well each fits the data, and conclude this chapter with a discussion of vowel sonority.

3.1.2 Linear order or sonority?

As discussed in the previous chapter, the syllable contraction process, at least disyllabic syllable contraction, is basically the deletion of some time slots of two syllables, causing the merging of the two syllables. Whenever possible, the *Edge-In Association* (Chung, 1996, 1997; Hsu, 2003) sets the leftmost and rightmost edges for the contraction, leaving a limited time slot

for the vowel nucleus. With regards to which vowel is more likely to be preserved in the contracted syllable, there are basically two types of analysis. I will collapse the LR-scanning model originally proposed in Chung (1996) complemented by Chung (1997) and the other possibility of languages favoring the Right vowel (RL-scanning). I will call this type of analysis the linearity-based analysis in general. At the same time, I will use the key point in the sonority-based model proposed by (Hsu, 2003): sonority shapes the vowel selection. As can be seen in the discussion in the previous chapter, the LR-scanning model and Sonority-based model differ in quite a few details, including how each treats Edge-In and the Nucleus. As I have already discussed the two models in the phonological analysis, I will only focus on the crucial differences here while omitting aspects that are not relevant for the experiment.

Despite the relatively minor differences in other aspects of these analyses, the crucial difference lies in how they treat the vowel nuclei, which I will call the vowel selection for vowels under competition. For di-syllabic words that have clear consonantal edges, there is no need for any arguments for the edges and the only focus is on the vowel selection. These words are thus the ideal data and environment for us to examine the vowel selection. Let us re-examine some data of Rugao here:

Word	LR-scanning Prediction	RL-scanning Prediction	Sonority-based Prediction	Attested
[tsha + xow]	[tshaw]	*[tshow]	[tshaw]	[tshaw]
[sən + teɪ]	*[səɪ]	[seɪ]	[seɪ]	[seɪ]
[ɕjæn + ɕjən]	[ɕjæn]	*[ɕjən]	[ɕjæn]	[ɕjæn]
[ma + sən]	*[man]	[mæn]	*[man] /*[maæn]	[mæn ₂₅]
[nun + xow]	[nuw]	[now]	[now]	[now]/[nuw]

²⁵ Note that [mæn] is attested in the Rucheng sub-dialect, while in the West sub-dialect, [ma + saŋ] → [maŋ] is attested.

For some data, especially words that have a vocalic right edge, e.g., [ɪəŋ.ka] (‘other’), the two models do not contrast, and both will predict the same output. But for the words above, the linearity-based model and the sonority-based model make distinct predictions. In particular, the linearity-based model specifies a positional preference for vowels in the pre-contraction word, either left or right. Even if a language has a left vowel preference or right vowel preference, it cannot choose both simultaneously. For the case of Rugao, the above data seems to suggest that both the left or the right vowels can be selected for the contracted syllable.

With sonority being generally better than the linear order, the vowel sonority also faces challenges from the variant contracted forms and vowels of the presumably same sonority. As there is a sonority hierarchy for a language, whether it is universal or language-specific, the sonority hierarchy predicts that the more sonorous vowel must be the winner. If the two vowels are not tied in the sonority ranking, there should be no variation, but this is not always the case in real word data. Take *nun* + *xow* as an example. Both *now*/*nuw* can be observed in natural speech contractions, but [o] and [u] are not tied in any sonority ranking; all sonority rankings that have been proposed rank high vowels below non-high vowels (Gordon et al., 2012; Kiparsky, 1979; Parker, 2011, 2012). Such variation seems to be phonologically unconditioned and cannot be explained by the linear order, either. Another problem is posed by words such as *ma* + *sæn*. Unless there is clear evidence that suggests that [a] and [æ], both low vowels, are ranked as [æ] > [a] in sonority, the sonority-based model would have to retain both vowels, or select the first vowel (Hsu, 2003). With *[man] unattested at least in the Rucheng dialect, such vowel selection is problematic for the existing sonority-based model as well. Given the fact that in the Rucheng sub-dialect and the West sub-dialect have exclusive contraction forms for this particular word, e.g., [ma + sæn] → [mæn] vs., [ma + saŋ] → [maŋ], it seems that the phonotactics of the

language plays a crucial role in the post-contraction phase. The final survival of vowels does not depend only on the linearity and/or sonority of the vowel, but also on the phonotactics of the language.

Real word data is always a mixture of many aspects of the phenomenon and the language, including the phonotactics of the language (Chung, 1996, 1997; Hsu, 2003; Tseng, 2005), the many factors that prohibit or promote syllable contraction (word/morpheme boundaries, word frequencies (Myers & Li, 2009), inner segments (Cheng & Xu, 2009) and even the variation in personal preference and inter-speaker variation (K. S. Chung, 2006; Kuo, 2010). The real word data is thus often hard to interpret, especially for a specifically focused question that requires the isolation of one factor from the many. It is uncertain whether vowel sonority indeed shapes the contracted output, or it is merely a pattern for specific words. On the one hand, data obtained under experimental settings also have problems because of the nature of experiments. On the other hand, experimental data may offer insights that natural speech data cannot, including the advantages of controlled stimuli and the use of non-sense words. The combination of well-controlled stimuli and nonce words would limit the influences from other aspects of the phenomenon that are not under study. If the role of sonority can be confirmed by generally all words (i.e., words that are phonologically well-formed but are free of the influences of syntactic category, semantic focus, etc.) then we can be more confident and validate the formal analysis. All these motivate the experimental study of the vowel selection in syllable contraction.

3.1.3 Vowel sonority

In order to fully discuss the sonority-based analysis, the vowel sonority needs to be explored in more depth here. Vowel sonority has long attracted the interests of researchers.

Sonority has been argued to be the most constant feature that predicts stress attraction by many studies²⁶ (de Lacy, 2010; Kenstowicz, 1994). For example, stress is attracted by high sonority vowel in languages such as Kobon and Takia, e.g., in /mo.u/ → [mó.u] ‘thus’, the more sonorant [o] is stressed, compared to the less sonorant [u] (Kenstowicz, 1994). Vowel sonority is also related to vowel reduction in languages like Bulgarian and Catalan (Crosswhite, 2000), as high sonority vowels, typically low and mid vowels, undergo qualitative reduction/neutralization to schwa or high vowels in unstressed position, while the low sonority vowels, typically high vowels, do not participate in such reduction. For example, /gradéts/ → [grədéts] (‘town’, Bulgarian) shows that the high sonority vowel /a/ gets reduced to a schwa, while low sonority vowels do not reduce, e.g., /əprimá/ → [əprimá] (‘to make thin’, Catalan). When analyzing height coalescence as a hiatus resolution, (Casali, 1998) proposed that PARSE(-high) is ranked above PARSE(+high) in height coalescence, showing a preference for more sonorous vowels in languages such as Yoruba, Bolia and Dangme, e.g. /a + i/ → [e]. Within the area of syllable contraction, previous studies have also provided evidence that sonority can shape the output of the contracted syllables and proposed possible sonority rankings of vowels in Chinese languages such as Southern Min.

There are various definitions of sonority, including phonology-based and phonetics-based approaches. The major difference being whether the sonority hierarchy is made based on the generalized phonological patterns or phonetic properties. Phonology-based definitions of sonority define it as the relative prominence of segments in phonological process such as stress

²⁶ It is important to acknowledge here that experimental evidence shows that at least for some vowels and some languages, i.e., [a] in Gujarati, the stress is not sonority-driven, but consistently falls on the penultimate position (Shih, 2016). However, due to very limited literature that supports this viewpoint, I still assume the relationship between stress and sonority.

attraction, vowel deletion, etc. (Casali, 1998; Crosswhite, 2000; de Lacy, 2010). The phonetic-based definition characterizes sonority as the acoustic intensity of sounds based on acoustic measurements such as intensity, durations, etc. (Gordon et al., 2012; Ladefoged & Maddieson, 1996; Parker, 2008). For example, Parker's (2008) study is based on measurements of intensity extremes, the peak intensity of vowels and the intensity nadir of consonants, which matches closely to the phonology-established sonority hierarchies. Gordon et al (2012) measured total duration, maximum intensity, first formant values, total acoustic intensity, and total perceptual energy and proposed a slightly different scale.

No matter how one approaches it, it is commonly accepted that vowel sonority adheres to a hierarchy that is generally predictable by height and centrality (de Lacy, 2010; Gordon et al., 2012; Kenstowicz, 1994): higher vowels are ranked lower in the sonority scale, as are central vowels. There is no commonly accepted sonority hierarchy for vowels of the same height. Whether there is only a [+high] and [-high] contrast or a detailed low > mid > high ranking still varies according to the type of data analyzed. The ranking of vowels of the same height is least agreed upon. They can simply be unranked, such as a > e, o > i, u (Parker, 2012; Selkirk, 1984). Some studies rank front > back, such as in a > e > o > i > u (Kiparsky, 1979), but there are also data that support a more complicated system, i.e., a > æ > ε > ɪ > u > i (Ladefoged & Maddieson, 1996). The syllable contraction study of Southern Min suggests a sonority hierarchy of a > ɔ > e > o > i > u (Hsu, 2003), and Mandarin syllable contraction data seems to suggest o > e (Tseng, 2005b).

Given the variation of vowel sonority scales and the massive variation of relevant data, the study of syllable contraction is another opportunity for the investigation of vowel sonority in general. First, although the sonority hierarchy of vowels is still in debate, it is usually common to

rank vowel sonority along the dimensions of height, which can also be attributed to formant information. The Rugao data (Xu et al., 2018) as well as Taiwanese Southern Min data (Chung, 1997; Hsu, 2003) both seem to suggest that height is a reliable attribute of vowel sonority. But there are also data that suggest that there is no general height ranking, but rather a high vowel vs. non-high vowel contrast. The study of syllable contraction may offer support to one of the two types of rankings with in-depth exploration of what type of vowel survives the competition, the generally lower vowel or the non-high vowel. Second, the Rugao data suggest that centrality may contribute to the sonority hierarchy. But these generalizations, based solely on syllable contraction data, need to be confirmed in some more general way due to the fact that the contraction form can be influenced by multiple factors. Third, the study of syllable contraction may also provide insights for the discussion of whether a detailed ranking can be made for vowels of the same height. Finally, although there is a universally agreed upon sonority hierarchy, data from different sources suggest that the sonority hierarchy can be language-specific, as reviewed in the previous paragraph. How much freedom there can be for language-specific variation and whether there is a baseline for such variation are questions that are still unclear. With the proposed experiment, I will try to explore these aspects of vowel sonority and add to the discussion of issues and questions with regards to this topic.

3.1.4 A preview of the experiment

As a summary, the analysis of Rugao syllable contraction data suggests that both the LR-scanning analysis and the sonority-based analysis each could explain a subset of the data, but both have exceptions. The corpus data suggests that the sonority-based explanation is better than the LR-scanning analysis in accounting for the observed patterns, but more evidence is required.

Because the real word contraction data receives influence from factors beyond phonology, it is hard to pinpoint to exactly how the vowels compete using the real words.

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:

- a. If speakers of Rugao are forced to judge the contracted form of nonce words, will they have a systematic preference of a particular type of vowel in the contracted form?
- b. If participants do have a systematic bias, is such bias well accounted for by the linear-order analysis or the sonority-based analysis?
- c. If the participants show a more consistent preference to the more sonorous vowels, what aspect of the sonority hierarchy specifically contributes to the preference?
- d. Even if the sonority hierarchy is a strong predictor of the vowel preference, does the linear order of the two vowels in the pre-contraction forms simply have zero effect on their choices?

In order to answer these questions, the stimuli need to be organized in a way that allows for the comparison of vowels for different aspects of sonority, i.e., height, centrality, and possible backness/roundedness. At the same time, the stimuli should also allow for examining the possible preference for linear order that is independent from the vowel qualities, i.e., whether there is a preference for the vowel on the left or the vowel on the right, no matter what the vowel is. In addition, all stimuli should consist of a pre-assumed syllable structure in order to facilitate the sole comparison of vowels. In the following section, I will elaborate on the methodologies of the stimuli construction, and then discuss the results in three blocks: the two models compared,

and the height and centrality aspects of vowel sonority. I will finally draw a conclusion on which model is more compatible with the experimental data and discuss the implications on vowel competition and vowel sonority.

3.2 A forced-choice contraction task

3.2.1 Stimuli

As seen in the previous sections, many other factors such as phonetics, morphonology and word frequency can affect the realization of contracted output in the contraction process. The nonce word task method was adopted in order to exclude other factors and solely test the phonological aspect of this process. All stimuli in this experiment are phonologically well-formed nonce words of Rugao. The nonce words have to be used for targeted testing, but they come with a sacrifice for the naturalness of the stimuli, which might increase the difficulty of the task. The increased difficulty might decrease the participants' willingness to contract, as contraction happens mostly in casual speech. Despite being nonsense, all the syllables are phonologically well-formed. This should increase the naturalness of the stimuli at least to some extent and in turn facilitate contraction.

As the experiment aims to test the vowel selection between competing vowels, the syllable structure of the stimuli needs to be kept as consistent as possible in order to control the phonotactics and eliminate the complications that stem from the syllable structure and other segments. The stimuli are thus constructed as if a di-syllabic word is contracted into a monosyllabic word, keeping the left and right consonantal edges of the pre-contraction disyllabic form while selecting one vowel from the two vowels available. C_1 , the onset of the first syllable, sets the left edge for the Edge-In process, while the coda of the second syllable, C_2 , sets the right

edge boundary. To imitate real word contraction, the inner consonants, i.e., the coda of the first syllable and the onset of the second syllable, are left optional depending on the vowel and syllable-structure restrictions. According to the Edge-In and Vowel Selection principles of real word contraction, all contracted syllables should have a $C_1V_?C_2$ syllable structure. Both consonants C_1 and C_2 are predictable, while the survival of vowel is unknown at the first step of contraction. The process of syllable contraction is shown below, with the syllable structure on top and the segments on the bottom:

$$\begin{array}{ccccc}
 C_1V_1(C_1) & + & (C_2)V_2C_2 & \rightarrow & C_1V_?C_2 \\
 | & & | & & | \\
 p_h V_1(n) & + & (x)V_2 n & \rightarrow & p_h V_?n
 \end{array}$$

The choice of segments also follows the principles of the phonotactics and the likelihood of contraction. The initial consonant C_1 of all the disyllabic words is kept $[p_h]$, as it is frequent in the language and can combine with most of the vowels. Only nasal and glottal stops are allowed in the coda in Rugao, but glottal stops have more limitations on the tone, as the vowels with a glottal coda must have a specific high check or low check tone. There is no such tonal restriction on the nasal coda, so $[n]$ is used for all the Codas whether C_1 or C_2 . The second syllable $(C_2)V_2C_2$ of all words except those with $[i]$ has a velar fricative $[x]$, as such words are likely to undergo contraction in the real word data. Zero onset was given for $[in]$ as $[*xin]$ is phonotactically illicit.

In such contraction, all contracted syllables should have a basic form of $[p_h V_?n]$, with the vowel being either V_1 or V_2 . With regard to the vowels, since this experiment focuses on contractions with clear consonantal edges, only vowels that can be followed by a coda consonant are used. While some vowels may appear both with and without codas, some vowels are more restricted. Table 5 summarizes the distribution of vowels regarding syllable structure:

	In open syllables	In closed syllables
Low	[a]	[æ]
Mid	[o] [e] [ɛ]	[ɔ] [ɛ] [ə] [e] [o] ²⁷
High	[i] [ɻ] [ʉ]	[i] [u]

Table 5. Distribution of Rugao vowels

Based on the distribution of vowels with regards to syllable structure, [i, u, ɛ, ɔ, ə, æ], vowels that can be followed by a nasal coda are used for stimuli construction, as seen organized in Table 6.

	Front	Central	Back
High	[i]		[u]
Mid	[ɛ]	[ə]	[ɔ]
Low	[æ]		

Table 6. Vowels used for stimuli construction

Among the selected vowels, there is a basic sonority ranking based on vowel height, assuming low > mid > high. The sonority of vowels of the same height varies crosslinguistically, but Rugao shows an inventory of mid vowels that have both featural contrasts: back vowels [o, ɔ, u] are all rounded vowels, while front vowels [e, ɛ, i] are unrounded. Despite the absence of relevant real word contraction data, such a vowel inventory introduces possible complications of the sonority ranking of vowels of the same height, comparing [ɛ] and [o], and [i] and [u], the rounded and unrounded vowels, or front and back vowels. With regards to the centrality of

²⁷ [e] only precedes [j]. [o] only precedes [w]. The other vowels may be followed by a nasal or glottal stop. All vowels can possibly be followed by /ɹ/.

vowel sonority, central vowels are generally ranked lower than peripheral vowels, so a mid-central vowel [ə] is used. According to the research questions laid out before, there were three blocks of stimuli, each targeting a specific vowel feature while controlling the others:

- Stimuli that compare vowels of different height, presuming the sonority hierarchy of low > mid > high. More specifically, [æ > ε > i] for front vowels, and [ɔ] > [u] for back vowels respectively.
- Stimuli that compare vowels of the same height, presuming the possible sonority hierarchy of [ɔ > ε] and [u > i].
- Stimuli that compare central and peripheral vowels, presuming the sonority hierarchy of [æ, ɔ, ε, i, u] > [ə].

To allow for a possible linear order comparison, each vowel is placed in both the first and the second syllable of the pre-contraction di-syllabic word. For example, [p_hε + xæn] and [p_hæɪn + xɛn] are both introduced, so that the linear-order effect can be examined whether [æ] is preferred over [ε] no matter what syllable it is in. The two possible contracted forms are: Option 1: [p_hV₁n], with V₁ in the nucleus, and Option 2: [p_hV₂n], with V₂ in the nucleus. The order of Option 1 and Option 2 is also flipped for each di-syllabic word, doubling the number of stimuli. In other words, there are two different triplets for each Word. For example, for the Word [p_hε + xæn], two types of triplet were included: [p_hε.xæn]-[p_hɛn]-[p_hæɪn] and [p_hε.xæn]-[p_hæɪn]-[p_hɛn]. This is to ensure that participants do not simply click on the first or the second option that they hear.

Table 7, Table 8 and Table 9 below are full lists of stimuli divided into three sub-groups according to the bulleted categorization above, each testing a specific local ordering of the

sonority hierarchy. Table 7 lists stimuli that compare low, mid and high vowels. Front vowels and back vowels are compared separately. Table 8 lists stimuli comparing the roundedness or front/back-ness using mid and high vowels. Table 9 lists stimuli that compare central vowels with other vowels using a mid-central vowel and peripheral high, mid and low vowels. In all these tables, the vowels are coded by height and front/backness. Unless specified as back, the vowels are front vowels. The only two rounded vowels used for stimuli construction are back vowels, so only backness is coded for them, and the roundedness can be inferred. Note that in the actual experiment, two types of stimuli, [p_hɛn] and [p_hɛ], are used for mid vowels in the first syllable because the syllable and tone combination [p_hɛn₂₁] has an extremely low frequency.²⁸ A preliminary check of the results suggests that participants show less contrast for comparison groups with the low frequency syllable [p_hɛn], although the basic conclusions still hold. For simplicity, groups with [p_hɛn] in the first syllable are excluded from the tables of stimuli and the analyses presented below. A slightly modified substitute CV syllable [p_hɛ] was used to prevent the low frequency issue. This affects groups for Mid + Low and Mid + High in Table 7, Mid + Mid-back in Table 8, and Mid + Central in Table 9.

All syllables here bear a low falling tone, for example, [xæn₂₁], which is the most frequent and most productive tone in Rugao. As the tone is both strictly controlled and irrelevant to the research questions of this project, the tone is not marked in the lists of stimuli that follow.

²⁸ There is only one word [p_hɛn₂₁] ‘攀’ (meaning: to climb), which is only pronounced this way in the *wendu* style, a way of formal reading that is heavily influenced by Mandarin. The commonly used informal pronunciation for this word is [p_hun₂₁] in Rugao.

Comparison group	Pre-contraction Word	Option 1	Option 2
Mid + Low	p _h ε + xæn	p _h εn	p _h æ̃n
Low + Mid	p _h æ̃n + xεn ²⁹	p _h æ̃n	p _h εn
High + Mid	p _h in + xεn	p _h in	p _h εn
Mid + High	p _h ε + in	p _h εn	p _h in
High + Low	p _h in + xæn	p _h in	p _h æ̃n
Low + High	p _h æ̃n + in	p _h æ̃n	p _h in
Mid-back + High-back	p _h ɔ̃n + xun	p _h ɔ̃n	p _h un
High-back + Mid-back	p _h un + xɔ̃n	p _h un	p _h ɔ̃n

Table 7. Stimuli for non-central vowels, compared by height

Comparison group	Pre-contraction Word	Option 1	Option 2
Mid-front + Mid-back	p _h ε + xɔ̃n	p _h εn	p _h ɔ̃n
Mid-back + Mid-front	p _h ɔ̃n + xεn	p _h ɔ̃n	p _h εn
High-front + High-back	p _h in + xun	p _h in	p _h un
High-back + High-front	p _h un + in	p _h un	p _h in

Table 8. Stimuli for vowels of same height, compared by roundedness or backness

Comparison group	Pre-contraction Word	Option 1	Option 2
Mid + Central	p _h ε + xæn	p _h εn	p _h ə̃n
Central + Mid	p _h ə̃n + xεn	p _h ə̃n	p _h εn
Mid-back + Central	p _h ɔ̃n + xæn	p _h ɔ̃n	p _h ə̃n
Central + Mid-back	p _h ə̃n + xɔ̃n	p _h ə̃n	p _h ɔ̃n
High-front + Central	p _h in + xæn	p _h in	p _h ə̃n
Central + High-front	p _h ə̃n + in	p _h ə̃n	p _h in
High-back + Central	p _h un + xæn	p _h un	p _h ə̃n
Central + High-back	p _h ə̃n + xun	p _h ə̃n	p _h un
Low + Central	p _h æ̃n + xæn	p _h æ̃n	p _h ə̃n
Central + Low	p _h ə̃n + xæn	p _h ə̃n	p _h æ̃n

Table 9. Stimuli for central vowel and peripheral vowels, compared by centrality

²⁹ Some speakers may produce a glide [j] before [ε].

The author (female, aged 28 at the time of recording), a native speaker of Rugao, produced all the stimuli in a quiet room using a high-quality microphone. All stimuli were produced in a carrier sentence: [ŋow su? ____ sɛn pʰin] ('I say ____ three times.'). To improve the naturalness of the reading, corresponding Chinese characters were used for the nonce words. Three repetitions for each stimulus were produced from which one was chosen. No manipulation was made to the stimuli in order to imitate the natural speech.

3.2.2 Participants and procedure

Participants for this experiment were recruited via an electronic flyer posted on the popular application WeChat. 42 native speakers of Rugao (11 males, 31 females, aged 18—33 at the time of the experiment) participated.³⁰ All participants were college students or had obtained a college degree at the time of the experiment. They were all born and raised in Rucheng Township, the downtown of Rugao, or were brought to Rucheng at a very early age. Other than the 3 or 4 years of college, the participants had been living in Rucheng and had not lived in other places for more than six months. All of them reported speaking Rugao exclusively at work and at home, or both Rugao and Standard Mandarin. Each participant received 30RMB (about \$5) for participation.

The experiments were conducted in a quiet room in downtown Rugao using a MacBook and a Bose QC5 noise-cancelling headset. PsychoPy (Peirce et al., 2019) was used for the

³⁰ Note that the experiment was conducted in two time periods, in August 2016 and May 2019. The first 21 participants were recruited in 2016 and Participants 22-42 were recruited in 2019. There is no observable difference between the two groups of participants with regards to age, gender distribution, level of education or language use.

presentation of the stimuli. Before the experiment, the researcher explained the concept of syllable contraction as “two syllables pronounced as one” and asked for their judgment on a short conversation of two natural sentences [nəj sən seɪ lə a? ɪəʊ mæn lə!] (‘When are you coming? I am coming right now!’), in which the underlined words [sən₃₁] [seɪ] and [mæn] are contractions of [sən.deɪ], [si.xeɪ] and [ma.sæn] respectively. The experiment proceeded only when the participant agreed that these sentences sound natural and understood that the three contractible words are examples of syllable contraction.



Figure 5. PsychoPy interface for training and test. [Note: The English translations were removed from the interface that the participants used.]

³¹ Note that this particular contraction does not follow the tested pattern. The complete deletion of the second syllable is likely due to the fact that the meaning of the word falls much more on the first syllable. However, the presence of this contraction should not decrease the naturalness of the sentence, as it is natural to have different patterns of contraction in the same sentence. For this reason, the author doesn't believe this particular word will cause any confusion for the participants.

The experiment consists of two phases: a *training* phase and a *test* phase. The two phases have the same interface and procedure demonstrated in Figure 5 (without the English translation). In the *training* phase, participants practiced with a short version of the test. As [a] was not used for the *test* phase, the stimuli in the practice session had [a]-[i], [a]-[ε] and [a]-[i] in the Word. The same template was used. A full list of stimuli can be found in [Appendix B](#). There were also two repetitions for each Word with flipped Options, just as in the *test* phase.

For each group of stimuli, participants listened to a triplet that consists of: A) a disyllabic word with two different vowel nuclei (Pre-contraction Word), B) a contracted, single syllable whose nucleus is one of the two vowels from the original disyllabic word (Option 1), and C) another contracted syllable with the other vowel nucleus (Option 2). They then decided between Option 1 and Option 2 clicked on the better contraction of the “Word” they just heard, with the first Option the “A” choice and the second Option the “B” choice. 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. All test stimuli triplets were randomly presented twice. All participants received the same stimuli.

3.2.3 Results and analysis

3.2.3.1 The two models compared

First of all, with regards to the linear position of the two vowels in the pre-contraction form, did the participants have a preference for the vowel on the left or the vowel on the right? Since the experiment was a forced choice task between two options, there were equal numbers of left vowel choices and right vowel choices. A hypothetical mean proportion of choices of 50%

and 50% for each pair of vowels and a similar distribution of data will be expected if null effects were assumed. Below and hereafter, I will present the results with plots that combine a boxplot and a violin plot. The boxplots show the means for each comparing variable, with connected lines and dots each representing one participant. The violin plots show the distribution of data; the wider the violin shape is, the more data there are in a certain range. Each line refers to a single participant. Each dot represents a single data point, which is the average proportion of certain choice for this participant.

Figure 6 shows the left and right choices of each participant. Overall, it suggests that participant's proportion of choices of left and right vowels are a near-even distribution. First, the mean proportion of responses for the left vowels and right vowels are both close to 50%, with the proportion of left vowels slightly higher. A paired Wilcoxon signed rank test³² comparing the proportions of the left choices and the right choices for each participant was conducted, and the result showed that there is no statistically significant difference between the left and right vowel choices [$V = 485.5$, $p = .48$ ³³]. An examination of the distribution of data (the violin plot) suggests slightly more data is on the >50% range for the left vowels than the right vowels, the latter having more data on the <50% range. An outlier participant chose the left vowel for more than 90%. Excluding this outlier participant will not shift the conclusion of the statistics [$V = 444.5$, $p = .65$]. It is safe to say that in this experiment, the left preference is not a bias, at least

³² For the statistics, Wilcoxon tests were run for this entire experiment because it does not make a normality assumption about the errors.

³³ P-values are Bonferroni-corrected. The same was done for all the Wilcoxon tests in this chapter.

not a *consistent* bias. There must be factors other than the relative positions of the pre-contraction vowels that have influenced participants' vowel choices.

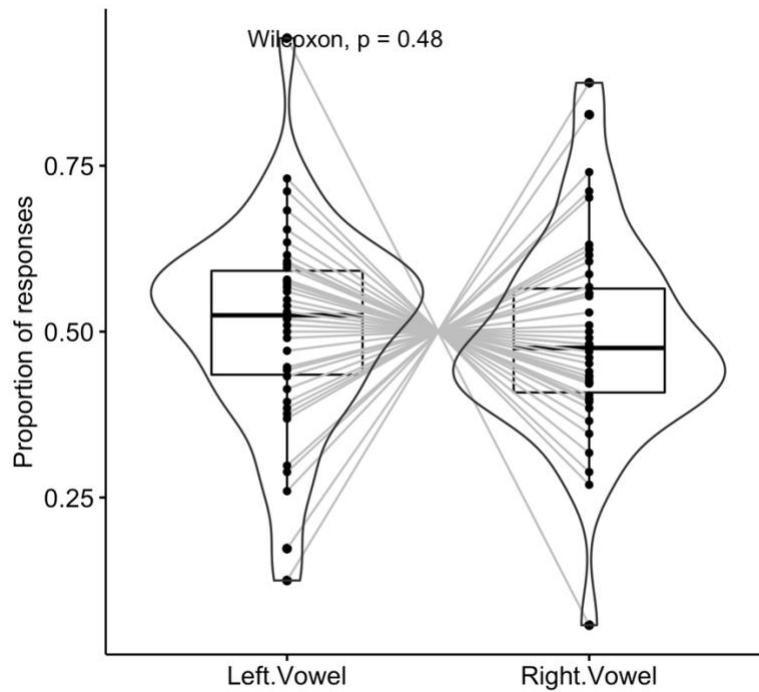


Figure 6. Proportion of responses for vowels on the left and vowels on the right. [Each line refers to a single participant. Each dot represents a single data point.]

Next, the sonority-based model is examined using the assumed sonority hierarchy based on height and centrality: $[\text{æ}] > [\text{ɛ}] > [\text{i}]$, $[\text{ɔ}] > [\text{u}]$ and $[\text{æ}, \text{ɛ}, \text{ɔ}, \text{i}, \text{u}] > [\text{ə}]$. Each comparing pair is coded according to the sonority hierarchy. For example, for $[\text{æ}]$ vs. $[\text{ɛ}]$, $[\text{æ}]$ is coded as more sonorous and $[\text{ɛ}]$ is coded as less sonorous. Figure 7 shows the proportion of responses to the more sonorous vowel and the less sonorous vowel based on the assumed sonority hierarchy, each line representing one participant. It suggests that there are generally a higher proportion of responses to the more sonorous vowels. A paired Wilcoxon test was conducted on the difference between the two sonority groups. The result shows that the difference between the two groups is

statistically significant [$V = 750.5$, $p < 0.0001$]. The proportional differences were relatively small, as most data cluster between 50% and 70% for the more sonorous vowels. However, considering the difficulty of the experiment, i.e., using nonce words for contraction on the fly while most speakers are not aware of contraction, such difference is still substantial. As expected based on random variation, a few participants went in the opposite direction and chose a larger proportion of the less sonorous vowels (Figure 7). They will not change the conclusion or the statistics, however, as they were too few compared to all the other participants who chose more sonorous vowels.

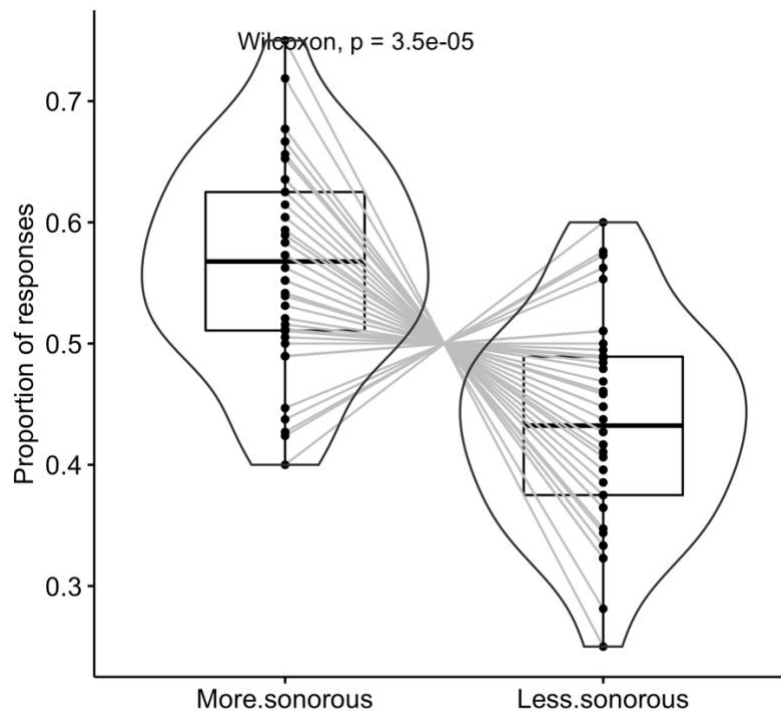


Figure 7. Proportion of responses of more sonorous vowels and less sonorous vowels. [Each line refers to a single participant. Each dot represents a single data point.]

The data presented here confirms that sonority in general plays a role in biasing participants' vowel choice for the contracted syllable, and that the participants were more likely to choose the more sonorous vowel than the less sonorous vowel. Such sonority bias is independent of the relative position of the vowels in the pre-contraction form, as the positions were counterbalanced. This bias is also independent of the order that the two contraction options were presented, as the order Option 1 and Option 2 was also flipped.

In summary, the linearity-based analysis is not confirmed by the data of this experiment, as the left vowel bias is too weak to be meaningful and absent on the exclusion of just a few outlier participants. The sonority-based model explains the experimental data well by revealing the preference of the more sonorous vowels. Next, I will examine the dimensions and aspects of vowel sonority respectively and find out whether the sonority bias hold in each one of them.

3.2.3.2 Height

Among the features that may have to do with sonority, height (or F1 value) is the predictor that most studies have agreed with (Gordon et al., 2012; Parker, 2008). Based on the universal hierarchy of vowel sonority, the ranking of low > mid > high is used for coding the stimuli. Five vowels are involved for vowel height testing: low [æ], mid [ɛ], high [i], high-back [u] and mid-back [ɔ]. Whether a vowel 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 example, for the Mid vs. Low comparison, [æ] is coded as the lower, more sonorous vowel, and [ɛ] is coded as the higher, less sonorous vowel. Figure 8 shows the proportions of choice of vowels of relative height and suggests that, generally, a higher proportion of choices was made to the lower, more

sonorous vowels than the higher, less sonorous vowels. A paired Wilcoxon test was conducted on the difference between the two height groups. The result shows that there was a statistically significant effect of height-based vowel sonority on vowel selection in syllable contraction [V = 20956, $p < 0.0001$].

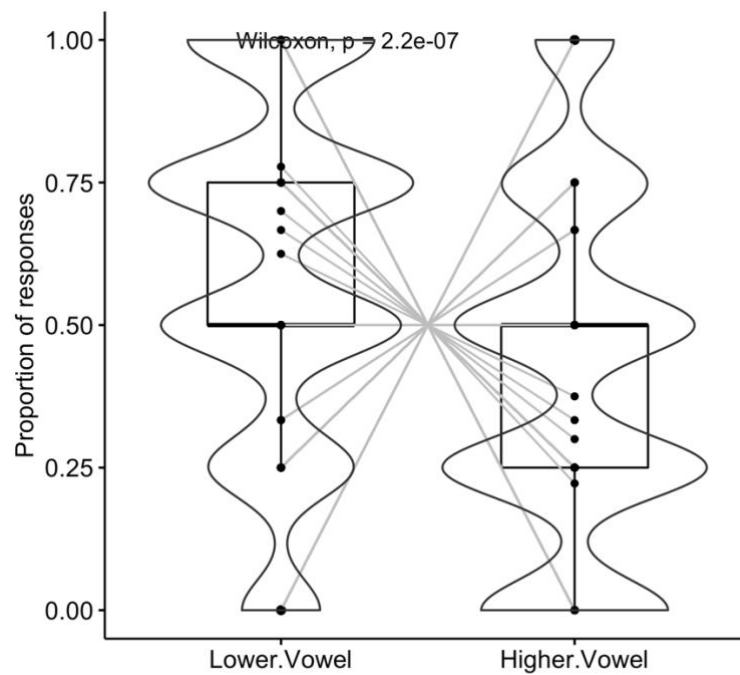


Figure 8. Proportion of responses of vowels of different height. [Each line refers to a participant; each dot represents a single data point. Some lines and dots are overlapped.]

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, as shown in Figure 9. Participants were consistently more likely to choose the comparatively lower vowels when other vowel features were the same. Individual results for each comparing group are summarized below:

- [æ] > [ɛ] (Low > Mid). Although the mean proportion of choices of the two vowels are identically 50%, the distribution of data is visually distinct, with data for [æ] more on the >50% side and more data for [ɛ] on the <50% side. A paired Wilcoxon test confirms that the difference between the proportion of choices of these two vowels is statistically significant [$V = 1240$, $p < .05$].
- [æ] > [i] (Low > High). The mean proportion of choices for [æ] was higher than that of [i]. The distribution of data also shows distinction, with more data for [æ] above 50% and more for [i] below 50%. A paired Wilcoxon test shows that the difference between the proportion of choices of these two vowels is statistically significant [$V = 1568$, $p < .05$].
- [ɛ] > [i] (Mid > High). The mean proportion of choices for [ɛ] was higher than that of [i]. Most data are clustered between 50% and 75% for [ɛ], while most data are clustered between 25% and 50% for [i]. A paired Wilcoxon test shows that the difference between the proportion of choices of these two vowels is statistically significant [$V = 1269.5$, $p < .001$].
- [ɔ] > [u] (Mid > High). Although the mean rate of responses are both 50%, the distribution of data is different, with more data for [ɔ] above 50% and more for [u] below 50%. A paired Wilcoxon test shows that the difference between the proportion of choices of these two vowels is statistically significant [$V = 1246.5$, $p < .05$].

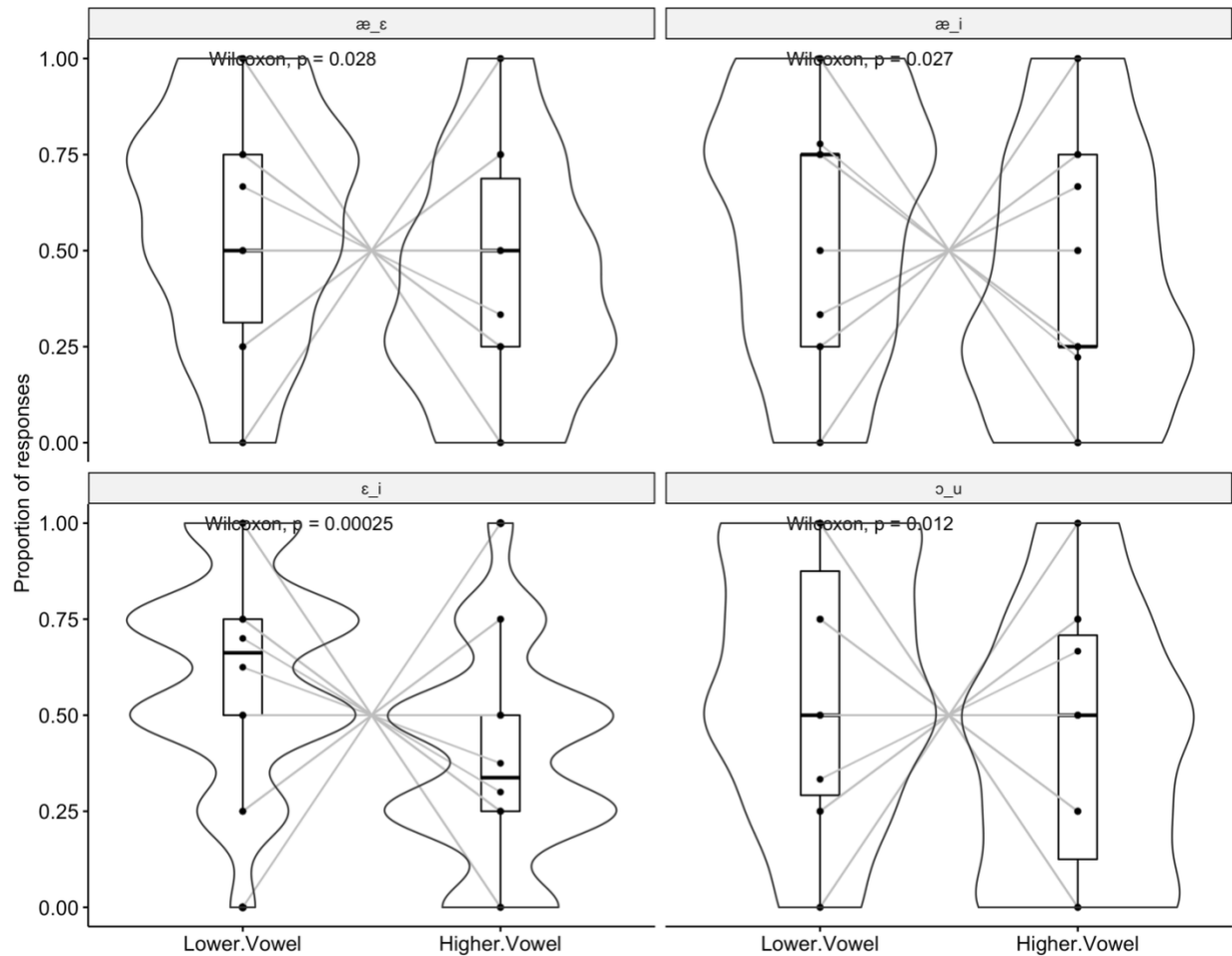


Figure 9. Proportion of responses arranged by height group. [Each line refers to a participant; each dot represents a single data point.]

In summary, the preference of lower vowel over higher vowel is not only generally true, but also consistently found in all comparing groups. Both the front, unrounded vowels and the back, rounded vowels exhibit the same height effect. The next question is whether there is any bias for vowels of the same height.

For vowels of the same height, the feature that can be tested is either roundedness or backness, as the back vowels in the experiment are all rounded vowels. Supposing there is a

sonority ranking of [u] > [i] and [ɔ] > [ɛ], the high vowels and mid vowels are compared respectively. The result here shows a general preference of rounded (or back) vowels over unrounded(front) vowels (Figure 10). A paired Wilcoxon tests confirms that there is a statistically significant difference between responses of rounded (back) vowels and those of unrounded (front) vowels (V = 1341, $p < 0.01$).

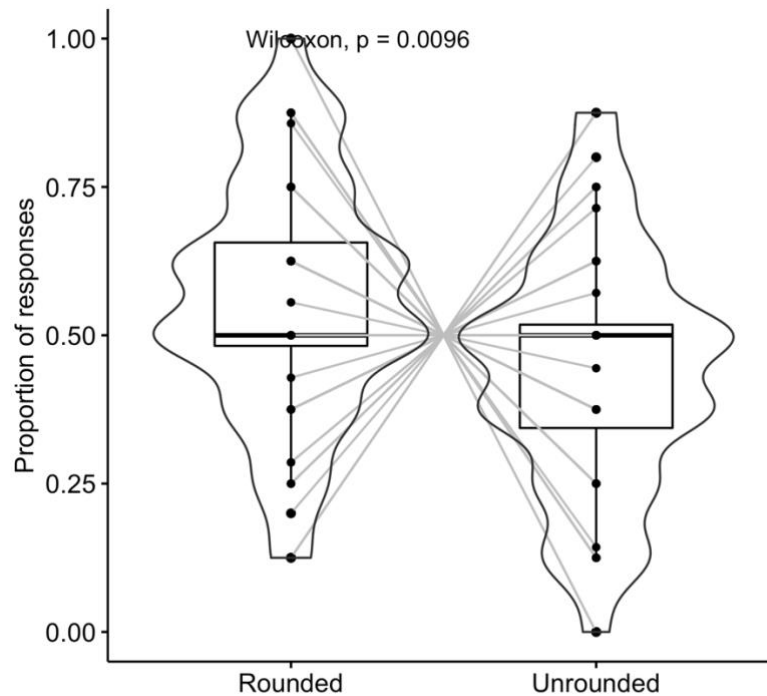


Figure 10. Proportion of responses of rounded/back and unrounded/front vowels. [Each line refers to a single participant. Each dot represents a single data point.]

Despite the general trend, however, details for rounding (or backness) comparisons points in two different directions, as shown in Figure 11.

- [ɔ] > [ɛ] (Rounded/back > unrounded/front). The mid vowel comparison is parallel to the general conclusion as in Figure 10 above in both means and data distribution. A paired

Wilcoxon test confirms a statistically significant difference [$V = 347.5$, $p < .01$] between the mean proportion of choices of mid-back and mid-front vowels.

- $[u] \approx [i]$. For the high vowels, there is simply no difference between $[i]$ and $[u]$ either in mean proportion of responses [$V = 322.5$, $p = .45$] or the distribution of data (symmetric distribution). Such null result might have been due to the fact that $[u, i]$ are both high vowels. See more discussion on high vowels in 3.3.2 later in this chapter.

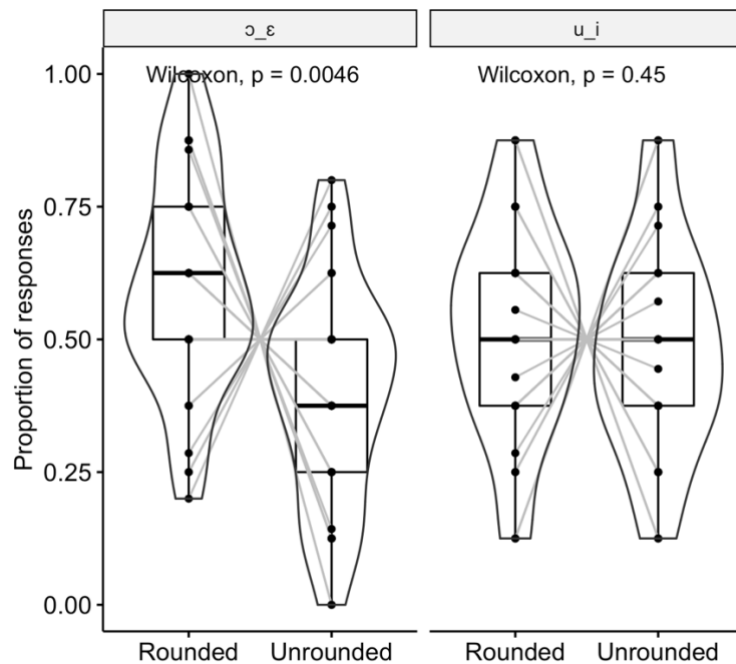


Figure 11. Proportion of responses arranged by roundedness/backness. [Each line refers to a participant; each dot represents a single data point.]

3.2.3.3 Centrality (peripherality)

According to the typically proposed sonority hierarchies, there is a hypothetical ranking of $[i, \epsilon, \text{æ}, u, \text{ɔ}] > [\text{ə}]$, with each peripheral vowel being more sonorous than the central vowel.

Each peripheral vowel is compared against the central vowel [ə] respectively and the proportions of choices are shown in Figure 12. As shown in Figure 12, the mean proportion of response is higher for peripheral vowels and the central vowel, which means the participants were generally more likely to choose peripheral vowels than the mid-central vowel. A Wilcoxon test was conducted to compare the proportions of a peripheral vowel and those of the central vowel and confirmed a statistically significant difference between the responses of peripheral vowels and the central vowel [$V = 9949$, $p < .0001$].

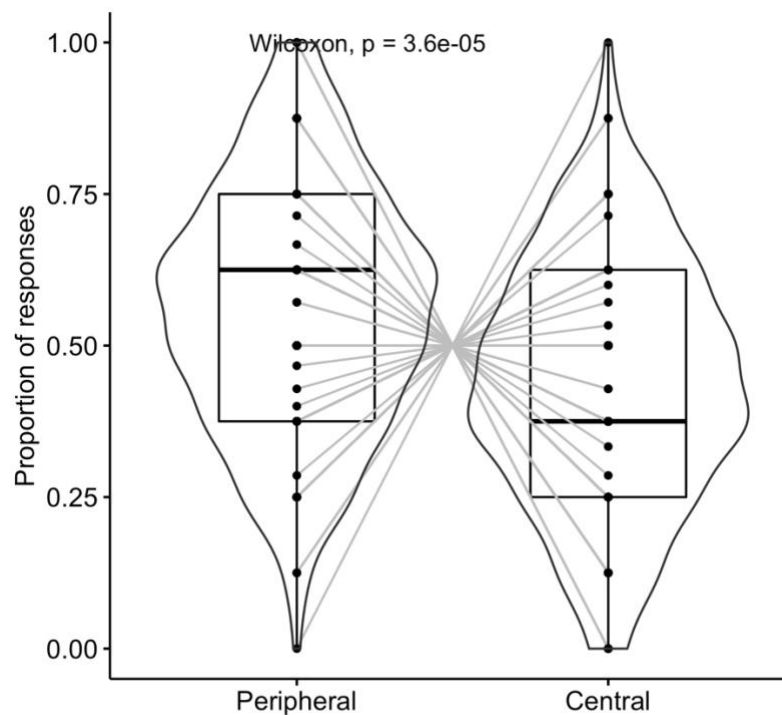


Figure 12. Proportion of responses to peripheral and central vowels. [Each line refers to a single participant. Each dot represents a single data point.]

In addition to the general effect of centrality, the preference for peripheral vowels is generally consistent, but with a few exceptions. Figure 13 shows the proportion of responses in each peripheral vowel and [ə] comparison. Results for each comparison are summarized below:

- [ɔ] > [ə]. The proportion of responses of mid-back [ɔ] is on average higher than that of [ə]. The majority of data for [ɔ] is also higher in value than the data for [ə]. A paired Wilcoxon test confirmed that the difference between the proportion of responses of [ɔ] and [ə] is statistically significant ($V = 598.5$, $p < .001$).
- [u] > [ə]. High-back [u] was on average chosen more than [ə], with choices for [u] generally above 50% and most data for [ə] below 50%. A paired Wilcoxon test confirmed that the difference between the proportion of responses of [u] and [ə] is statistically significant ($V = 385$, $p < .01$).
- [i] > [ə]. The mean proportion of responses for high-front [i] and [ə] are both 50%. However, more data for [i] is distributed in the upper values than [ə]. A paired Wilcoxon test shows that the proportions of responses of [i] and [ə] are significantly different ($V = 355$, $p < .05$).
- [æ] $\approx$ [ə]. The responses to low vowel [æ] and [ə] are generally no different in both the mean proportion of responses (50% vs. 50%) and the distribution of data. A paired Wilcoxon test confirmed that the difference between the proportion of responses of [æ] and [ə] is not statistically significant ($V = 273$, $p = .90$).
- [ɛ] $\approx$ [ə]. Except for a few outliers, the mean proportion of responses (50% and 50%, respectively) and the distribution of data are both near-identical for mid-front [ɛ] and [ə]. A paired Wilcoxon test confirmed that the difference between the proportion of responses of [ɛ] and [ə] is not statistically significant ($V = 323.5$, $p = .89$).

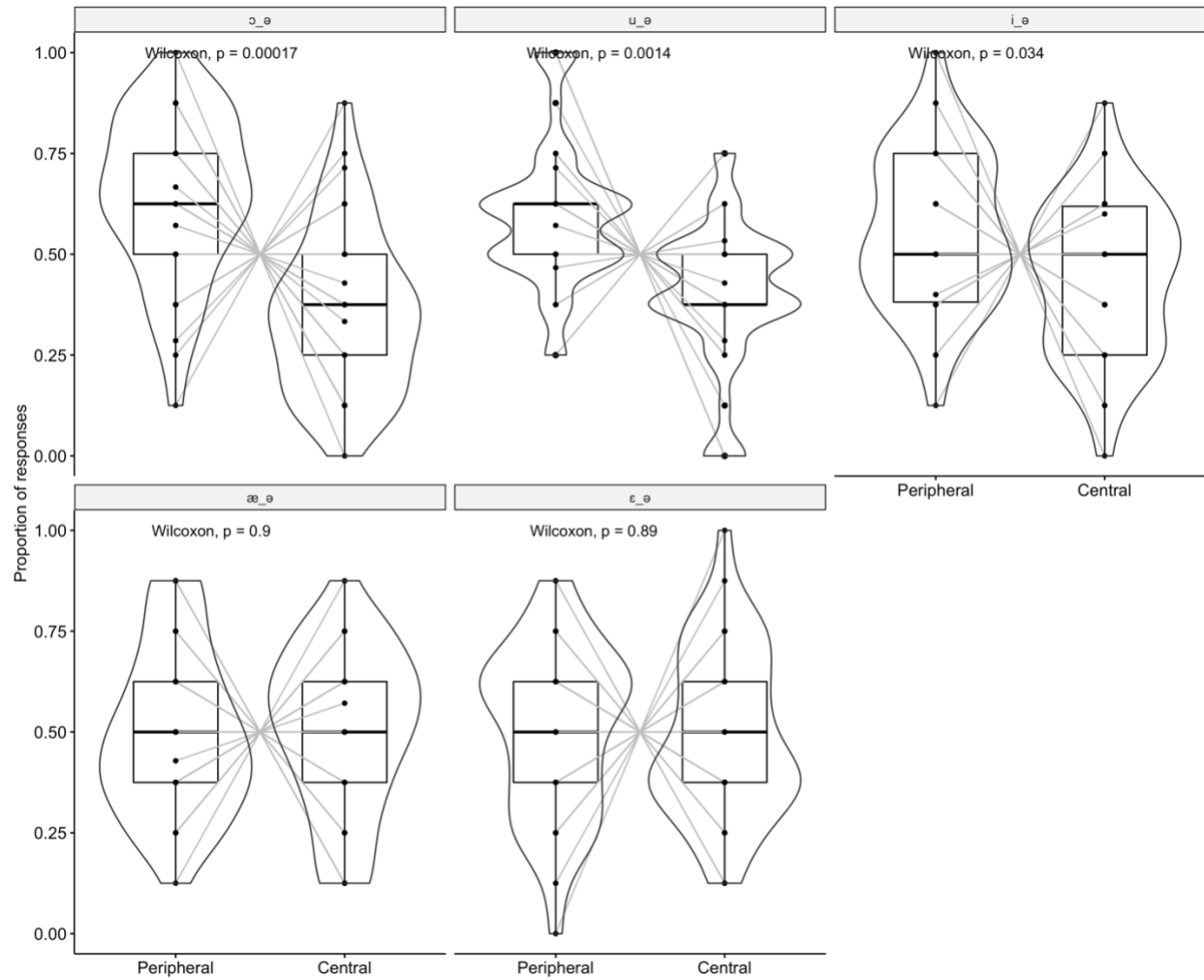


Figure 13. Proportion of responses to vowels arranged by peripheral/centrality. [Each line refers to a participant; each dot represents a single data point.]

In summary, the peripheral vowels are generally more likely to be chosen for the contracted syllable than the central vowel, but the peripherality bias is not consistently found across all the comparing groups. A closer examination of each comparing group suggests that the general trend holds for three groups: High-back vs. Central, High-front vs. Central and Mid-back

vs. Central. However, the Low vowel and Mid-front vowel did not receive more responses than the central vowel.

3.3 Discussion

3.3.1 Sonority in syllable contraction

This experiment examined Rugao speakers' vowel selection preference 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 other order-based preference. The difference between the more sonorous vowels and less sonorous vowels may seem small. The largest difference is between 75% and 25%, while most groups yield a 60% vs. 40% to a 65% vs. 35% difference; some comparing groups even had a mean 50% and 50% for the two options, although statistical analyses confirmed the difference between the two sonority groups, as the data for the more sonorous vowel is generally on the higher end while data for the less sonorous vowel is on the relatively lower end. Despite being small, the difference is consistent and in the same direction. The same preference for more sonorous vowels is found in the majority of comparing groups and in both aspects of vowel sonority. More importantly, syllable contraction occurs in casual speech and mostly in high-frequency words and phrases. This could have made the experiment a relatively difficult task for the participants, 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.

Height and centrality are known to contribute to 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, which is exactly how the sonority is ranked. Second, the results presented above clearly show that high vowels, the supposedly least sonorous among all the height groups, were less preferable to mid vowels and low vowels respectively. This is true for both the rounded and unrounded vowels respectively. A ranking of [ɛ] > [i], [æ] > [i], and [ɔ] > [u] (“>” means vowel receiving more responses) suggest that [i] and [u] lose to all other non-high vowels with the same roundedness/backness feature. Casali (1998) argued for a [+high] vs. [-high] contrast for the vowel sonority ranking based on the vowel coalescence data. However, the results of this experiment unmasked a detailed sonority ranking of [-high] vowel as well, as mid and low vowels competition was also in the same direction, with the more sonorous low vowel more likely to be the winner: [æ] > [ɛ]. Due to the limitation of the vowel inventory, only mid-front and low-front vowels were compared because there is no corresponding low-back vowel in Rugao. Assuming the relation between vowel sonority and vowel selection in syllable contraction, the fact that there is a detailed ranking of vowel height preference of low > mid > high supports a detailed ranking of vowels of different height (de Lacy, 2010; Kenstowicz, 1994; Ladefoged & Maddieson, 1996; Parker, 2008, 2011, 2012). This further suggests that, rather than only distinguishing [+high] and [-high], one should consider the full vowel ranking based on the more general concept of *height* instead of one specific feature, such as [high] or [+high]/[-high].

For vowels of the same height, the back and rounded vowels are more likely to survive in the contraction. For the mid vowels [ɛ] and [ɔ], [ɔ] is preferred to [ɛ] in the nonce word contraction. The null difference observed for the high vowels [i] and [u] might have been due to the Glide Formation of high vowels. See more discussion for high vowels in 3.3.2. Unfortunately, the other two mid vowels of Rugao [e] and [o] cannot be tested using this experiment design due to the syllable requirements, as [e] and [o] cannot be followed by a nasal

coda. On the basis that the vowel sonority and vowel competition are correlated, the rounded/back vowel bias that is found in the experiment may imply that the sonority ranking of vowels of the same height can be made, at least for mid vowels [ɔ] > [ɛ], i.e., back > front or rounded > unrounded. Previous analyses have different rankings of vowels of the same height. Some studies do not rank them, for example, [a] > [e, o] > [i, u] (Parker 2002, Selkirk 1984) while some studies rank front vowel above back vowels, e.g. [a] > [e] > [o] > [i] > [u] (Kiparsky, 1979). Studies of syllable contraction of other Chinese languages/dialects also seem to indicate a language-specific variability for vowels of the same height. Taiwanese Southern Min has a hierarchy of [a] > [ɔ] > [e] > [o] > [i] > [u] based on syllable contraction data, which suggests front > back (Hsu, 2003). But Mandarin syllable contraction data seems to suggest [o] > [ɛ], which is back > front (Tseng, 2005a). There is also no sonority hierarchy that clearly distinguishes between backness and roundedness. For the limitation of the Rugao vowel inventory, there is no way to tease apart the backness and roundedness for this experiment. There may be a group of phonological features and phonetic cues that contribute to the sonority differences among vowels of the same height. These cues can be so weak that languages can choose freely, resulting in the great variation among languages. More research is needed for both the vowel competition and vowel sonority of vowels of the same height, possibly studies that utilize languages with a more contrastive vowel inventory and another experiment design.

Second, peripheral vowels were generally more preferable to central vowels. The non-high central vowel [ə] was generally less preferable to peripheral vowels [æ, ɛ, ɔ, i, u]. However, such peripheral bias is less robust as seen for the height bias, as only three out of five comparing groups exhibit an observable difference between the peripheral and central vowel, while no preference was found for [æ] and [i] over [ə]. Gordon et al., (2012) concluded that centrality is a

more consistent predictor of sonority across the languages being studied compared to the commonly accepted correlation of sonority. The somewhat inconsistent data of this experiment seems to pose a potential challenge to the status of centrality in the vowel sonority. However, the fact that the centrality bias exists in a weaker way in syllable contraction might be the result of the phonological and phonetic status of [ə] in Rugao, which in turn indicates that it may not be really weak in sonority. First, as stress is not contrastive in Rugao, [ə] does not have to appear in an unstressed syllable as it typically does in languages such as English (e.g., [ə'baʊt] <about>, [ə'phʌn] <upon>). This is not rare cross-linguistically even in languages that do have contrastive stress, such as in Lushootseed, which allows stress to fall on the first schwa if there is no “full” vowel in the stem (Urbanczyk, 2006). Second, correspondingly, many phonetic attributes that make [ə] phonetically weak in other languages may not apply for Rugao. As can be seen in Figure 14, the Rugao [ə] is not shorter than other vowels, such as [æ], when both vowels are followed by a nasal coda. Also shown by Figure 14, the [ə] does have a lower intensity than [æ], which will indicate that [ə] is somehow weaker than [æ]. The intertwined contributors of sonority might have caused the more complicated data for the central vowel bias, which in turn suggest that vowel sonority is more complex than simply two or three dimensions. I will discuss this issue in more detail in Chapter 5.

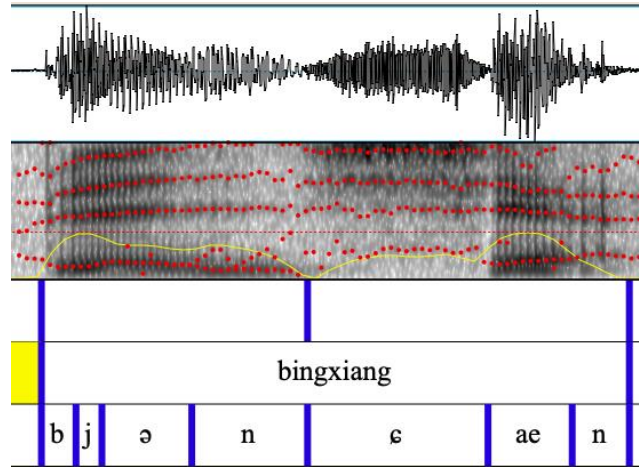


Figure 14. Praat screenshot for [pjən.ɛæn] ('fridge'). [Note: the Praat annotation may not be accurate for ease of annotation.]

The mid-central vowel [ə] is typically ranked higher than the mid high vowel [ɪ] in the sonority scale, but this experiment was not able to test whether there is a bias between the two central vowels. No structurally consistent stimuli can be used for [ɪ] because the high central vowel cannot be followed by a coda. The phonetic properties of [ɪ] in Rugao are also not clear, as it sounds more fricative and apical than the commonly seen [ɪ] in other languages. With the case of [ə] already being complicated, there needs to be better ways to test the details of the competition between the central vowels, and I will leave this for future projects.

3.3.2 Remaining issues

The data that are presented in this chapter are generally easy to interpret except for a few comparisons. In particular, data for the comparisons that involve high vowels seem less consistent than those without (Figure 11, Figure 13), as the vowel preference predicted by sonority is either small or simply not observable. This is probably because of the Glide Formation process that satisfies the Maximal Syllable constraint: contracted forms follow

Maximality (McCarthy & Prince, 1994) and construct the largest possible syllable (Hsu, 2003; Xu, 2014). One way to keep the maximal syllable in the contracted output is to form a glide whenever possible, especially when the high vowel is on the left. High vowels [i, u] are possible candidates for the glides [j, w]. For example, real word [p_hin₂₁ + ɛən₂₁] (‘favor’) is contracted as [p_hjən₂₁], preserving the high vowel as a glide. None of the stimuli with high vowels undergo glide formation as a sacrifice to keep the CVC syllable structure template consistent. For the [i]/[ə] competition in the experiment, the stimulus [p_hin + xən] has a more natural contraction form [p_hjən], but only [p_hin] and [p_hən] were given. For the [i]/[u] comparison, stimulus [p_hin + xun] has a more natural contraction [p_hjun], which was not one of the given options ([p_hin] and [p_hun]). Meanwhile, stimulus [p_hun + in] can be contracted as [p_hwin], despite the fact that [p_hwin] is an illicit syllable in Rugao. As a consequence, participants might have been reluctant to choose between the given two options without glide formation because neither of them was an optimal choice. This could have shifted the data closer to the 50%/50% distribution and decreased the difference that otherwise might have been more observable.

This experiment showed a null effect of order-based preference. Contrary to the real word data that exhibits either a Left vowel or Right vowel preference (R.-F. Chung, 1996; Sun, 2014), the participants of this experiment did not have a bias for either left or right vowels. But the experimental data alone is not enough for us to make the statement that LR-scanning is simply missing from the vowel selection in syllable contraction in general. In fact, there is a substantial amount of inter-subject variation and nearly half of the participants chose the left vowel and other half chose the right vowel. Although it seems like an even distribution, it is unclear whether the variation is truly free or influenced by the sociolinguistic background of the speaker. The variation question is beyond the scope of this project but can be answered by future studies,

where more demographic data could be collected. An alternative interpretation of the experiment result is that the linear order of the vowels in the syllable is not the most predictable factor, but a very low-ranking one that affects the vowel output choice only when needed. The experiment also set the stimuli in a strictly controlled frame and used nonce words. Such design might also have masked the possible effects and discouraged the use of some tools that are otherwise utilized in the real word contraction. Nonce words are good tools for testing the vowel pattern but are different from real words after all. Furthermore, this experiment tested on Rugao contraction. Whether the analyses on the two models can extend to other languages still needs further study. It still remains a question whether the sonority-based preference is language-specific. Given that the sonority hierarchy can well be language-specific, it would not be surprising to find the sonority-based selection is so as well. These open questions call for more study on the linear-order effect.

This experiment tested on a very specific type of syllable contraction, i.e., CVN + CVN → CVN, with clear consonantal Edge-In, single vowel selection, and the same low-falling tone. Such manipulation makes it possible to test on a single aspect of this process while keeping the others consistent but oversimplifies the actual complexity of syllable contraction. Many aspects of phonology and phonetics can be involved in the process of syllable contraction. The sonority-based vowel selection only predicts the winner and loser of the vowel competition that are more likely, but it does not guarantee the winner of the competition. There are still real word cases that are against the sonority-based predictions; as discussed before, [nun.xow] → [nuw]/[now] ('warm') seems to support both vowels being selected, although [u] should be ranked lower in sonority than [o].

There are other more general issues regarding the limitations of the corpus data as well as the experimental data, and I will come back to these questions later in this dissertation. This is the first study of vowel competition and vowel sonority using experimental data. Although it simplifies the issue, it provides a first step towards understanding the syllable contraction data in Rugao and the vowel selection mechanism in the vowel competition, in particular.

3.4 Conclusion and implications

This chapter is based on the phonological analysis of the role that vowel sonority plays in the vowel selection process of syllable contraction using real word Rugao syllable contraction data. Using a nonce word experiment, I confirmed the sonority-based analysis but did not find supporting evidence for the linear-order based analysis. With all other possible factors well controlled, when two vowels compete for the limited vowel nucleus position in the contracted syllable, the more sonorant vowel is more likely to be selected than the less sonorous vowel. More specifically, sonority hierarchies of both height and centrality exhibit a consistent and robust effect in biasing the vowel selection, with a preference ranking that goes in line with their sonority hierarchy. The roundedness (or backness) also affects the vowel competition but requires more study.

The vowel that gets preserved in di-syllabic contractions is essentially the survivor of two vowels in competition for a single timing slot. Vowel competition may go beyond the realm of syllable contraction and generalize any vowel-related phonological processes in a broader perspective. The study of syllable contraction provides new insights to understanding the competing vowels and how the survivor is determined. This experiment, in particular, offers a

strong tool, i.e., the vowel sonority, that needs to be taken into consideration when vowel competition is involved.

Sonority as a single property is thus a way to unify different vowel features, height, centrality and possibly roundedness or backness, in accounting for the corpus and experimental data. In stress attraction, no stress system could refer *only* to height features, which means simply specifying low vowel or high vowel will not suffice in predicting the place of the stress (de Lacy, 2010). Similarly, in the case of syllable contraction, no single feature can account for all the patterns; one cannot take a height feature like [low]/[mid] or a roundedness feature [rounded] to predict the contracted output. If the vowel selection pattern is simply due to markedness constraints, we would need an array of features in order to determine one single output. Sonority is not a subsegmental feature. In contrast, it behaves like manner features and does not have a root node (McCarthy 1988). Sonority functions as a relative manner: instead of specifying one vowel feature, it generalizes over a couple of features and “floats”.

In turn, the study of syllable contraction and vowel competition may offer baseline data for the study of vowel sonority. Compared to the vast literature of consonant sonority, vowel sonority receives much less attention, partially due to the shortage of empirical evidence. The vowel sonority hierarchy is also less agreed upon, with different studies providing different rankings. The study of syllable contraction provides both real word data and experimental data to support a more detailed ranking of vowel sonority. Sonority is known to be a mixture of several dimensions and aspects, including height and centrality. Based on the findings of the experiment, height seems to be the most stable predictor of vowel sonority, with all data consistently pointing to the same direction of low > mid > high. Centrality exhibits a pattern of peripheral > central, but with less supporting data. As discussed before, the status of central vowels may vary cross-

linguistically. As sonority consists of different but intertwined aspects, it is not surprising to find inconsistency for the centrality bias. Last but not least, whether backness predicts the sonority ranking of vowels of the same height is debatable. The syllable contraction data implies that there can be a detailed ranking of vowels of the same height as back > front or rounded > unrounded, but it could not pinpoint backness or roundedness for the limitation of the experiment design.

This experiment, as well as the sonority-based vowel selection in general, treat the vowel selection and deletion as complete processes and stipulates that one vowel must survive and the other must delete. However, the sonority-based vowel selection does not make any prediction on the detailed phonetic properties of the surviving vowel or even the deleted vowel. Therefore, in the next chapter, I will introduce another set of experiments that specifically test on the phonetic properties of the contracted syllable with a focus on the vowels.

4 THE PHONETICS OF VOWELS IN SYLLABLE CONTRACTION

4.1 Introduction

4.1.1 Motivation for this study

As seen in Chapter 1, phonetic studies of contracted syllables are not abundant, and most studies were focused on the gradience of contraction, and/or the factors that affect the degree of contraction (Cheng & Xu, 2009; Myers & Li, 2009). It is a clear conclusion based on these studies that the syllables undergo different degrees of contraction. Wong (2006) adopted the qualitative approach and categorized two basic forms of syllable fusion on the continuum of contraction: *bisyllabic fusion*, where there is “the deletion of at least one segment contiguous to the syllable boundary between two syllables without affecting the vowel count” and *vowel coalescence*, which changes the syllable count, is detected in cases where “a single intermediate quality or deletion of one of the vowels” results in a “merging” of vowel qualities of adjacent syllables. This way of categorization is based on the examination of the spectrograms by the author. Differently from Wong (2006), Myers & Li (2009) adopted a quantitative approach by using the *trough depth* measurement (Mermelstein, 1975), an algorithm for detecting syllable boundaries, to define the degree of contraction. The *trough depth* measures the maximum difference in intensity between the convex hull and the intensity levels at the syllable boundaries. The measurement in dB in turn defines whether the two syllables are partially contracted or fully contracted in a way that is independent of human judgment and the phonotactics of the language. Kuo (2010b) studied the fully contracted syllables in Taiwan Mandarin and defined a fully contracted syllable as those that share the same syllable structure with an actual lexical word and also used the *trough depth* measurement to determine the degree of contraction. She also pointed

out that, in this sense, the process of syllable contraction can be regarded as a process of neutralization.

The scarcity of existing research on the detailed phonetic properties of the contracted syllable leaves a large potential for further studies and many questions unanswered. First, what does a contracted syllable look like? More specifically, is it generally longer or shorter compared to a ‘normal’ lexical syllable? Is any constituent of the syllable structure longer or shorter than a lexical counterpart? For example, does the nucleus consistently take a bigger proportion of the contracted syllable? Second, are the segments the exact same as their lexical counterparts? For example, is the [a] in [ɿa], the contraction form of [ɿəŋ.ka] (‘other’) the exact same vowel as in [ɿa] (‘hassle’)? One could expect there to be some vowel coalescence in the output of contracted syllables, based on previous studies in Cantonese and Mandarin (Myers & Li, 2009; Wong, 2006). This means that, first, both vowels in partially contracted syllables should exhibit some differences from the underlying vowels, and second, the vowel may even change to a different category in this process. Furthermore, the possibility of vowel coalescence may indicate that the one and only vowel in a fully contracted syllable can also be significantly different from a lexical vowel in some ways, including vowel quality. On the other hand, if the process of contraction, at least for the fully contracted cases, is true neutralization, there shouldn’t be much difference between a contracted syllable and a corresponding monosyllabic word with regards to the syllable structure, lengths and quality of segments in production, and listeners should not be able to distinguish a fully contracted syllable with a lexical syllable with seemingly same segments and tones (Kuo, 2010).

Yet, based on the vast literature on *incompleteness* of phonological changes (Ellis & Hardcastle, 2002; McCollum, 2019; Port & Leary, 2005; Warner, Jongman, Sereno, & Kemps,

2004), there is a third possibility: there *is* some neutralization in the phonological process, but the neutralization is incomplete, in that the final output and the comparing unit are discretely different in aspects that are not easily captured by human ears. Such incompleteness makes the fully contracted syllables similar enough to a lexical word, but fine phonetic differences are retained in order to maintain the underlying lexical contrast. The production and perception studies have revealed subtle phonetic cues that are preserved in the spoken output, and listeners can make use of these cues to make distinctions. Take the most famous case of German final obstruent neutralization for example (Port, Robert F. & O'Dell, 1985; Port & Leary, 2005). Word pairs such as 'Bund' and 'bunt' seem to be pronounced the same [bunt] in German. As the UR for "Bunt" is /bund/, the final stop /d/ is devoiced and surfaces as [t], the latter also having a corresponding lexical UR /t/. The devoicing of the voiced final obstruent /d/ → [t] makes a voiced obstruent sound like the SR of another phoneme /t/, a process canonically called neutralization. However, modern experimental studies have revealed the many characteristics of the derived and neutralized obstruent are different from a lexically voiceless obstruent, with the one being neutralized having a longer preceding vowel, longer voicing into the closure, shorter closure duration, and weaker bursts (Port, Robert F. & O'Dell, 1985; Port & Leary, 2005). All of these differences point to the incompleteness of the neutralization. Similar results were found for similar process of consonant alternations in other languages such as Dutch (Ellis & Hardcastle, 2002). The incompleteness of the phonological process then expanded beyond final obstruent devoicing and is found in processes such as vowel harmony (McCollum, 2019) and Place of Articulation assimilation ((Ellis & Hardcastle, 2002).

In the context of Chinese-style syllable contraction, data of empirical analysis and production and perception experiments all indicate that, even if the same set of IPA symbols for

lexical vowels is used for the transcription of a fully contracted syllable, making contraction seem like a neutralization process, many phonetic properties of the contracted syllables are distinctive from lexical syllables (Kuo 2010, Wong 2006). It could also be a similar case for a contracted syllable vs. a lexical syllable such as in [mæn₂₁₃] (contraction of [ma₂₁.sæn₃₃], ‘right away’) vs. [mæn₂₁₃] (‘aggressive’) in Rugao, although the two syllables have the same segment and tones as well, the latter being proven to help both speakers and listeners maintain the contrast between a contracted and a lexical syllable (Kuo 2010, Wong 2006). The seemingly same vowel [æ] may not be exactly the same in the contracted form and in the underlying form, manifested in phonetic cues such as differences in duration and quality. Such difference, if it exists, can be captured by the acoustic measurements. If the incomplete neutralization is proven true in syllable contraction, it will connect syllable contraction with other cases of incomplete neutralizations in the literature in the fact that what seems to be ‘complete to the ear’ may be in fact incomplete. Kuo (2010) provides some data for this discussion, but uncertainty remains because of the inconsistency of the results (see more discussion below).

In addition, the phonological analyses presented in the previous chapters of this dissertation as well as in other studies are all based on the assumption that the impressionistic transcriptions made by the linguists faithfully reflect the pronunciations, including the deleted, the changed, and the remaining segments. However, the vast literature of the mismatch between the perception and the production informs us of the possibility of misperception, taking auditory illusion (Dupoux, Hirose, Kakehi, Pallier, & Mehler, 1999) and loanword adaptation (Davidson, 2007) as examples. In the context of syllable contraction and neutralization, the missing of subtle phonetic cues that cannot be possibly described by any phonetic symbols may present a problem for all the analyses, especially as the perception of speech signals may be influenced by the

phonological knowledge of the listener. As Port & Leary (2005) pointed out after investigating multiple cases of incomplete neutralization, even a well-trained phonetician can miss the subtle phonetic and sub-phonemic differences, such as differences in duration between neutralized lexical items. They agree with Manaster Ramer (1996) that one cannot and should not rely solely on the phonetic symbols based on anyone's auditory transcription. For this reason, we should use caution in the investigation of all cases of neutralization and look closer before concluding anything, especially for cases which look like "the same."

In summary, there needs to be more in-depth studies on the phonetic properties of the segments in the contracted syllables. This chapter is dedicated to analyzing the phonetic properties of contracted syllables, starting with vowels in the fully contracted syllables. Compared to the relatively good predictability of consonants, vowels may undergo more complicated changes in this process, such as deletion, lenition, length change, and quality change. Vowels thus provide first-step information for us to look closer at contracted syllables and the process of syllable contraction. In particular, a fully contracted di-syllabic word may sound similar to a monosyllabic word, making full syllable contraction seem like a process of neutralizing a disyllabic word and a monosyllabic word. The vowel in this extremely reduced form is thus comparable to a lexical vowel. For example, in Rugao, [sa_{l21}], the contracted form of [sə₂₁₅. a_{l21}] ('twelve'), sounds very similar to the monosyllabic word [sa_{l21}] ('spoon') to the ear of the author and transcriber. However, the seemingly same [sa_{l21}] may not be exactly the same phonetically. In particular, the two [a]'s may be similar enough to belong to the same vowel category /a/, but this does not guarantee exactly the same phonetic qualities. In addition, it is also unclear whether the seemingly deleted vowel, [ə] in this example, is completely gone as reflected by the transcription. The incomplete deletion could be discrete enough to escape human

perception. Intuitively, although there are no other choices of phonetic *symbols* to represent the contracted form, the contracted syllable still sounds somewhat different from a lexical syllable.

Kuo's (2010) series of perceptual and production experiments have provided pieces of evidence for such suspicion. First, listeners can distinguish between the fully contracted disyllabic word and the mono-syllabic lexical word. For example, [njaŋ₅₁] (contracted form of [na₅₁.jaŋ₅₁], 'that way') and [njaŋ₅₁] ('brew') are distinguishable from each other. Second, the contracted form preserves the length and the tones of the underlying disyllabic form. This means the contracted form is longer in duration than a monosyllabic lexical word although the fully contracted syllables are usually transcribed as single syllables. The tone is also distinguishable from a lexical tone, meaning that the contracted tone is often an illicit tone. This echoes the similar pattern of tone reservation in Cantonese found in Wong (2006) in that no matter what the degree of fusion/contraction, the basic tonal contour and pitch range both match the underlying non-contracted form to a large extent. The preserving of length and tone may provide reasons for why the perceptual differences exist for the listeners. It is likely that a cluster of differences is maintained in the contracted form, so that listeners can use these cues to make distinctions. The length and the tone that match the underlying disyllabic form are two cues of such kind. Rugao syllable contraction seems to have these two characteristics as well. Take another word 'then' [ɿn₁₅. xej₂₁] → [ɿej₁₅₁] for an example. First, the contracted form [ɿej₁₅₁] sounds longer than a lexical word and longer than either of the two single syllables in the pre-contraction form. While the consonants are roughly of the same duration, the vocalic portion can be bigger for the contracted form of [ɿej₃₅₁] than either syllable [ɿin₃₅] or [xex₂₁]. The tone on the contracted form [ɿej₃₅₁] is also an illicit tone, making this word stand out from a lexical word.

On top of duration and tone, there may also be cues on the segment itself. With regards to the vowel quality, Kuo's (2010) experiment found inconsistent patterns between male and female participants in the production of the vowels. For male speakers, vowels in contracted syllables have a wider vowel space: while [a] shows no difference, [i, u] in the contracted forms are higher than their lexical correspondents. No significant difference was observed for female speakers. However, the experiment did not control the deleting vowels. For example, in the contraction of [na₅₁.jaŋ₅₁] → [njaŋ₅₁] ('that way'), the deleted vowel and the surviving vowel are both [a], while in [pu₃₅.jaw₅₁] → [pjaw₅₅], [u] is deleted and [a] survives. It is possible that the deleting vowel either does not change anything for the surviving vowel because it is the same vowel, or that it 'pulls' the remaining vowel to different directions in the vowel space, which neutralizes the differences. A different picture may be seen if the deleting vowel is controlled.

In a controlled experiment, if the fully contracted syllables do have remnant elements from the deleted segments, the surviving vowel should differ from their lexical correspondents in some ways. Since the difference is too subtle to be captured by human ears, acoustic measurements of the vowel formants can shed light upon what the surviving vowel really is and whether the vowel undergoes quality changes in the process of contraction. Besides, if the deletion of a vowel is in fact incomplete, it could leave traceable cues, such as durational and quality differences, on the surviving vowel. The study of the latter can thus be used to infer the incompleteness of the vowel deletion. As a summary, the impressionistic analysis and relevant literature both suggest that the contracted vowels in the fully contracted syllables may be different from the vowels in lexical words. If proven correct, such a finding may provide evidence for the argument that the vowel deletion of syllable contraction is incomplete, and that the contracted form and the corresponding lexical form are incompletely neutralized. Finally,

these arguments together may implicate that the phonological representation is gradient rather than categorical (Port & Leary, 2005).

Finally, yet another motivation for collecting more naturalistic production data is to collect additional evidence for the centrality bias as shown in the previous chapter. In the sonority bias experiment (Chapter 3), I showed that height and centrality both shape vowel preference in a nonce word contraction experiment. The height preference low > mid > high is clearer with statistical evidence. Although the peripheral > central preference showed a predicted difference in the expected direction, not every pair of peripheral and central vowels in competition yielded the same statistical significance or the exact same pattern. The pattern pertaining to centrality is thus less clear. The patterns in the real word data collected via the production experiment proposed for the phonetic study can thus be used to compare with the nonce word forced choice data regarding vowel choice. If centrality bias is confirmed in the controlled naturalistic data, this experiment will lend more support to the sonority bias analysis.

4.1.2 A preview of the experiments

To find out about the phonetics of vowels in the syllable contraction process, two production experiments were run. The major goal of these experiments was to elicit comparable word tokens which are then used to compare the quality of the surviving vowel in the fully contracted word and the corresponding lexical vowel. For instance, is the [i] in [teɪn₂₁] (contracted from [tɛən₂₁.ɪn₂₁], ‘experience’) the same as the [i] in [teɪn₂₁] (‘pointy’)? This general goal can be divided into these following questions:

- a. Is a contracted syllable different from a lexical syllable with seemingly the same segments?

- Specifically, are there durational differences for these vowels? Does the vowel in a contracted syllable take up a bigger proportion than a lexical vowel does?
 - Is the vowel in the contracted syllable exactly the same quality (height, backness) as a corresponding lexical vowel?
- b. Does ‘fully contracted’ mean the other vocalic segment is completely deleted? Or are there remnants of the deleted vowel on the surviving vowel?
 - c. Broadly, what is a contracted syllable? What is syllable contraction?
 - d. Ultimately, does the case of syllable contraction support the categoricity or gradience account of phonological representation?

In order to elicit contracted items and their corresponding lexical words in a well-controlled while naturalistic setting, two major experiments were run, including: Experiment I, a natural sentence repetition task, in which participants were asked to repeat the sentence they hear, and Experiment II, a self-paced free word contraction task, followed by a non-contracted word list reading task, in which participants were expected to read each word clearly, presumably without contraction. The natural sentence repetition task elicited naturally contracted tokens without explicitly telling the participants to do so. The word contraction task explicitly prompted participants to contract words, and comparable controlled non-contracted words were elicited by the word list reading task. Although the basic stimuli, i.e., the vowels tested and words used, are kept consistent in the two experiments, the actual stimuli and procedures are different. By following the order of natural sentence repetition task → self-paced free word contraction task → word list reading task, more instructions were given while the experiment progressed, and the formality of speech is increased. The experiments were arranged this way for several reasons. First, as syllable contraction occurs subconsciously in natural speech, any

manipulation of the stimuli can discourage participants from contracting. Different methodologies were used to elicit as many contracted items as possible, and to compare the contracted and non-contracted items of the same or at least similar speech style. Second, it will be informative for understanding the details in the process of syllable contraction if any differences can be found among different types of speech. For example, vowel qualities or degree of contraction in a sentence may be different from those in single words. It will shed light upon how speech style affects the production. Third, as the participants of the two experiments overlap, this order was made specifically to minimize the influence of each task on the other and to encourage contraction, which are especially important for the second experiment. Details for each block are discussed separately below. Both experiments were run on a PsychoPy (Peirce et al., 2019) interface, and all speech was recorded on Audacity (Audacity Team, 2008).

Below I will discuss the methodologies and results of the two experiments separately. After that, I will compare the results of the two experiments and discuss the issues and implications.

4.2 Production Experiment I: natural sentence repetition task

4.2.1 Stimuli

4.2.1.1 Test items

The focus of the experiment is on the vowels in the contracted forms and the lexical words. Based on the vowel inventory of Rugao, 8 vowels well distributed along the vowel space are chosen for the experiments, including two low vowels (front and central), four mid vowels (front and back) and two high vowels (front and back): [a, æ, ε, e, ɔ, o, i, u]. All vowels are their surface forms. Test items consist of:

- a. 16 contractible words that are attested contractible in the corpus (see details in previous chapters), including 14 disyllabic words and 2 trisyllabic³⁴ words.
- b. 16 corresponding monosyllabic lexical words with the target vowels [a, æ, ε, e, ɔ, o, i, u] as nucleus and similar neighboring consonants as each of the contracted forms in a.

Contractible words are chosen based on two criteria. First, full syllable contraction is at least possible, i.e., contractible for at least some speakers. Second, contraction for these words is attested according to the corpus and the author's intuition. Two different words are used for each combination of central vowel [ə] and peripheral vowels [a, æ, ε, e, ɔ, o, i, u]. Every contractible part has [ə] in the first syllable and one of the eight target vowels in the second syllable³⁵. This arrangement, based on previous analyses on syllable contraction, is made to ensure consistency and promote syllable contraction. First, the analyses can be focused solely on the 8 target vowels when the deleting vowel is controlled and does not change. This will avoid the difficulty of interpreting data as in Kuo (2010), where deleting vowels were not controlled. Second, participants are most likely to delete [ə] and preserve the peripheral, according to the sonority preference (Chapter 2, Chapter 3). Third, by putting the deleting vowel [ə] consistently on the first syllable, any positional preference and/or linear effect (R.-F. Chung, 1996, 1997) is not a concern.

³⁴ Only the disyllabic part of the trisyllabic words are contractible, so these are not different from the disyllabic words. Trisyllabic words were used only when there is no such disyllabic word in the corpus.

³⁵ For the trisyllabic words, this is the second syllable of the contractible part.

Vowels	Word	Contraction	Meaning
[ə, a]	[səʔ ₁₅ + aI ₂₁]	[saI ₂₁]/ [saI ₃₅₁]	‘twelve’
	[Iən ₁₅ + ka ₂₁]	[Iaa ₃₅]	‘other people’
[ə, æ]	[pjən ₂₁ + ɛjæn ₂₁]	[pjæn ₂₁]	‘fridge’
	[jəʔ ₅ + jæn ₅₅]	[jæn ₅₅]	‘same’
[ə, ɛ]	[pəʔ ₅ + ɲɛ ₂₁]	[pɛɛ ₅₁]	‘no problem’
	[pən ₂₁₃ + lɛ ₁₅]	[pɛɛ ₂₁₃]	‘originally’
[ə, e]	[tɛjən ₂₁ + xej ₂₁]	[tɛje ₂₁]	‘thereafter’
	[ɛjən ₂₁ + nej ₂₁]	[ɛej ₂₁]	‘in mind’
[ə, ɔ]	[ɛjə ₅₅ + jɔŋ ₅₅ + kha ₂₁₃]	[ɛjɔŋ ₅₅ + kha ₂₁₃]	‘credit card’
	[tsəʔ ₅ + jɔ ₅₅]	[tsjɔ ₅₅]	‘only if’
[ə, o]	[tsəʔ ₅ + jow ₂₁₃]	[tsjow ₅₁₂]	‘only’
	[phən ₁₅ + jow ₂₁₃]	[phjow ₁₅]	‘friend’
[ə, i]	[jən ₁₅ + iʔ ₅ + phu ₂₁]	[jiʔ ₅ + phu ₂₁]	‘Business Dep.’
	[tɛən ₂₁ + in ₂₁]	[tɛin ₂₁]	‘experience’
[ə, u]	[pəʔ ₅ + xuj ₂₁]	[puj ₅₁]	‘cannot’
	[tsən ₂₁ + juʔ ₂₁]	[tsjuʔ ₂₁]	‘first Lunar Month’

Table 10. Contractible words, numbers indicating tones

As the control group, the 16₃₆ monosyllabic lexical words each match the contracted forms of the contractible words (Contraction Column, Table 10) to the largest extent, including the same vowel nucleus and neighboring consonants around the vowel. For example, monosyllabic word [saI₂₁] (“spoon”) corresponds to the contracted form of [saI₂₁] (‘twelve’, contracted from [səʔ₃₅ + aI₂₁]). A completely matching (consonants, vowel and tone) word to the contracted syllable is given whenever possible. However, many of the contracted syllables are phonotactically illicit or there is a lexical gap, which means there is no matching lexical word for them. When that is the case, the closest lexical word is given following the preference hierarchy of modified tone > modified onset > modified tone and onset > modified tone and coda. If the

³⁶ In fact, 15 different monosyllabic words were used, because the same word was used for [tsjow₂₁₃] and [phjow₁₅].

contracted syllable with a modified tone is a word, the word with the modified tone is chosen; if modifying the tone does not produce a word, a modified onset is considered. If there is still a lexical gap with the modified onset, a word with both the modified tone and modified onset is used. Modifying the coda will be the last choice. This hierarchy is chosen for three reasons. First, some attested contracted forms simply have irregular, illicit tones. These tones have to be changed in order to have a lexical word. Second, although tone can change the length of the syllable, it is less likely to change the vowel qualities compared to onset and coda. This could avoid vowel quality changes due to their adjacent segments. Third, as coda is part of the rime, coda is not modified unless tone or onset changes still produce an irregular syllable. With all the changes that have to be made, even when the onset or coda must be modified, the articulatory features, including place of articulation, manner and voicing are kept as much as possible. For example, for [ɕej₂₁], [sej₂₁] only the modified Place of Articulation is used. A full list of monosyllabic words is in Table 10.

Contraction	Exact Matching	Modified tone	Modified onset	Modified tone & onset/coda	Meaning
[saɪ21]	[saɪ21]				‘spoon’
[ɭaa35]		[ɭa213]			‘annoy’
[pɿæn21]			[ɕjæn21]		‘aromatic’
[jæn55]		[jæn15]			‘Yang (name)’
[pɛɛ51]		[pɛ55]			‘greet’
[pɛɛ213]	[pɛ213]				‘put’
[tɛej21]			[xɛj21]		‘thick’
[ɕej21]			[sej21]		‘search’
[ɕjɔŋ55]		[ɕjɔŋ21]*			‘competent’
[tsjɔ55]	[phjɔ55]				‘ticket’
[tsjow512]		[tɕhjow213]*			‘prank’
[phjow15]		[tɕhjow213]*			‘prank’
[jiʔ5]	[jiʔ5]				‘leaf’
[tɛin21]	[tɛin21]				‘pointy’
[puj51]				[thuj55]	‘return’
[tsjuʔ21]				[tɕhjuʔ5]	‘short of’

Table 11. Word list—monosyllabic lexical words. [Numbers indicate tones. Tokens with an

asterisk are dialect words that do not have standard written forms.]

4.2.1.2 Stimuli

Wong’s (2006) dissertation and Kuo’s (2010) series of experiments laid out a basic guideline for designing such studies in Chinese languages, both of which used the speeded sentence repetition technique to elicit contracted syllables. In this experiment, all test items are embedded in carrier sentences. The purpose of the experiment is to elicit as many contracted tokens as possible, but it may still be difficult for the participants given the experimental setting. One way to induce contraction is to encourage casual speech by providing chat-like sentences of everyday communication. 48 short sentences are used as carrier sentences, all of which sound like part of an everyday conversation/chat between two women. Topics include weather, shopping, banking, and gossiping. Unlike in Kuo (2010), no consistent carrier phrase/sentence

was used in this experiment. There are several reasons for not using the same carrier phrases. First, the use of a consistent carrier phrase will certainly reduce the naturalness of the sentences, as the comparing word pairs can be of different syntactic categories, making it hard to put them in the same sentence frame while keeping the sentence natural. Second, if the target words always appear in the same position of the sentence, e.g., right after a certain word, the elicited tokens are not only less natural, but also unpredictable. Undesirable emphasis might be put on the word when the participants figure out the position of the target word in the sentence. In that case, the duration measurement as well as some formant measurement might be untrustworthy. For these reasons, all carrier sentences are different. In natural speech, sentences vary in how many contractible words they may have, and the speakers can choose to contract certain words but not others. In order to mimic natural speech as much as possible, the number of contractible words in each sentence varies from 1- 4, including the contractible test items listed in Table 10, and filler contractible words that are not targets for analysis. Each test item (contractible word or monosyllabic word) is embedded in sentence-initial, sentence-medial and sentence-final positions of multiple carrier sentences, creating three tokens for each word for each round of sentences. All these manipulations were made so that participants could navigate into the casual speech register more easily without paying attention to any specific word.

Below are some examples. Where there are target words, the contractible words are bolded, and the monosyllabic lexical words are underlined. Note that not all contractible words are target words; only target words are marked. The first line is the standard *pinyin* annotation. The second line is the actual pronunciation in IPA. See Appendix C, Appendix D and Appendix E for full lists of sentences in the practice, *burn-in*, and test phrases, respectively. Note that conventional *pinyin* annotation is used for ease of transcription, but the actual pronunciation can

be distinct from the *pinyin* annotation. Rugao also has slightly different syntactic and vocabulary systems from Standard Mandarin. Standard *pinyin* annotation is used here to represent the rough syntax, but not necessarily the actual word choice or pronunciation.

a. *Mashang jiushi nongli zhengyue shi'er.*

masæn tɕhjowsi nɔŋljəʔ tsəŋjuʔ səʔaɪ

soon be lunar January twelve

‘It will be Lunar January 12 soon.’

b. *Zhengyue li chichi shuashuar benlai manhao de.*

tsəŋjuʔ li tɕhəʔtɕhəʔ ʃwaʃwaɪ pənle mənɬəʔ teɟ

January in eat-eat play-play pretty-good PCL³⁷

‘It is nice to eat and have fun in the Lunar January.’

c. *Dongtian nuanhe de shihou you shi dao shi'er du.*

tɔŋthin nɯnhow de siheɟ jow səʔ tɔ səʔ.aɪ thu

winter warm DE time has ten to twelve degree

‘It can be 10-15 degrees when it is warm in winter.’

d. *Hou mian'ao dou zai dazhe ni dao wangshang sousou*

xɛɟ tentsi dow tshɛ tatɕiʔ nei tɔ vænsæn seisei

thick coat all -ing on-sale you go online search-search

‘Thick coats are on sale; go search online.’

e. *Zhege mian'ao zhengyue chuan benlai xian hou*

Tsakow tentsi tsənjuʔ tshun phatei ɛin xɛɟ

³⁷ Sentence-ending particle, similar to *de* in Standard Mandarin.

this coat January wear probably too thick

‘This coat may be too thick for Lunar January.’

One disadvantage of the sentence repetition methodology is the shadowing effect (Marslen-Wilson William, 1973), which means certain aspects of the audio stimuli may prime the listener and participant to repeat what they hear and thus confound the results. To minimize the influence of any certain speaker, two different speakers were recorded producing the sentences. Both speakers are female, native speakers who grew up in downtown Rucheng City and aged 30 and 33 respectively at the time of recording. The voice qualities and speech rates of the two speakers were distinct and easily identifiable by the listener. Note that I also collected data that does not involve audio stimuli, i.e., all sentences were presented orthographically, but I leave this part for future study.

Besides the casual speech nature of the sentences themselves, another way to encourage contraction is to expose the participants to speech that contains as many instances of contraction as possible. To facilitate casual speech and syllable contraction during the stimuli recording process, all sentences were practiced several times before the recording. During the recording, the stimuli speakers were instructed to read out the list of sentences “fast”, “in a lively tone” and “as if you are chatting to a friend”. For each speaker, two recordings were made for each sentence by repeating the list of sentences twice; the one with more contraction items was chosen as the stimuli. Another native speaker checked all the recordings for naturalness. The recordings of the two speakers were fully randomized during the experiment.

4.2.2 Participants and procedure

Only female participants were recruited, as women are more likely to be contraction users; even if both genders contract, women may contract to a larger degree compared to men (Xu & Mao, 2017). Besides, the previous study on Taiwan Mandarin (Kuo, 2010) only found differences in vowel space for male speakers, while female speakers pronounced all three vowels the same. It will be of interest to find, first, whether female speakers of Rugao also show no difference in vowel space in this process when the stimuli are well controlled. It will be informative to investigate the gender factor, but as the first experiment of this kind in the context of Rugao, I start with female speakers and leave the male speakers for future projects.

All participants were recruited via posts and re-posts using an app called WeChat. Some participants had to be excluded due to recording issues such as excessive background noise and headphone issues. Some participants were excluded because they were simply not contraction users or entirely misunderstood the task.

A total of 20 participants were analyzed. For these participants, mean age = 31.75 years, range = 24—36 years. All participants grew up in and around downtown Rucheng City. They also reported to currently live in this area and speak the Rugao dialect exclusively at home and at work or only at home. It is worthwhile to note here that inter-speaker variation was observed. Even for the 20 speakers who are contraction users at least subconsciously, the words that are contracted vary substantially across speakers, as does the degree of contraction. Although the variation question is interesting to explore, it is not a focus of this study. As the first phonetic study of Rugao syllable contraction, the measurement methodology as well as the reported results focus on the general trend and ignores such inter-speaker variation for the conciseness and simplicity of the analysis. I leave the variation questions for future projects.

The experiment took place in a relatively quiet office in downtown Rugao. The location is not noise free, however. Traffic passing by, car horns honking, people talking, and an automated answering machine were major sources of noise. Noise-cancelling Bose headphones were used for playing all stimuli. The built-in microphone of a Macbook Air was used for all the recordings to avoid any fine microphones catching too much noise from the environment. Participants were instructed to sit about 30 centimeters away from the computer for clearer recording.

No information for syllable contraction was given throughout this experiment. There were three basic sessions in this block: practice, *burn-in*, and the real task. Three sentences were used for practice (full list in [Appendix C](#)), and another two/three were used for *burn-in* (see [Appendix D](#)). The use of *burn-in* sentences was to ensure a smooth transition to the task while avoiding the first few sentences not being useful (participant being nervous, too many errors, too formal). The sentences in practice and *burn-in* sessions were similar to the test sentences, but none of them were used for the real task or analyses. Although different sentences were presented in the three sessions, the same instruction was given throughout the experiment: participants were asked to listen to the sentence and “say the sentence you just heard again” “like you are talking to a close friend” and “as fast as you can”. To increase speech rate, there were only 5 seconds for them to record each sentence. If they wanted, there was one chance to repeat the sentence. All the 96 (48*2) sentences produced by the two speakers were blended as one list and fully randomized. Instead of repeating each sentence twice in a row, the list of sentences was repeated twice.

4.2.3 Measurement

All measurements were made using Praat (Boersma & Weenink, 2018). The length information, including vowel duration and word duration were measured for each relevant vowel and word, and the vowel/word ratio was calculated by dividing the vowel duration with the word duration. The formant information, including the F1 and F2 values, were measured at the midpoint of the vowel duration.

In order to simplify the analyses and stay focused on the main goal, the non-contracted words and partially contracted words were excluded from the measurement for this experiment. All monosyllabic lexical words were measured, which was to gather information about all the lexical vowels, but only the fully contracted contractible tokens were measured. This means if a certain disyllabic word was pronounced uncontracted or only partially contracted, this token was left out. Particularly, a token must pass three criteria to be considered ‘fully contracted’. First, there is no audible [ə] in the audio file to the ear of the author. Second, there is no visible [ə] on the Praat spectrogram. Third, according to Myers & Li (2009), the degree of contraction can be measured in *Trough-depth*, i.e., the “dip” in intensity where the two syllables are expected to break. A *Trough-depth* that is equal or close to zero means the two syllables are like one single syllable. On the spectrogram, this means the intensity line is a more or less smooth curve, or at least does not have a visible dip where the boundary between the two syllables is expected. The author used this as a baseline and visually examined each token. An example of a fully contracted syllable is illustrated in Figure 15 below. As can be seen from the Praat screenshot of the word [tsjuʔ] (contraction of [tsəŋ.juʔ], ‘lunar January’), there is no visible [ə], and there is no dip in the intensity (the yellow line) throughout the word.

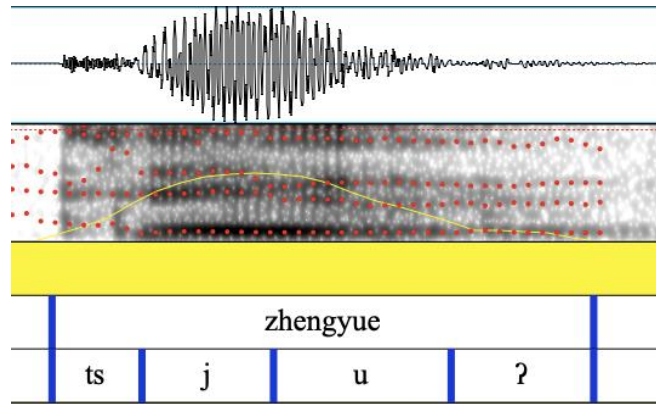


Figure 15. Example of fully contracted token, [tsju?], contracted from [tsən.ju?]

For less perfect cases where the intensity line is not as smooth as in Figure 15, i.e., showing a slight dip or plateau or trough-depth not equal to 0, precautions were made by the author to ensure that [ə] is indeed deleted using the three criteria that combine audial and visual examination, as described in the previous paragraph. As expected, not every participant contracted every word. Even though all the non-contracted users were excluded, as discussed before, the contraction users exhibited great variation in what words they would contract and to what degree. As the comparison is made within participants, the mean values for each vowel is calculated for the contracted and the lexical words respectively. If a participant did not contract a certain word, there are no comparable values, thus this specific comparison is excluded, decreasing the total number. Furthermore, since only fully contracted tokens were analyzed here, the number of usable tokens further decreases. Despite some loss of data, a good amount of data was still obtained for each vowel. A total of 1070 contractible words were obtained, together with 1541 monosyllabic words in the control group.

4.2.4 Results

4.2.4.1 Length and vowel ratio

One big problem for absolute length data has always been that the absolute values may not be comparable unless the same speech rate is assumed. In the sentence repetition task, all the contracted and lexical words were produced in carrier phrases. The duration values were not altered because, presumably, participants used the near-same speech rates to produce all sentences, given the time limit for producing each sentence and the help of practice sentences and *burn-in* sentences. The former prompted speakers to use a fast speech rate all the time, and the latter greatly decreased the hesitation that slows down the speech. The author also spot-checked some participants and found no observable speech rate variation.

As can be seen in Figure 16, the contracted words are generally longer than the monosyllabic words. This difference is likely due to the underlying syllable count, as the contracted words were underlyingly disyllabic, but it is also possible that the seemingly contradicting data³⁸ could simply be the result of random variation. Note that in Figure 16, the contractible word and the corresponding lexical word are placed right next to each other, with the contracted on the left and the lexical on the right.

³⁸ No observable difference is found for [tsjo] (“zhiyou” in the figure) vs. [tɕhjo] (“qiu”). An opposite trend was found for [tsjo] (“zhiao”) vs. [phjo] (“piao”), [buj] (“buhui”) vs. [tɕuj] (“tui”) and [tsju?] (“zhengyue”) vs. [tɕju?] (“que”). What these pairs have in common is that the lexical word has an aspirated consonant in the Onset, while the contracted form has unaspirated consonant there. These pairs had to be used because of the lexical gaps, but most likely caused this set of data to not pattern with the rest.

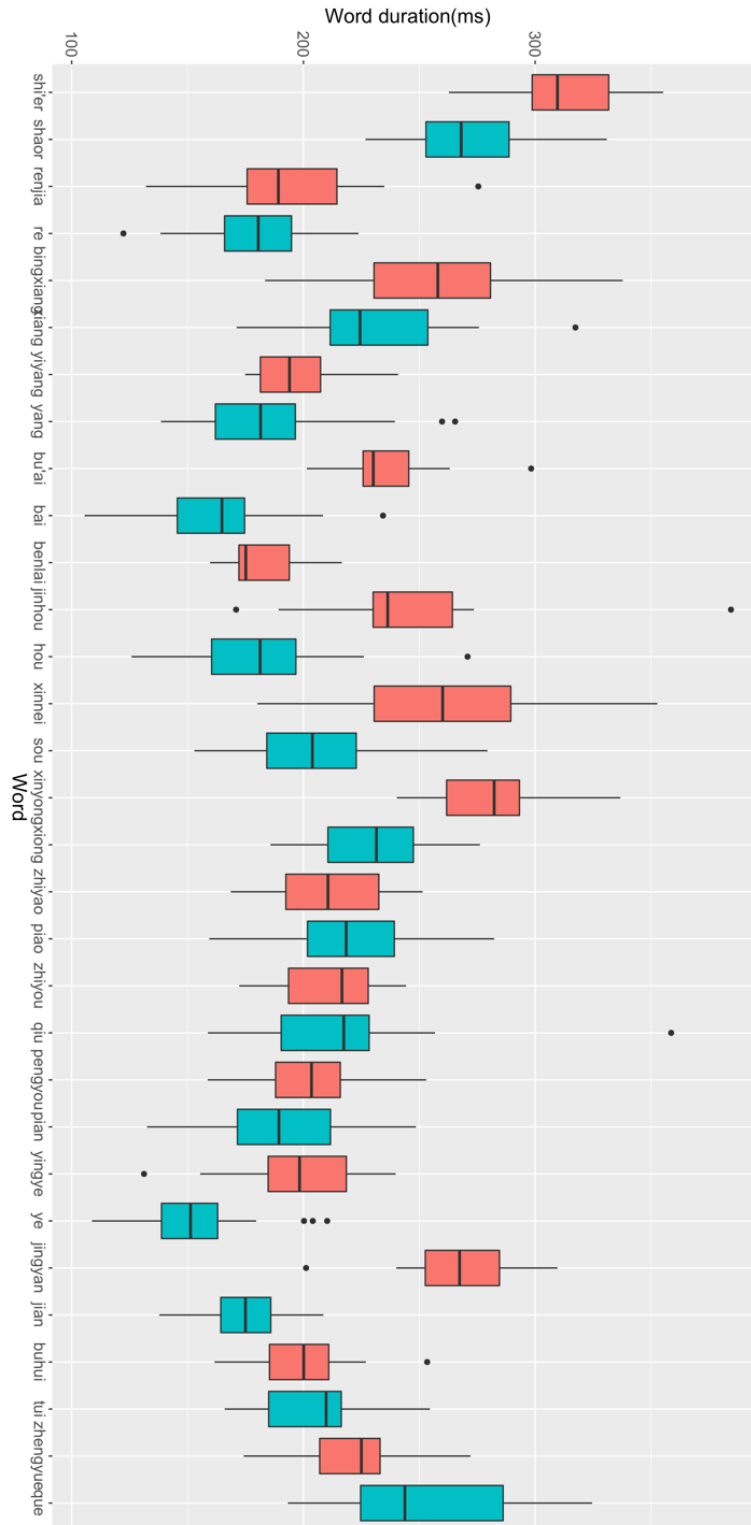


Figure 16. Word duration—Experiment I. [Red boxes present contracted tokens. Green boxes represent lexical tokens.]

Based on the fact that the contracted words are generally longer than the lexical word, it is not surprising to find that the vowels are correspondingly longer in the contracted words, as shown in Figure 17. The contracted vowel being longer seems to be a consistent pattern for all the vowels in question here. However, paired Welsh tests were conducted comparing the vowel length in the contracted syllable and lexical syllables, and the results show that only half of the vowels have a significantly longer vowel in the contracted syllables, summarized as the following. Note that the number of the df (degree of freedom) varies because different numbers of participants were analyzed for each vowel. If a participant did not contract a certain word, there were no comparable values, thus this participant is excluded.

Vowel longer in contracted syllable	No significant difference
[a]: $t(17) = 5.17, p < .001$ ³⁹	[æ]: $t(17) = 2.1, p = .05$ ⁴⁰
[e]: $t(13) = 4.9, p < .001$	[ɛ]: $t(14) = 0.6, p = .46$
[i]: $t(17) = 10.1, p < .0001$	[u]: $t(17) = 1.41, p = .18$
[o]: $t(17) = 3.1, p < .01$	[ɔ]: $t(17) = 1.74, p = .10$

³⁹ P values for this chapters are not adjusted.

⁴⁰ This is marginally significant.

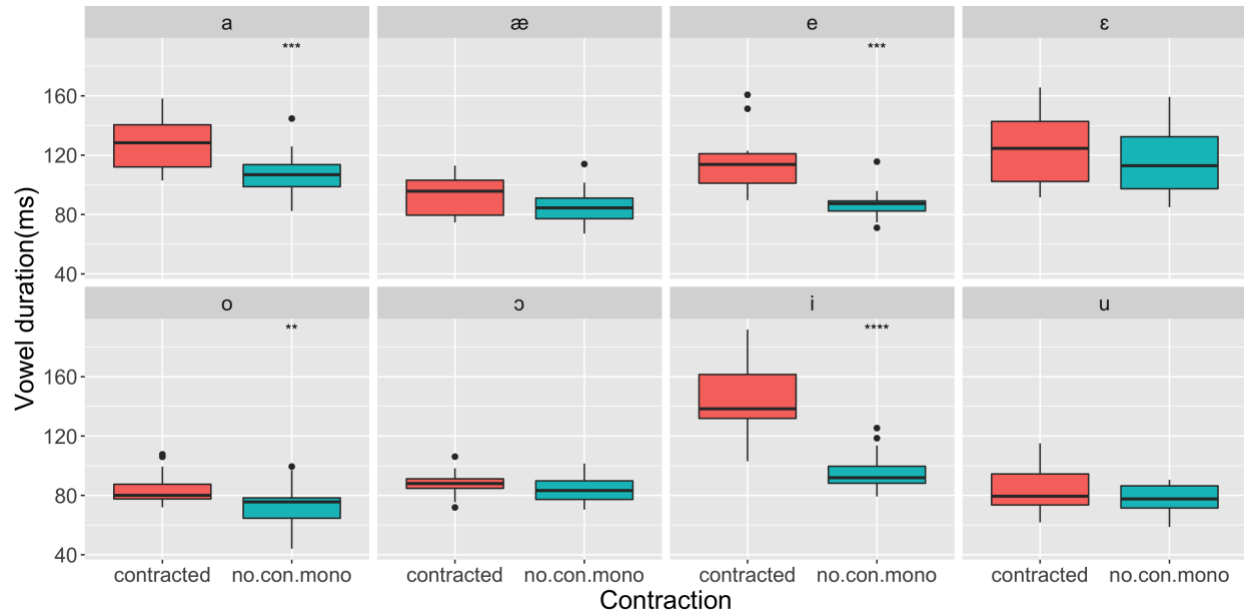


Figure 17. Vowel duration—Experiment I. [Asterisks indicate significance levels.]

If the contracted syllables have the same syllable structure, i.e., the same proportion of the vowel nucleus, the contracted vowels should be longer than the lexical vowels, based on the fact that the contracted words are generally longer than the lexical word. However, such duration difference is not consistently observed. It is likely that the status of the vowel nucleus in the contracted syllables is different in the way that the vowels take up a different proportion in the syllable, similar to what was reported in Taiwan Mandarin (Kuo, 2010). The vowel ratios are calculated by dividing the vowel duration with the word duration. As only fully contracted words and monosyllabic words are included in this analysis, the Vowel/Word ratio equals the Vowel/Syllable ratio. Figure 18 reveals that, in general, the proportional difference of vowels in the syllable is not consistently observed in Rugao contracted syllables. The mean vowel proportion of a fully contracted word is greater for [o, i, u], but no difference is found for [æ, e, ε, ɔ]. Paired Welsh tests were conducted comparing the Vowel/Syllable ratio of these vowels in the contracted syllable and lexical syllables, and the results are the following:

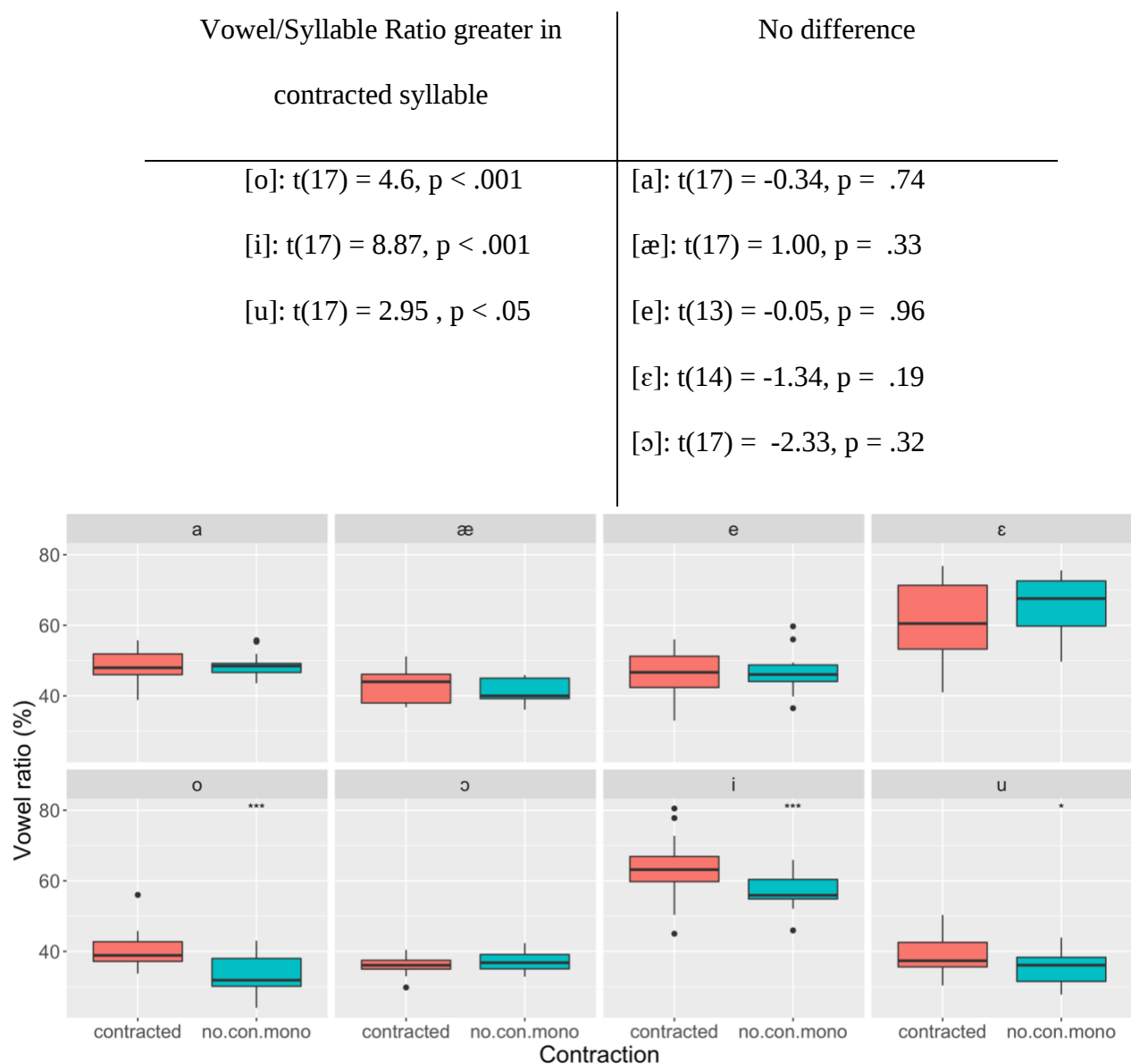


Figure 18. Vowel ratio—Experiment I. [Asterisks indicate significance levels.]

For this complicated result, one important note is that words that contain [i, o, u] may have caused the difference in Vowel/Syllable ratio:

For [i], a tri-syllabic word [jən.iʔ.pu] was used. Although only the disyllabic part is contractible, the last syllable might have caused the [jən.iʔ] to lose the glottal stop, a process that

is common in Rugao. The lexical word [iʔ], however, is less likely to lose the coda. This likely makes the vowel ratio larger in [iʔ]/[i], the contracted form, than in [iʔ], the lexical form.

For [o], the contraction form [tsjo] of the word [tsəʔ.jo] is compared with [tɕhjo] because of the lexical gap *[tsjo]. The aspiration of the lexical word [tɕhjo] takes up a bigger proportion than [ts] would have, making the vowel portion comparatively smaller in [tɕhjo].

Similarly, for [u], [puj] (contraction of [pəʔ.xuj]) and the lexical word [tɰuj] have aspiration contrast as well as Place of Articulation contrast in the first consonant, i.e., [p] vs. [tɰ].

It is likely that the observed vowel proportion difference for some vowels are exceptions that are results of the stimuli construction instead of the contrast between the contracted and the lexical words. However, there needs to be more evidence to confirm this claim.

4.2.4.2 Vowel quality and vowel space

Both F1 and F2 are different in vowels in the fully contracted words, compared to the corresponding monosyllabic lexical words. In the following sub-sections, I report the detailed data of F1 and F2 measurements at the midpoint.

4.2.4.2.1 F1

As mentioned before, if the [ə] deletion is incomplete, the central vowel should “drag” the peripheral vowels to the center of the vowel space, which means that at least the non-mid vowels may become lower or higher, with the F1 values being higher or lower than that of the lexical vowels, respectively. However, one cannot jump to any conclusion before each vowel is examined separately, as a different picture arises, shown in Figure 19. For vowels [a, æ, ɛ, o], the F1 in the contracted word is either lower or higher than that in the monosyllabic lexical word, while [e, ɔ, i,

u] do not change in F1. Within-subject paired Welsh tests were run, and the results are summarized as follows:

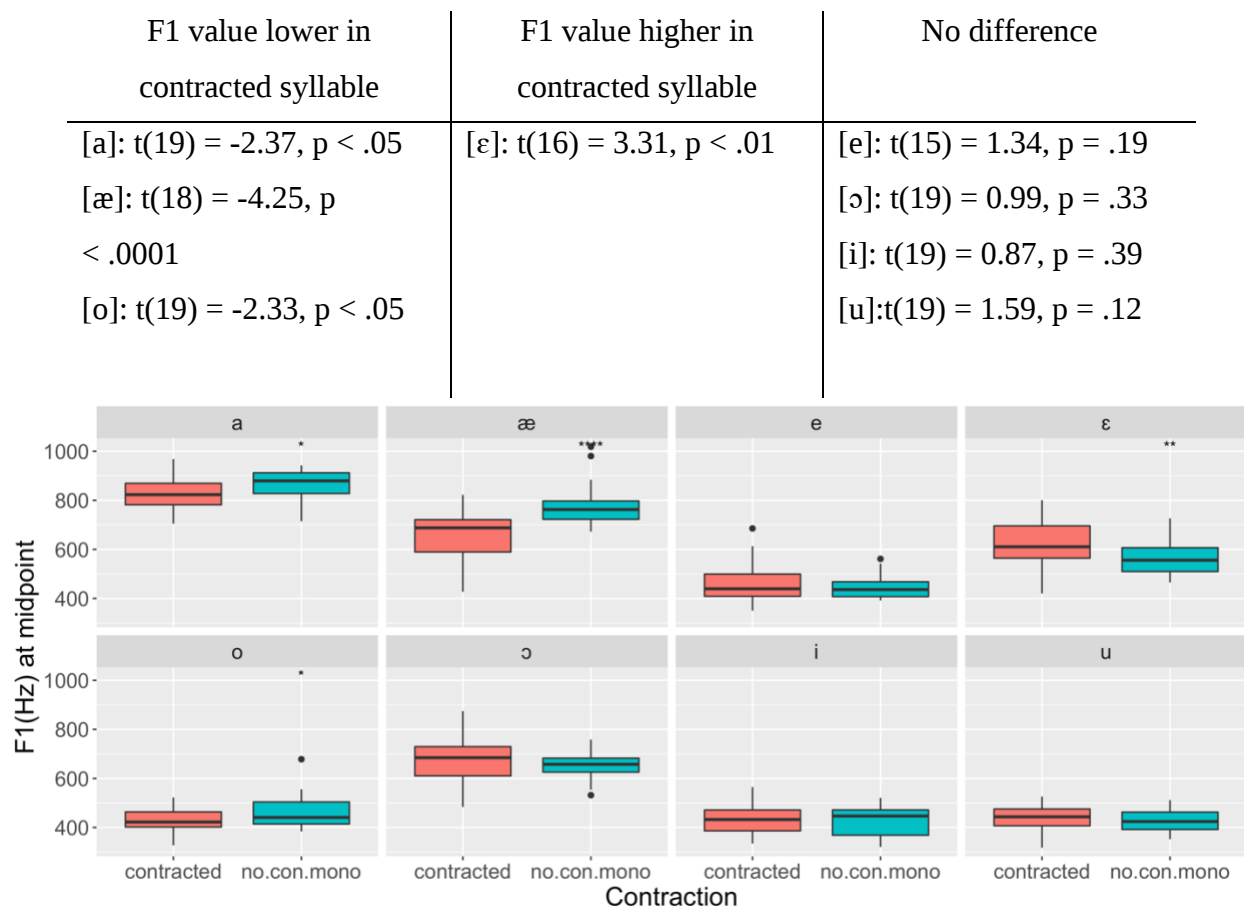


Figure 19. F1 (Hz) at midpoint—Experiment I. [Asterisks indicate significant levels.]

This means the height of the vowels can change in the contracted forms. The low vowels [a, æ] are higher⁴¹ in the vowel space in the contracted forms than in the lexical forms. The mid vowels [ε, o] point in two different directions: in the contracted word, while [ε] is lower, [o] is higher. No change was observed for the other two mid vowels [e, ɔ] and high vowels [i, u].

⁴¹ Terms “higher/high” and “lower/low” refer to the height of the vowel, unless otherwise specified, e.g., “higher in F1”.

Crucially, for the four mid vowels, there is no consistent pattern with regards to vowel height, with either no change or changes in different directions. The height changes observed for [ɛ, ɔ] should be either exceptions or the artifact of experimental design. Mid and high vowels in general do not change height in the contraction process.

4.2.4.2.2 F2

Once again, as the deleted vowel [ə] is mid-central, it should “pull” the peripheral vowels to the center. With regards to F2 values, there should be a change, either higher or lower, depending on the backness of the vowel. As shown in Figure 20, the F2 difference is mostly observed for mid vowels, but not much for low or high vowels. For [ɛ, u], the F2 value is lower in the contracted words than in the monosyllabic words. For [ɔ, ɔ], the F2 in the contracted word is higher than that in the monosyllabic lexical word. No statistically significant difference was observed for [a, æ, e, i]. Within-subject paired Welsh tests were run, and the results are summarized as follows:

F2 value lower in contracted syllable	F2 value higher in contracted syllable	No difference
[ɛ]: t(16) = -5.63, p < .001	[ɔ]: t(19) = 7.01, p < .0001	[a]: t(16) = 0.26, p = .79
[u]: t(19) = -2.65, p < .05	[ɔ]: t(19) = 2.3, p < .05	[e]: t(12) = -0.68, p = .51
	[æ]: t(16) = 2.16, p < .05 ⁴²	[i]: t(16) = 1.234, p = .23

⁴² The exact p-value is 0.04587.

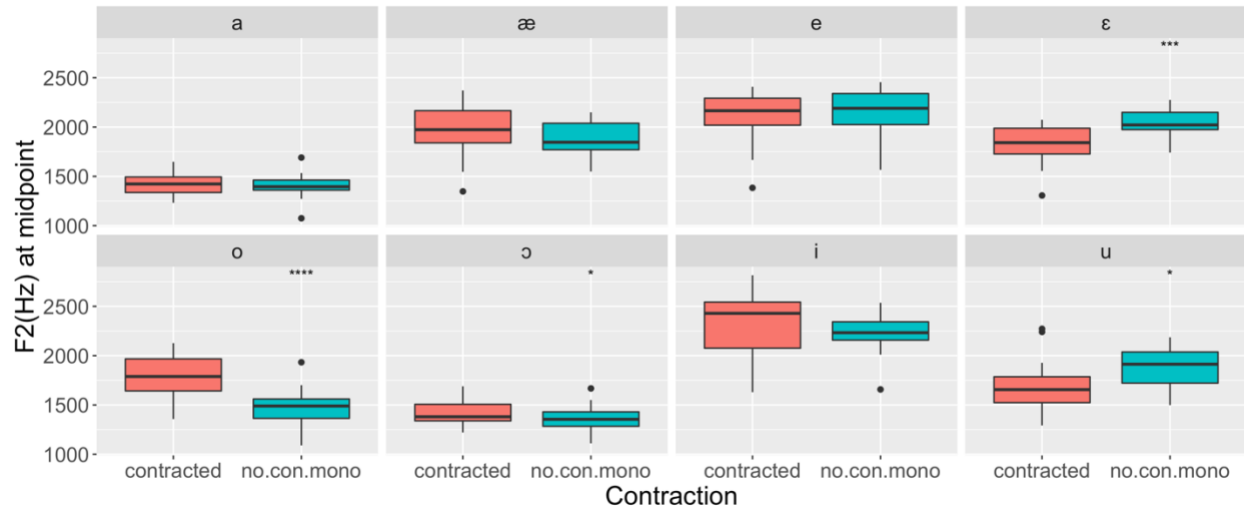


Figure 20. F2(Hz) at midpoint—Experiment I. [Asterisks indicate significant levels.]

4.2.5 Interim summary

In summary, the results of Experiment I show that the vowels in the contracted syllables are similar enough to the lexical vowels, but subtle differences exist. The contracted vowels are generally longer, but the vowel/syllable ratio, as well as the vowel/word ratio, are similar to that of a single lexical syllable. However, the vowel quality may be different. Generally, vowels in the fully contracted words are centralized, but to a limited degree, as most vowels are different in either height or backness, but not both.

4.3 Experiment II: Self-paced word contraction task

As mentioned before, the second experiment was designed to elicit comparable contractible and non-contracted monosyllabic and additional disyllabic lexical words. Importantly, this additional experiment collects data without immediate influence from audio

stimuli. The test items were similar to Experiment I. So, in the following paragraphs, I will simplify the descriptions of the methodologies and be focused on the differences and the results.

4.3.1 Stimuli

Test items for this experiment are the same di-syllabic words and monosyllabic words as used in Experiment I (as in Table 10, Table 11). Differences lie in how these items were presented. First, all these items were presented as single words, and no carrier phrase was used in Experiment II. Second, all stimuli were shown in *hanzi*, the standard Chinese orthography. For words without standard written forms or the written form does not demonstrate the accurate pronunciation (words with an asterisk in Table 11), a short sentence that explains the word was given. Unlike in Experiment I, where filler sentences and *burn-in* sentences were used, no filler item was added in Experiment II. If the participants understood the task correctly, the nature of the contraction task had already made the experiment goal, although not necessarily the vowels, explicit to them. It is not worth the added time to the experiment to add in fillers.

4.3.2 Participants and procedure

As mentioned before, participants who did not understand the tasks or those who are simply not contraction users (i.e. those who produced all words in their non-contracted forms for either reason) are excluded. If the participant produced contraction in Experiment I, but not in Experiment II because of style shift or other sociolinguistic factors, this participant was not included in the analyses below. 26 participants were analyzed for Experiment II. For these participants, mean age = 31 years, range = 26—36 years. All participants reported to have grown up and lived in downtown Rugao for most of their life except the three to four years in college.

They also speak the Rugao dialect exclusively at work and at home or use a mixture of Mandarin and Rugao.

This experiment consists of two tasks: a self-paced word contraction task and a wordlist reading task. At the very beginning of this experiment, the participants listened to an introduction to the concept of syllable contraction and the task of this block. This introduction session serves two purposes. First, compared to a written instruction, an audio with real word examples is supposedly a better way to introduce the concepts as well as the task because participants are able to hear real word contraction examples and have to go through the entire instruction without being able to skip. Second, this one-minute long audio creates a break that is far enough from the previous experiment, so that the participants were less likely to still remember details of any certain words that they were exposed to before.

The author speaking in Rugao recorded the introduction. Syllable contraction was explained as “a two-syllable word spoken in one syllable, but still the same word. And the meaning will not change.” The explanation was accompanied by an infamous Mandarin example, *jiangzi* ([tɕjaŋ⁵¹.tsi⁰], contraction of [tɕɤ⁵¹.jaŋ⁵¹.tsi⁰]), and some frequently used Rugao contraction words embedded in a short everyday conversation. The author said the sentences with all the contractible words contracted, and then none of them contracted, to highlight the difference, as shown below. Comparable contractible words are underlined. A full transcript of the introduction can be found in [Appendix E](#).

Contracted: *nej sen sej lɛ a? ɲo mæn lɛ.*

Non-contracted: *nej sen.dej seihej lɛ a? ɲo ma.sæn lɛ.*

you what time come PAR I right-away come

‘When will you come? I will come right away.’

The participants had the time to ask questions about the concept of syllable contraction and the task before they proceeded. In the self-paced word contraction task, participants saw the randomly presented wordlist, one word each time on the screen in standard *hanzi*. They were instructed to say the contraction form of each word. There was no time limit for each word, and participants had the option to record multiple tokens until they were satisfied with the recording.

4.3.3 Measurements

The same measure as in the previous experiment was made in Praat (Boersma & Weenink, 2018). All target tokens were measured in vowel duration, word duration, ratio of vowel in the word/syllable, and F1 values and F2 values at the midpoint of the vowel. The same precaution was made to discriminate among the non-contracted, partially contracted, and fully contracted tokens, as used for Experiment I. For the word contraction task, only the fully contracted tokens are measured. If there were repeated tokens for the same word, only one fully contracted token was chosen for each word. This experiment provides an added set of data by allowing the measurements for the non-contracted disyllabic words, which was not feasible in Experiment I. All tokens in the wordlist reading task were measured, including the monosyllabic words and the disyllabic words that are not contracted. In some cases, speakers contracted some words even in the wordlist reading task. Such contracted tokens were excluded. A total of 226 contracted tokens, 606 non-contracted disyllabic tokens, and 641 monosyllabic tokens were measured and compared.

Once again, not every participant contracted every word. In this experiment, fewer contracted tokens were elicited than in the previous experiment, and even fewer for the fully contracted tokens. However, even with far fewer data points, the results of Experiment II are

quite similar to that of Experiment I, although differences arise in the many details. The data presented below thus provides another perspective into this phenomenon. Below I will show the results in the same format, although some details will be omitted if already discussed in the previous section.

4.3.4 Results

4.3.5 Vowel/Word ratio

As mentioned before, the absolute duration values depend heavily on the speech rate. Although this is not a concern for Experiment I, the design of Experiment II decided that any comparison of absolute length is problematic here. Words that are elicited with a contraction task are by nature shorter than words that are elicited with a wordlist reading task. The vowel duration shouldn't be directly compared due to such big differences in the word duration. For these reasons, I will not discuss these two sets of data here and will focus on reporting the vowel/syllable ratio data in this section, shown below in Figure 21.

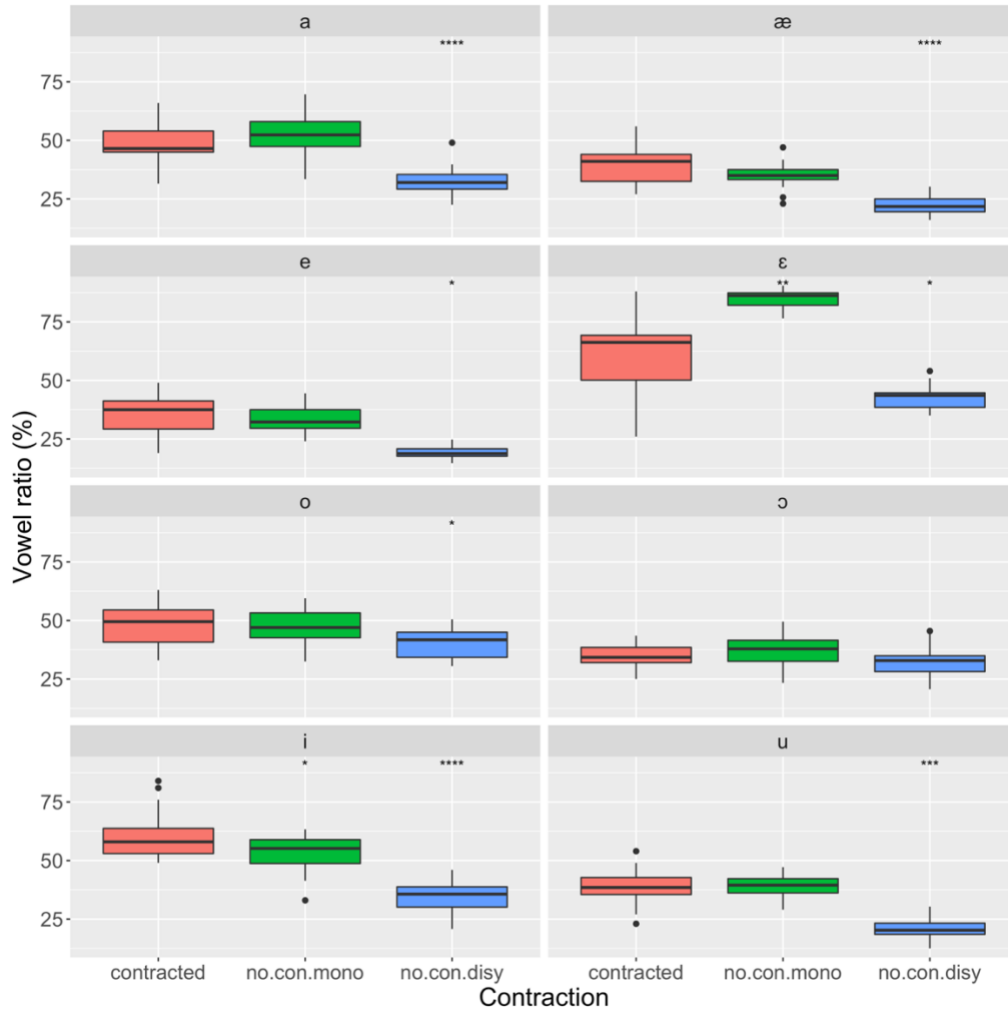


Figure 21. Vowel/Word ratio—Experiment II. [No.con.mono = non-contracted monosyllabic, no.con.disy = non-contracted disyllabic. Asterisks indicate significant levels.]

4.3.5.1.1 Contracted vs. monosyllabic lexical

Generally, the vowel proportion of a fully contracted word is not different from that in a monosyllabic lexical word (red box vs. green box, Figure 21). This pattern holds true for [a, æ, e, o, ɔ, u], but [ε] and [i] are two exceptions. Paired Welsh tests were conducted and the results are summarized below:

Vowel proportion larger in contracted syllable	Vowel portion smaller in contracted syllable	No difference
[ɛ]: $t(9) = -3.85, p < .01$	[i]: $t(22) = 2.39, p < .05$	[a]: $t(22_{43}) = -1.37, p = .18$ [æ]: $t(16) = 1.89, p = .08_{44}$ [e]: $t(7_{45}) = .38, p = .78$ [o]: $t(9) = .23, p = .82$ [ɔ]: $t(21) = -1.64, p = .11$ [u]: $t(18) = -.43, p = .67$

4.3.5.1.2 Contracted vs. disyllabic lexical

The comparison between the contracted vs. non-contracted disyllabic tokens are new data that is only possible based on the data of Experiment II. Recall that the vowel ratio is calculated by dividing the vowel duration by the word length. In this experiment, the Vowel/Word ratio does not equal the Vowel/Syllable ratio. As seen in Figure 21 (red box vs. blue box), the vowel ratios in the non-contracted disyllabic words were smaller than those in either contracted or monosyllabic lexical words for all the words except for those that contain [ɔ]. And this dramatic difference is most likely due to the syllable count difference (2 vs. 1).

Recall that the comparing contracted and non-contracted tokens are the same word underlyingly, for example, [tsjuʔ] ('first Lunar month') vs. [tsən.ju] ('first Lunar month'). Results presented above indicate that the fully contracted disyllabic word is indeed no longer a disyllabic

⁴³ Note, again, that the *df* values vary because participants that did not contract the words for a certain vowel were excluded.

⁴⁴ This is marginally significant.

⁴⁵ Only a few participants contracted words with [e], [ɛ] and [o], making their *df*'s relatively small. Word frequency may be a cause for the low contraction incidences. This is expected, however, considering the forced contraction task is by nature a more difficult environment for syllable contraction.

word, but patterns with some other type of word, the corresponding monosyllabic word, just the same as seen in Experiment I (Figure 18). This added set of data here completes the picture for the vowel proportion of the contracted words. Contrary to the belief in the former literature that the contracted syllable is somewhere between its corresponding di-syllabic form and a monosyllabic word, the contracted word/syllable actually has the syllable structure of a monosyllabic word.

4.3.6 Vowel quality and vowel space

For the formant values, the same general pattern was found for the comparison between the contracted and the monosyllabic tokens: both F1 and F2 change in vowels in the fully contracted syllables, compared to the corresponding monosyllabic and disyllabic lexical words.

Below I will report the detailed data of F1 and F2 measurements of Experiment II one by one. Once again, as Experiment II provides the non-contracted disyllabic data, I will report the monosyllabic data first and then the disyllabic data.

4.3.6.1.1 *F1*

- Contracted vs. monosyllabic lexical

Generally, the F1 values in the fully contracted syllables are different from that in the monosyllabic lexical words for non-mid vowels, which is similar to Experiment I. As shown in Figure 22 (red bar vs. green bar), for vowels [a, æ, u], the F1 in the contracted word is lower than that in the monosyllabic lexical word, suggesting raising. For all mid vowels [e, ε, o, ɔ] and high vowel [i], no difference was observed. Paired Welsh tests were run, and the results are:

F1 lower in contracted syllable	No difference
[a]: $t(22) = -4.66, p < .001$	[e]: $t(7) = -1.0, p = .35$
[æ]: $t(16) = -2.15, p < .05$	[ɛ]: $t(9) = 1.88, p = .09$
[u]: $t(18) = -3.54, p < .01$	[o]: $t(9) = -1.77, p = .11$
	[ɔ]: $t(21) = -0.12, p = .09$
	[i]: $t(22) = -0.48, p = .64$

- Contracted vs. disyllabic lexical

As shown in the red bar vs. blue bar in Figure 22, three vowels, [a, o, u], show a lower F1 in contracted syllables than in disyllabic lexical syllables. For the rest of the vowels [e, ɛ, ɔ, i, æ], no difference was observed. Paired Welsh tests were run and the results are:

F1 lower in contracted syllable	No difference
[a]: $t(22) = -4.01, p < .001$	[æ]: $t(16) = -1.62, p = .12$
[o]: $t(9) = -2.51, p < .05$	[e]: $t(7) = -1.24, p = .25$
[u]: $t(18) = -3.51, p < .01$	[ɛ]: $t(9) = -0.74, p = .48$
	[ɔ]: $t(21) = 1.5, p = .09$
	[i]: $t(22) = -0.47, p = .64$

This suggests that [a, u, o] are higher in the contracted form than in their corresponding non-contracted disyllabic forms, but there is no difference between the contracted and the non-contracted disyllabic lexical words with regards to the height for [æ, e, ɛ, ɔ, i].

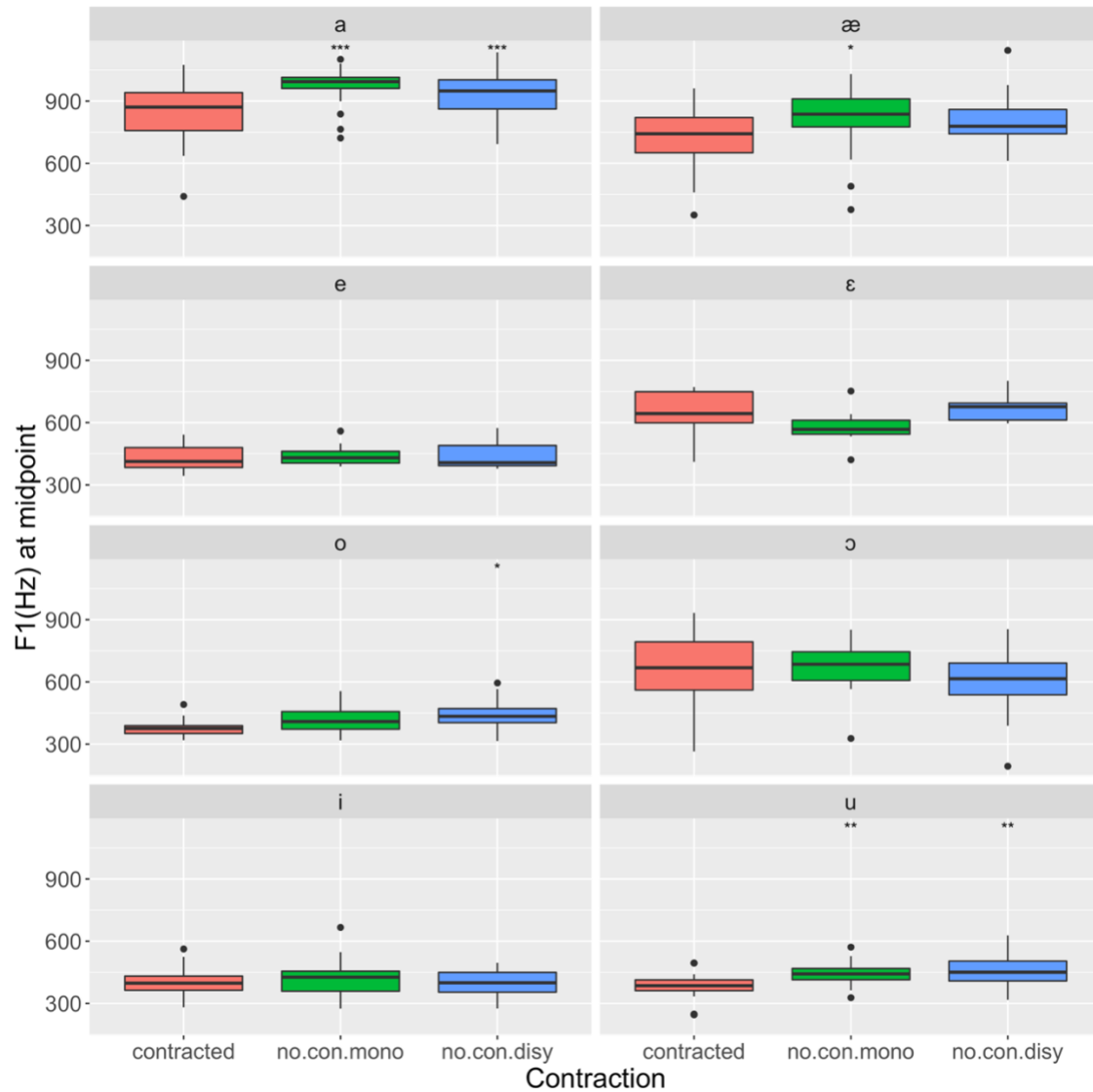


Figure 22. F1 (Hz) at midpoint—Experiment II. [No.con.mono = non-contracted monosyllabic, no.con.disy = non-contracted disyllabic. Asterisks indicate significant levels.]

4.3.6.1.2 F2

- Contracted vs. monosyllabic lexical

The F2 values for [e, o, ɔ, u] are different in the contracted form than in the monosyllabic lexical form, as shown in Figure 23 (red bar vs. green bar). More specifically, for vowels [o, ɔ,

u], the F2 in the contracted word is higher than that in the monosyllabic lexical word, suggesting fronting. For [e], the F2 value is lower in the contracted words than in the monosyllabic words ($t(7) = -2.55, p < .05$), suggesting backing in the contracted form. No difference was observed for [a, æ, ε, i]. Paired Welsh tests were conducted, and the results are:

F2 higher in contracted syllable	F2 lower in contracted syllable	No difference
[o]: $t(9) = 2.29, p < .05$	[e]: $t(7) = -2.5, p < .05$	[a]: $t(22) = 1.6, p = .12$
[ɔ]: $t(21) = 2.57, p < .05$		[æ]: $t(16) = 1.59, p = .13$
[u]: $t(18) = 3.5, p < .01$		[ε]: $t(9) = 0.7, p = .65$
		[i]: $t(22) = -1.3, p = .31$

- Contracted vs. disyllabic lexical

As shown in Figure 23 (red bar vs. blue bar), three vowels, [e, ɔ, u], exhibit differences between fully contracted and disyllabic words. For [ɔ] and [u], the F2 values in the contracted word are higher than those in the disyllabic lexical words, suggesting fronting. Mid-front vowel [e] has a lower F2 value in the contracted form compared to the disyllabic form, suggesting it is more back in the contracted forms. F2 values are not different for vowels [a, æ, ε, o, i]. Paired Welsh tests were conducted, and the results are:

F2 higher in contracted syllable	F2 lower in contracted syllable	No difference
[ɔ]: $t(21) = 2.31, p < .05$	[e]: $t(7) = -2.98, p < .05$	[a]: $t(22) = 0.3, p = .98$
[u]: $t(18) = 5.8, p < .0001$		[æ]: $t(16) = 0.6, p = .95$
		[ε]: $t(9) = 0.99, p = .35$
		[o]: $t(9) = 0.69, p = 0.51$
		[i]: $t(22) = -0.2, p = .84$

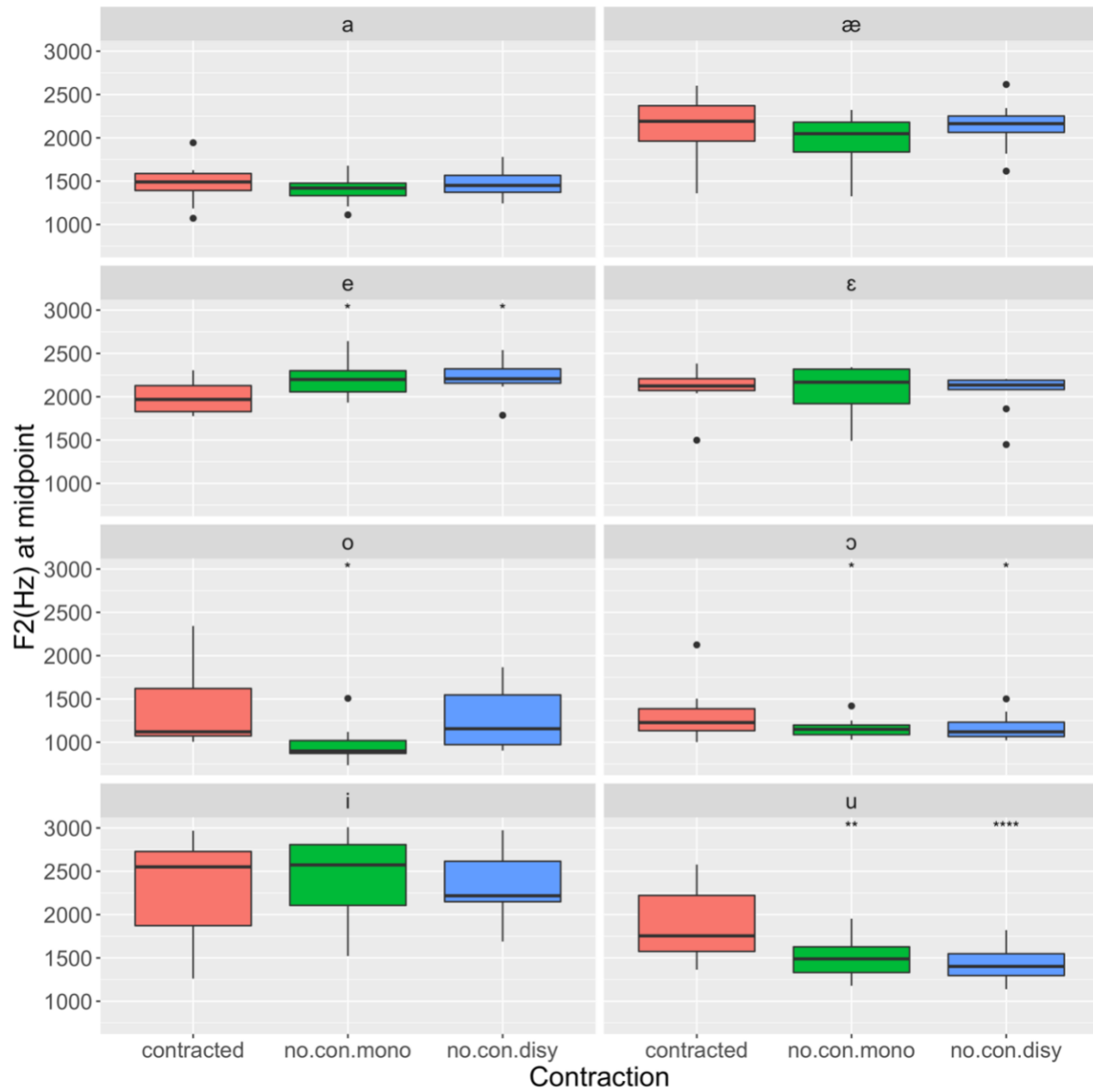


Figure 23. F2 values (Hz) at midpoint—Experiment II. [No.con.mono = non-contracted monosyllabic, no.con.disy = non-contracted disyllabic. Asterisks indicate significant levels.]

4.3.7 Interim summary

Experiment II confirmed the findings of Experiment I, showing that the vowel ratio in fully contracted words patterns better with that of a monosyllabic lexical word than a disyllabic word, and that the vowels in the fully contracted words are different from those in the monosyllabic words in height and/or backness. More importantly, by adding data comparing the fully contracted syllables and their non-contracted disyllabic forms, Experiment II further explored the vowels in these conditions and showed that the vowel ratio does not pattern with the disyllabic forms, but the vowel qualities, either manifested in F1 or F2, are different. The vowel space for contracted words is generally smaller than that of lexical words, whether monosyllabic or disyllabic.

4.4 Discussion

The results of the two experiments were slightly different yet similar to a larger extent. They all point in the same direction and to general conclusions, which suggests that the different methodologies and procedures produced similar results. This consistency of data suggests that the findings are reliable and independent of experimental designs. Below I will first compare the findings and address the general implications, and then discuss the possible cause for the differences and the implications. Finally, I will discuss the other issues of the experiment and possible explanations for these complications.

4.4.1 Length and vowel ratio

With regards to the absolute length values, the two experiments were less consistent because of the nature of the experiments. The durations of either vowel or syllable in the

sentence repetition task of Experiment I are both comparable and unpredictable, as all target words were elicited in the same experimental setting and embedded in carrier sentences. In this style of speech, the contracted words are longer than the monosyllabic lexical word, possibly due to the underlying syllable count (i.e., two syllables versus one). Relatedly, the vowels are longer in the contracted forms than in the lexical form. However, such difference only holds for about half of the vowels and is inconsistent across the vowels and words. But in Experiment II, the contracted words were elicited with a contraction task, while the monosyllabic words were elicited in a wordlist reading task. Although both tasks require attention to speech, the latter may be more formal and thus give more attention to the details of pronunciation. Furthermore, the most common strategy that participants adopted in the self-paced contraction task was to say the target word faster and faster until they reach the contracted level, which was not observed in the wordlist task. Given these contrasts, it is not surprising that the word lengths, as well as vowel lengths, are drastically shorter in the contracted as opposed to a non-contracted di-syllabic word.

The problem of comparing absolute duration values is not an issue when the calculated vowel ratios are compared, as the ratio values are independent of speech rate that may be different for different tasks especially for Experiment II. The two experiments, although conducted under different experimental settings and presumably with different formalities of speech, agree that except for a few tokens, the vowel ratio cannot differentiate contracted words from lexical words. Both experiments challenge Kuo's (2010) conclusion that the vowels in the fully contracted syllables take up a much bigger proportion in the syllable than the vowels in the lexical monosyllabic word. The results of the two Rugao experiments exhibit the opposite trend, suggesting that the vowel ratio of a fully contracted disyllabic word is not, at least not consistently, different from the vowel ratio in a monosyllabic lexical word. And this similarity is

true, independent of the nature of the experiment, whether it is natural sentence repetition or explicit contraction and conscious reading. This in turn means that a fully contracted syllable is proportionally and structurally adapted to a lexical syllable at least for some vowels and words.

On the basis that a vowel takes up a certain portion of the syllable duration whether it is fully contracted or lexical, the total syllable durations *can be* different between the fully contracted and lexical word. From the findings on absolute values, one can see that at least some contracted words are longer than their monosyllabic counterpart. It is a reasonable stipulation that the vowels are fully adapted to the lexical syllable structure, and if the syllable is longer, so is the vowel. It is possible that other elements, such as the consonant duration and the corresponding tonal contour, make the distinction between a contracted and a lexical syllable.

4.4.2 Incomplete vowel deletion

Unlike the duration data, the vowel quality data are relatively consistent. The results of both experiments suggest that, generally speaking, vowel qualities in fully contracted syllables are different from those in lexical mono-syllabic words in at least one dimension. Combining the two experiments, similar vowel quality differences were observed for both contracted vs. monosyllabic lexical and contracted vs. disyllabic lexical comparisons. At the same time, there is much less or no difference between the lexical forms, whether in the non-contracted disyllabic form or the monosyllabic forms, the two of which show great consistency in the vowel qualities. This further suggests that the vowel quality differences are generally between contrasting contracted forms and lexical forms, no matter how many syllables the lexical form has.

More specifically, the vowel quality difference manifests mostly in one dimension along the vowel space, height and/or backness. Height-wise, as indicated by the F1 differences, the low

vowels are higher⁴⁶ in the contracted forms than the lexical forms, whether the lexical is monosyllabic or disyllabic. But critically, mid vowels are more stable with regards to height. Even when there is a height difference for some mid vowels, as shown in Figure 19 and Figure 22, the difference is much smaller than seen in the cases of low vowels. Mid vowels mainly exhibit backness differences between the contracted form and the lexical form. Mid-back vowels are more front and mid-front vowels are more backed in contracted forms than in lexical forms, all pointing to the center of the vowel space.

Crucially, great consistency in the direction of difference is observed, even for changes that do not have statistical strength. The low vowels, if a difference exists, are always higher, and sometimes more back. Mid-front vowels are always more back, while mid-back are vowels always more front, if difference exists. This consistency verifies that the vowels are not randomly different, but different because of something consistently there, the ‘deleted’ [ə] in the first syllable. Only the seemingly deleted mid-central vowel can make such a consistent change to the surviving peripheral vowels. The vowel quality differences further suggest that the deletion of [ə] is *incomplete* even in the fully contracted syllables, shown by a changed height and/or backness of the surviving vowel.

According to the definition of a fully contracted syllable in this dissertation, the seemingly deleted [ə] vowel is not only auditorily absent but also invisible on the spectrogram. But it still leaves its quality remanence on the remaining vowel and centralizes these non-schwa vowels. This means that vowel coalescence must have been involved in some ways in this process, a phonological process that changes the quality of vowels by merging the features, as

⁴⁶ High vowels are also raised; I will discuss high vowels in the following subsection.

commonly seen in Bantu languages. For example, $a + i \rightarrow e$; $a + u \rightarrow o$, in Shona (Harford, 1997). The type of coalescence in syllable contraction is different from the canonical cases of vowel coalescence in that it alters quality of the surviving vowel, instead of merging and producing “another” vowel. The category of the vowel does not change, but the vowel bears the remanences of the deleted vowel and gets “pulled” to the direction of the deleted vowel.

The incomplete deletion links the syllable contraction process, at least for the fully contracted syllables, to the widely discussed issue of incompleteness in the phonological changes (Ellis & Hardcastle, 2002; Manaster Ramer, 1996; McCollum, 2019; Port & Leary, 2005; Warner et al., 2004). As Kuo (2010) put it, the fully contracted output looks/sounds like a lexical monosyllabic word, making contraction look like a process of neutralization. Kuo’s work also suggested that such neutralization is incomplete, but with different aspects of evidence. Despite the differences of experimental results between Kuo (2010) and the current study, it is a clear conclusion that the contracted syllable and the lexical syllable are different in some ways, such as syllable duration, vowel duration, vowel ratio, and vowel quality.

A participant at MidPhon 24⁴⁷ (Milwaukee) expressed his concern that the neighboring segments might have contributed to the vowel quality difference, as the immediately adjacent segments might change the formant values. I agree that some consonants, especially glides, nasals, and rhotics, are known to change the frequencies of adjacent vowels. However, it is unlikely that these segments changed the vowels to the degree that the comparisons were unmeaningful because the consonants surrounding the comparing vowel are carefully controlled. All segments of the corresponding monosyllabic lexical words were based on the expected

⁴⁷ The 24th Annual Mid-Continental Phonetics and Phonology Conference.

contraction form of the disyllabic word. For example, for the matching [jæn] (contracted) vs. [jæn] (lexical), even if the glide and the nasal do change the formant values of the vowels, they change the formants in *both* the surviving vowel and the lexical vowel. Even for tokens that modification was a must, special attention was paid to the consonants that are adjacent to the target vowel, especially /j, w, n, ŋ, ɹ/, and these are not substituted in any token. Instead, other segments are modified to make a lexical word. In all such cases, the two comparing syllables, i.e., the contracted vs. lexical are comparable and the comparisons of the vowels are meaningful. When differences arise, there must be some other factor(s) than the consonants.

Another participant at MidPhon 24 pointed out that the differences in vowel space shown in Experiment II may be caused by the simple fact the vowels are generally more centralized in careful speech. I acknowledge such possibility, as it is common to observe vowel centralization in careful speech cross-linguistically, and that there was style change in Experiment II. But I would like to add to this discussion by pulling the results of both Experiment I and Experiment II into the view. Both experiments tested the same target vowels and words and found very similar patterns of vowel quality change. Particularly, the speech style in Experiment I, unlike in Experiment II, did not change throughout the procedure, but the same vowel quality change was observed and with greater effect size. This in fact completes the bigger picture of how consistently speakers behave under different speech registers with regards to maintaining the vowel quality contrast.

Finally, with the altered vowel quality of the contracted syllables, it seems a clear conclusion that the vowel space that the contracted vowels take is smaller than that taken by lexical vowels under this experimental setting. Since the deletion of [ə] is incomplete, the remanences of the seemingly deleted mid-central vowel pull the surviving peripheral vowels to

the center of the vowel space, resulting in a smaller vowel space. However, it is too early to take this to any further conclusion about the general vowel space of the contracted vowels. Kuo (2010) concluded that the contracted syllables use a wider vowel space to mark the contraction, but this was based on the conclusion that only [i, u] are higher in the contracted words than in the lexical words. It is a fair prediction based on the findings of this study that when the competing vowels are all peripheral vowels, the actual vowel space may not change drastically. Put in short, whether the vowel space becomes bigger or smaller depends on the deleting vowel and the surviving vowel, and no clear general prediction can be made based on the available evidence.

4.4.3 Syllable contraction

With the discussions of length, ratio, and formants, now we can return to the original question: what is syllable contraction in the context of Chinese languages, based on the findings above? First, is a contracted syllable really one syllable? Does it resemble its underlying disyllabic form or is it somewhere between one and two syllables? Past research is more inclined to say that a contracted syllable is a one-and-half syllable, with a longer vowel, a bigger vowel ratio, and the tonal contour of the underlying disyllabic form (Cheng & Xu, 2009; Kuo, 2010; Wong, 2006). However, this study challenges this conclusion by revealing that, although the vowels in the contracted syllable are usually longer than a lexical vowel, the vowel ratio of the contracted syllable is more similar to that of a single syllable of a lexical word than a disyllabic lexical word. Based on the results of the experiments presented above and the narrow definition of a fully contracted syllable that is adopted for this entire chapter, I would like to put forward an opposing conclusion here: at least in Rugao, a fully contracted syllable resembles a single syllable structurally. The partially contracted syllable may fall between one syllable and two

syllables, supposedly with longer total vowel lengths and bigger vowel ratios than a single syllable, but not to the extent that it is a two-syllable word.

To become a structurally single syllable, the underlying disyllable word has to lose some segments, but this process does not only involve segment deletion, but also vowel quality change. The results shown here suggest the seemingly deleted vowel is still “there” to affect the quality of the remaining vowel. This again suggests that the contraction process also involves coalescence of the vowels, a process that is necessary in order for the surviving vowel to bear the vowel features of the deleted vowel.

Based on the discussion above, a fully contracted syllable should mean the following:

- The contracted syllable has a sonority contour that is close to a single syllable.
- The contracted syllable has a length and vowel ratio that are more like a single syllable than its underlying disyllabic form.
- Some segments of an underlyingly disyllabic word get deleted, making it similar to a monosyllabic word with one vowel in the nucleus, but the lexical distinctions remain on the surviving vowel.

The process of syllable contraction then starts with a disyllabic word with two vowels. The first step is the loss of inner consonantal segments, which feeds the next step by pulling the two vowels closer and making the resyllabification possible. As the inner consonants are no longer at the original inner syllable boundary, the sonority dip is lost, the resyllabification happens and the syllable count reduces. Then the two vowels compete for the limited time slot, the more sonorous vowel is likely to win the competition and survive. Finally, vowel coalescence happens, with the two vowels influencing each other’s height and/or backness. Optionally, the losing vowel deletes and leaves a vowel trace on the surviving vowel. Otherwise, both vowels survive, and the less

sonorant vowel reduces. The last step separates the fully contracted syllable from the partially contracted syllable, as the former ends up having one single syllable with one “colored” surviving vowel, and the latter with one-and-half syllables with one full vowel and a reduced vowel. Two example derivations are shown below:

- | | |
|---|--|
| 20) /səʔ ₁₅ .aɪ ₂₁ / | disyllabic word, with two vowels |
| [sə ₁₅ .aɪ ₂₁] | inner consonantal segments delete |
| [₄₈ səaɪ _{151/21}] | resyllabification |
| [səaɪ _{151/21}] | vowel competition, based on sonority hierarchy |
| [səaɪ _{151/21}] | vowel coalescence |
| | vowel deletion (optional) |
| | |
| 21) /bəʔ ₄₄ .uɔj ₂₁ / | disyllabic word, with two vowels |
| [bə ₄₄ .uɔj ₂₁] | inner consonantal segments delete |
| [bəuɔj ₄₂₁] | resyllabification |
| [bəuɔj ₄₂₁] | vowel competition, based on sonority hierarchy |
| [bəuɔj ₄₂₁] | vowel coalescence |
| | vowel deletion (optional) |

Such derivation will not only apply to the words that have appeared in the experiments, but the syllable contraction in general, such as the example below:

⁴⁸ The tone is marked 151/21 because of the variation.

22) /tʃa21.xoW21/	disyllabic word, with two vowels
[tʃa21.oW21]	inner consonantal segments delete
[tʃaow21]	resyllabification
[tʃaow21]	vowel competition, based on sonority hierarchy
[tʃaow21]	vowel coalescence
	vowel deletion (optional)

Lastly, although the investigation of partially contracted syllables is beyond the scope of this project, the fact that the definitions and derivations above still apply means that the partially contracted syllables are fundamentally not different from the fully contracted syllable. The process of contraction should be similar. The difference between a partially contracted syllable and a fully contracted syllable just lies in the treatment of the less sonorant vowel in the last step. The case of partial contraction is spotted when vowel deletion does not happen, but both vowels end up being preserved. The coalescence happens before the deletion applies, so I will predict that the vowels in the partially contracted syllable are also different from the lexical vowels: the losing vowel will reduce/centralize, while the winning vowel moves towards the other vowel.

4.4.4 Remaining issues

4.4.5 High vowels

One big challenge of this study comes from the high vowels, as the high vowels [i, u] do not seem to follow the general trend, especially for the vowel quality data. For [i], there is no difference in either height or front/backness for both experiments. For [u], no difference was found for height, but it is more back in Experiment I. By contrast, [u] is both higher and more

fronted in Experiment II. Kuo (2010) also found the high vowels are higher in the lexical forms. In short, compared to the corresponding lexical vowels, [i] and [u] in the contracted forms are either no different or even more peripheral rather than being centralized.

The high vowels are special because of the likelihood of them becoming glides. It is possible that at least in some cases, high vowels become [j]/[w], and are thus exempt from processes that target vowels, such as vowel coalescence. However, this will only possibly affect [bəʔ.xuj] → [buj]/[bwi], but not the other words, the latter will necessarily need the vowel in the nucleus. Even [bəʔ.xuj] → [bwi] does not seem likely, seen in Figure 24 and Figure 25, as the contracted form has similar F1 and F2 values to the [u] in its non-contracted form [bəʔ.xuj]. Meanwhile, the comparing lexical word [tui] can be pronounced as [twəj] in careful speech. Without the adjacent [j], the [u]/[w] in [twəj] is less fronted compared to the [u] in [buj] (Figure 25). This could explain why contracted [u] is more fronted in Experiment II but not in Experiment I because only the wordlist reading in Experiment II is careful speech, making [twəj] more likely to be elicited.

Tense and lax vowels are not contrastive in Rugao, and check tones and closed syllables are strongly correlated with tense vowels: [iʔ], [uʔ], [fɯ], [pɪ]. The glottal stop and check tone can be lost at the right edge of this disyllabic word, rendering [tsən.jʊ], although the underlying form is [tsən.juʔ]. It is possible that, in order to maintain the lexical contrast, the contracted form of this word, [tsuʔ], keeps the underlying tense vowel. With the [ə] centralizing the vowel, the underlyingly high-back vowel finally results in a strange vowel that is both higher and more front than a lexical [u]. Different from /u/, the other high vowel /i/ does not have this tense/lax alternation, even though the check tone and glottal stop may also get lost after /i/.

Finally, as can be seen in Figure 24 and Figure 25, there is great inter-speaker variations in the high vowels, shown in the height of each of the boxes. It is possible that some speakers pull the data to the two ends. There needs to be more studies on high vowels and how they pattern in the syllable contraction process.

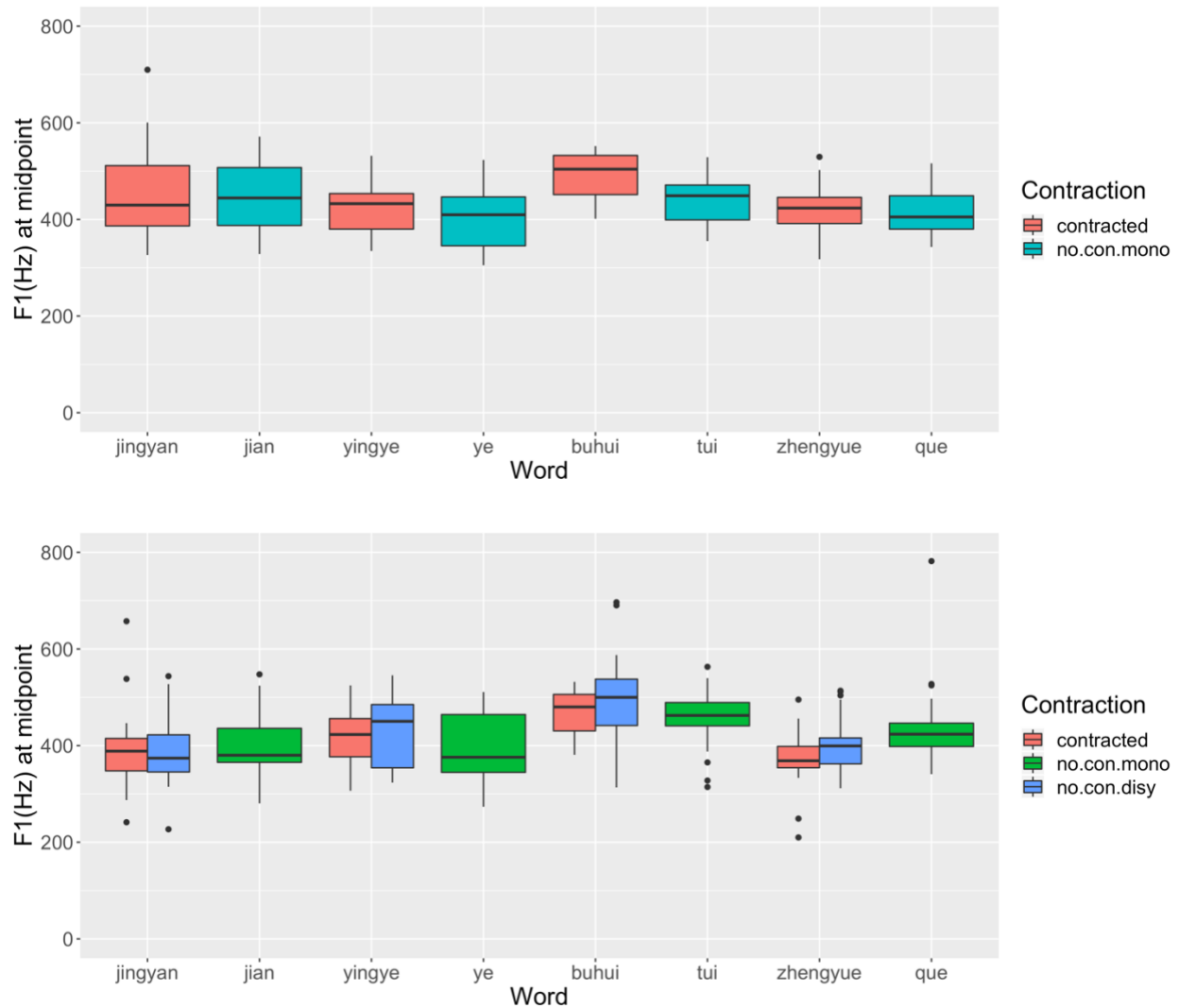


Figure 24. F1(Hz) for words with high vowels [i, u]. [Experiment I data on top, and Experiment II data below]

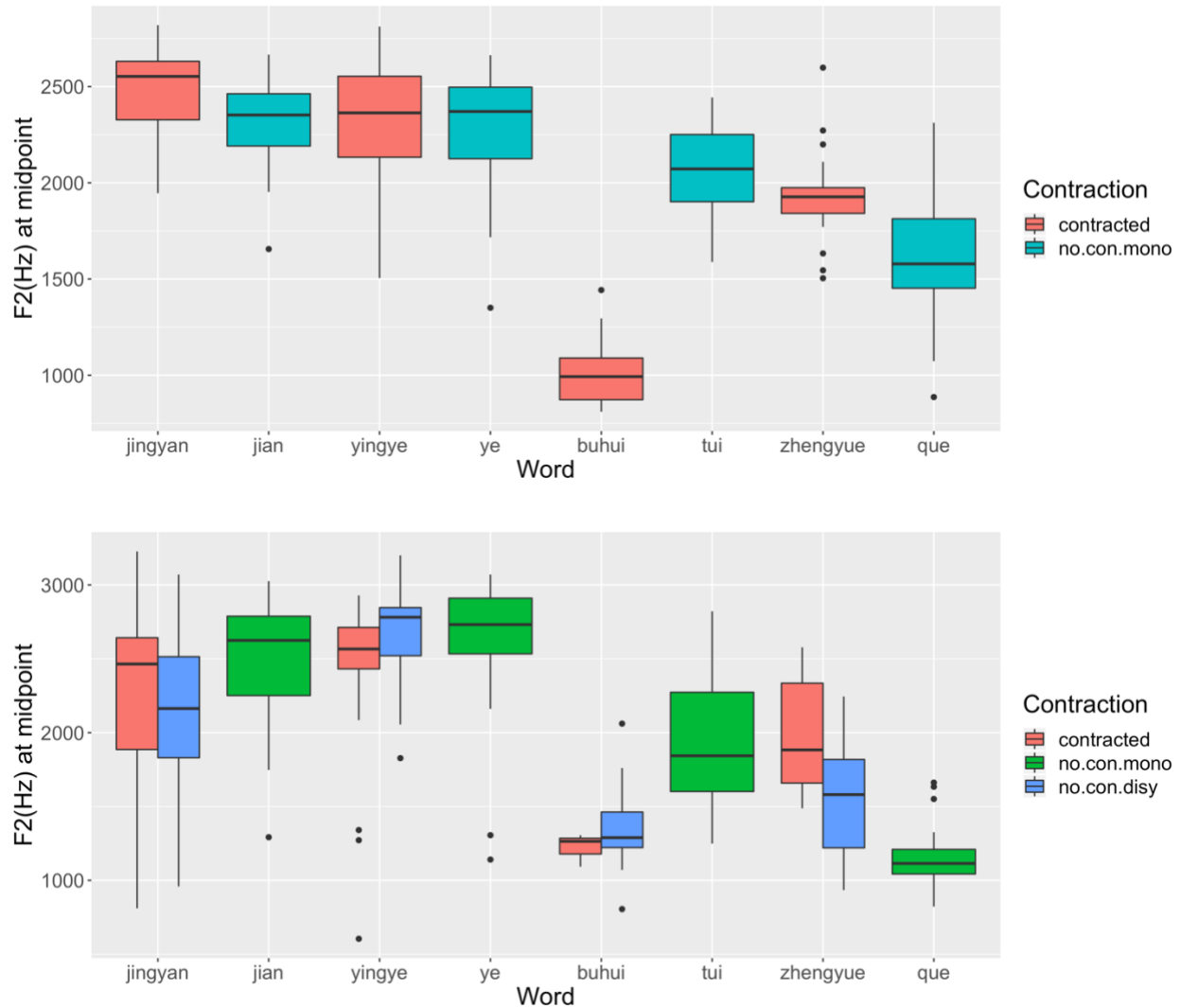


Figure 25. F2(Hz) of words with high vowels [i, u]. [Experiment I data on top and Experiment II data below]

4.4.6 How much difference can the experimental methodology make?

Generally, how much can methodology affect the result? The study presented here employed two different experiment designs for the same research goal. The results of the two experiments, as can be seen above, are generally consistent.

The comparable part is the fully contracted vs. monosyllabic word comparison. For this comparison, the differences between the two experiments are summarized below:

- For vowel ratio, [o, u] are different in Experiment I, but not in Experiment II. [ɛ] is different in Experiment II, but not in Experiment I.
- For F1 values, [ɛ, o] are different in Experiment I, but not in Experiment II. [u] is different in Experiment II, but not in Experiment I.
- For F2 values, [ɛ] is different in Experiment I, but not in Experiment II. [e] is different in Experiment II, but not in Experiment I.

It is evident that vowels [ɛ, e, o, u] are the major contributors to the differences in the results. First, by “not different”, I actually mean “no statistically significant difference”. For all the vowels except for [u], the generalizations that are made on all the vowels still apply: the mid-front vowels are more back and mid-back vowels are more front in the contracted syllable. The lack of difference may simply be due to the lack of statistical power. As shown in Table 12, the number of tokens for [ɛ, e] is relatively small in Experiment I. Meanwhile, the unpredictability spikes simply because of the lack of sufficiently large data, so the significance for [ɛ] in F1 of Experiment I could also be the result of some extreme speakers. Second, as seen in the previous subsection, the high vowels do not generally follow the general trend for various reasons. It is not surprising to find [u] not behaving consistently.

Word	N fully contracted— Exp. I	% fully contracted— Exp. I	N fully contracted— Exp. II	% fully contracted— Exp. II
[səʔ15 + aɪ21]	109	10.7%	24 ⁴⁹	9.4%
[ɪən15 + ka21]	46	4.5%	14	5.5%
[pʲən21 + ɛjæn21]	68	6.7%	22	8.7%
[jəʔ5 + jæn55]	103	10.2%	26	10.2%
[pəʔ5 + ɲɛ21]	16	1.6%	17	6.7%
[pən213 + lɛ15]	36	3.6%	14	5.5%
[tɛjən21 + xej21]	34	3.4%	15	5.9%
[ɛjən21 + nej21]	8	0.8%	15	5.9%
[ɛjə55 + jən55 + kha213]	140	13.8%	11	4.3%
[tsəʔ5 + jə55]	101	10.0%	22	8.7%
[tsəʔ5 + jow213]	115	11.3%	23	9.1%
[phən15 + jow213]	17	1.7%	0 ⁵⁰	0.0%
[jən15 + iʔ5 + phu21]	94	9.3%	15	5.9%
[tɛən21 + in21]	70	6.9%	24	9.4%
[pəʔ5 + xuj21]	30	3.0%	19	7.5%
[tsən21 + juʔ21]	83	8.2%	17	6.7%
1070	256			

Table 12. Number of tokens analyzed for each experiment

As a summary, I do not believe these minor differences in the results pose a substantial challenge to the general conclusions of this chapter, as the differences are small and explainable. This study adopts a definition of fully contracted syllable and the methodology of experiments that are distinctive from other studies, however. Compared to the minor differences in the results

⁴⁹ Only one contracted token was chosen for each participant in Experiment II, so the numbers in this column are small.

⁵⁰ The variation of pronunciation for ‘friend’ [pən.iw]/[pən.jiw] was overlooked and this caused no usable token for this part.

of the two experiments, the general conclusion of this study is substantially different from Kuo (2010).

Both the vowel ratio data and formant data presented in this chapter are different from those of the similar study in Taiwan Mandarin (Kuo 2010), the latter found the opposite pattern. Vowel ratio wise, Kuo (2010) found that 1) the length of the contracted syllable is still similar to the original disyllabic form. 2) the remaining segments in the contracted syllable, especially vowels, are longer. 3) the fully contracted syllables have a larger proportion of vowel nucleus than a monosyllabic lexical word. She then implied that such differences may be cues the speakers can employ to distinguish the contracted vs. lexical words in the perception experiment. With regards to vowel quality and vowel space, however, only high vowels were higher in contracted syllables and this difference was only observed for male subjects. These differences were likely caused by the following reasons.

The slightly different methodologies used in the two studies might have made the difference. As discussed in the methodology section, the deleting vowel of the target words was not consistent in Kuo's (2010) study, which might have masked the formant differences that are otherwise observed with better controlled stimuli. The other difference in methodology was that the same carrier phrase was used for the comparing contractible word and lexical word, which means the pair of words are in the same position in the sentence. The setback of using a consistent carrier phrase is that the target positions are prone to be figured out by the participants. If that happens, it is unsurprising that they would hyper-correct or perform for these words to make the pair somehow different. The durational/ratio difference might be one tool that was ready to be utilized to be utilized to make the distinction. Another difference is in the measurement methods. The author of this dissertation referred to three criteria, including the

auditory judgment, the visual examination of the spectrogram, and the intensity curve, to determine the fully contracted syllables. Instead, the Kuo (2010) experiment adopted the trough depth measurement (Mermelstein, 1975; Myers & Li, 2009) and largely relied on the value to decide on the degree of contraction. Due to the nature of the phonetic realization of speech, it is possible that some partially contracted tokens were treated as fully contracted. The author witnessed such complication during the Praat annotation process and often spotted cases where the intensity line was simply not a good indicator of the number of syllables.

Despite the differences in the experimental designs, however, the two languages may be just different in how they each treat syllable contraction. Both languages employ some strategies, whether maintaining the vowel ratio difference or vowel quality difference. After all, the two studies are still unified in a way that incomplete neutralization is observed in the process of full syllable contraction.

4.4.7 Back to the centrality issue

Finally, do these experiments confirm the centrality effect? The simple answer is yes. In the previous chapters, the peripheral > central bias seems to be present but with slightly weaker statistical support. The production data in the experiment show that the centrality effect does exist as a strong factor, but exceptions are also found in many tokens, as the [ə] deletion rate is not 100%. Even though the vast majority of contracted tokens preserve the non-schwa vowel, some contracted words had [ə] preserved as the only segment in the nucleus. As shown in Figure 26, the word [pən.ju] (‘friend’) is contracted as [pəj], instead of the predicted [pju], by this speaker. While [pju] is the attested form of contraction in the corpus data, this type of contraction is unexpected, but quite common in the experiment. This poses a challenge not only to the

peripheral vowel bias, but also to the Edge-In Effect (Yip 1988), as the right edge segment is not preserved. The existence of such forms of contraction adds to the complexity of the syllable contraction data. But other than this, all the contractible tokens had only three types of output: a. the fully contracted form with the non-schwa vowel, b. the partially contracted form with both vowels, and c. the disyllabic form without contraction. One can still stipulate that the forms that are not confirmative with the centrality bias are outliers that require more sophisticated investigation. In this sense, the production data presented right above are more consistent with the “more real speech” data in the corpus and supplements the experimental data in Chapter 3.

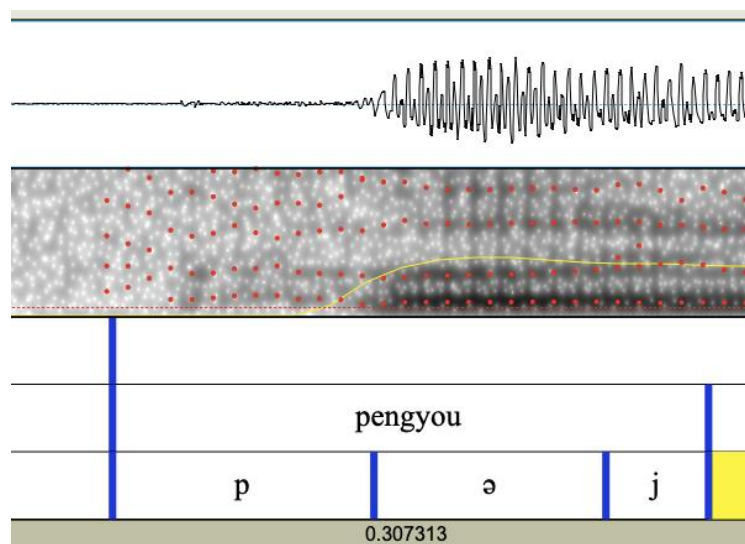


Figure 26. Screen shot of Praat interface on [pəj], contraction of [pən.jow] ('friend')

4.5 Conclusions and implications

In conclusion, the vowels in the fully contracted syllables are to a large extent similar to a lexical vowel but maintain subtle phonetic contrasts. First of all, although the vowel durations are longer in fully contracted syllables than in corresponding lexical monosyllabic words, the units of higher level, i.e. the words, are also longer. The vowel ratio in a fully contracted syllable

patterns with a monosyllabic lexical word and not with a disyllabic lexical word, the latter being the non-contracted form of the contracted syllable. The fully contracted syllable is proportionally similar to a lexical syllable, as the vowel takes up the same percentage of the syllable duration. Meanwhile, the vowel quality, reflected by the first and second formants, are different in some ways. Either F1 and F2, or both F1 and F2, are seen as different in contracted syllables versus in lexical words.

Crucially, the vowel does not become another vowel in the process of full contraction. For most vowels, namely the low and mid vowels, the quality along only one dimension changes, either height or backness. The high vowels and sometimes mid vowels change both height and backness. These one-dimension changes are too subtle for the overall vowel quality and thus not sufficient to shift the perception of the vowel to a completely different vowel category. In this sense, the transcriptions using the phonetic symbol may be in fact quite close, although they cannot faithfully reflect the subtle differences in the actual segments. This explains the reason why contraction may seem like a process of neutralization, and the lexical contrast, at least segment-wise, is lost in the surface form.

Despite the similarities, however, the quality differences in the vowels are maintained. The vowels in the contracted forms are not identical to a lexical vowel. The fully contracted syllable, largely resembling a lexical monosyllabic word, maintains its phonetic differences as a resolution to the loss of lexical contrast in the surface forms. The underlying lexical distinction between the words, as well as the contrast between the vowels, are not neutralized. Although different patterns were observed in Rugao and in Taiwan Mandarin (Kuo, 2010), both languages provide some evidence for full syllable contraction in Chinese languages being a process of incomplete neutralization (Port & Leary, 2005).

The broader implication here is that, if the syllable contraction is a process of incomplete neutralization, then it may provide another way into the discussion on the gradience of phonological representation, as noted in many other studies (McCollum, 2019; Port & O'Dell, 1985; Port & Leary, 2005; Warner et al., 2004). As a long-established topic for discussion, the nature of phonological representation seems to have obtained considerable empirical evidence for both the categoricity argument and the gradience argument. Phonological categoricity would predict that any process that alters the underlying representation will lead to another category of segment. In the context of syllable contraction, this would mean that the contracted form of any lexical word would only contain segments that are categorized as ones in the underlying inventory of this language. However, the fact that the vowel in the contracted form is not identical to its lexical counterpart indicates that the non-typical vowel is existent in the language. It is an open question whether the variation of the output vowel quality is simply free variation or conditioned by factors such as speech rate and sociolinguistic background, but great variation of resulting F1 and F2 values is observed from the experiments. The coalescence of vowels does not change the underlying vowel to a completely different vowel category, but to a quality that is *close* to the lexical vowel in the vowel space. The existence of a non-typical vowel may be linked to the gradience of phonological representation. Due to the design of the experiments, I refrain from making a strong claim on this issue, but I would like to add to this discussion by providing data from a less studied phonological phenomenon.

This study focuses on the fully contracted syllables but calls for further investigation of the partially contracted syllables. According to the analysis presented above, the partial syllable contraction should not be fundamentally different from the full syllable contraction, which predicts the contracted vowels to have a different quality as well. Many other issues and

questions remain, such as how the high vowels are unpredictable in contraction, the possible correlation between the degree of contraction and the degree of vowel changes, the inter-speaker and intra-speaker variation in the vowel production, etc. The investigation of the phonetic details of the contracted syllable enlightens us and helps us better understand not only the contraction itself, but also broader theoretical questions such as incompleteness and gradience. The interesting but conflicting results calls for more research on these topics.

5 CONCLUSION

5.1 Summary of findings and further discussions

In this dissertation, I investigated the phonological patterns and phonetic details of syllable contraction in Rugao with regards to how the segment selection process determines the contracted output. The emphasis was on the vowels, including principles of vowel selection and the phonetic properties of the surviving vowels. On the basis of the phonological analysis, I further studied two issues concerning vowels using experimental techniques and acoustic measurements, 1) how do the two vowels compete for survival in the contracted output regarding the linear order of the vowels and the sonority ranking, and 2) whether the deletion of the losing vowel is a true complete deletion process.

The phonological analysis laid the foundation for the ensuing investigation on vowels. Based on the phonological system of the language of Rugao, I summarized the general patterns of Rugao syllable contraction in the corpus data with regards to the edge segments and the vowel nuclei. First, the initial and final segments of the pre-contraction form are most likely preserved in the contracted output. This Edge-In Effect (Yip, 1988) is similar to the findings in other Chinese languages. In particular, the contracted syllables are irregular syllables by nature, with much more freedom in distributing the segments in the syllable, which is contrary to the pre-specified syllable structure of the lexical syllables. I further updated the Edge-In analysis to account for Rugao data by specifying that the edge segments must be preserved, whether they are consonant or vowel, unless the phonotactics of the language overrides them. For contractible words that end with a vowel, this means the vowel gets the final timing slot and nucleus position,

while for those that end with a consonant, this means this consonant gets the final timing slot while the two vowels compete for one nucleus in the next step.

With regards to the vowel selection, I primarily concluded based on the corpus data that the two vowels in competition follow the sonority hierarchy in determining the winner and loser and that the more sonorous vowel is more likely to be the winner. Following this conclusion, the forced-choice experiment further confirmed the influence of sonority in the vowel competition and ruled out the factor of relative linear order of the vowels. Generally, the vowels of higher sonority are more likely to survive than the competitor of lower sonority ranking, assuming the universal vowel sonority hierarchy based on height and centrality. The experiment also confirmed that the vowel sonority along the dimensions of both height and centrality can bias the vowel selection: the lower vowel, as well as the peripheral vowel, is more likely to be selected for the contracted nucleus. The competition results of vowels of the same height further suggest that vowel sonority may even go beyond height and centrality. The front/backness or roundedness of the vowel may also contribute to the sonority ranking in this language, although more evidence is needed. The controversy over sonority and sonority ranking has been actively discussed throughout the development of modern linguistics (Ohala, 1990; Ohala & Kawasaki-Fukumori, 2011). In this dissertation, I did not participate in the debate on whether sonority exists or not. Instead, I assumed that sonority does exist based on the many phonological analyses and phonetic evidence (Parker, 2008, 2011, 2012). Both empirical and experimental data in this dissertation are best accounted for if the vowels are ranked in sonority based on height, centrality and possibly backness or roundedness. In this sense, sonority does not only exist, but also can be a good tool for explaining the patterns of vowel selection in syllable contraction and vowel competition in a broader sense. Further phonetic studies on Rugao vowels

may provide more evidence for supporting the existence of sonority if the phonetically defined sonority ranking coincides with the vowel competition patterns. At this point, the basically consistent patterns of vowel selection in Rugao syllable contraction have already been confirmed to follow all dimensions of vowel sonority. Compared to using three or four independent features (height, centrality and possibly backness/roundedness), using *sonority* as one single non-feature characteristic of the vowels is not only more efficient, but also more accurate. Furthermore, the sonority debate is mainly about the role of sonority, primarily consonant sonority, in segment sequencing (Ohala & Kawasaki-Fukumori, 2011; Zec, 1995). Vowel competition in syllable contraction is focused on the vowels and this issue is fully independent of segment sequencing. Exploring the connection between vowel competition and sonority thus provides another type of data for the discussion of the sonority issue.

On top of the findings of the vowel competition which show that the winning vowel survives in the contraction while the loser deletes, the phonetic details of the winner/surviving vowel is further investigated with acoustic measurements of the duration and vowel formants for a deeper look at the contracted syllable. The distinct phonetic properties of the contracted syllable and the surviving vowel suggest that the deletion of the losing vowel is incomplete. The incompleteness is manifested in two ways. First, the contracted vowel is longer than the lexical vowel in general, although the ratio of vowel in the duration of the syllable may or may not be different. Second, the contracted vowel has different F1 and F2 values than their lexical counterparts, suggesting the vowel quality is not the same for the contracted syllable and the corresponding lexical syllable, the two of which are seemingly neutralized especially to the ear of the listeners. The seemingly lost underlying, lexical distinction is thus preserved in the phonetics. The study of formants adds on top of duration and shows that the incompleteness of

vowel deletion in syllable contraction can also be manifested by vowel quality. These findings pair well with the previous results that highlight the perceptual distinction between a fully contracted syllable and a lexical syllable with seemingly the same segments and possibly even the same tones (Kuo, 2010). This leads to the discussion of incompleteness of phonological processes in a broader sense, following the findings on incomplete neutralization of final obstruents (Dmitrieva, Jongman, & Sereno, 2010; Port, Robert F. & O'Dell, 1985; Port & Leary, 2005), incomplete vowel harmony (McCollum, 2019), incomplete place assimilation (Ellis & Hardcastle, 2002) and so on. Crucially, the incompleteness of both the vowel deletion and syllable neutralization in syllable contraction is free from the influence of orthography, as the dialect of Rugao does not have its own written forms and the lexicalized contraction forms were carefully excluded from this study. The incomplete neutralization of final obstruents can be influenced by orthography (Warner et al., 2004), and Rugao syllable contraction adds to the family of incomplete phonological processes with a type of data that is free from the influence of orthography. In this sense, the studies of syllable contraction as well as other cases of incompleteness seem to suggest that, without orthography, true complete neutralization cannot occur in either phonology or phonetics.

Furthermore, the study of incomplete vowel deletion in syllable contraction opens a window for the discussion over the nature of phonological representation Ohala (1990). Different types of phenomena and data may support either the categoricity or the gradience account. Syllable contraction adds to this discussion with a less studied phonological process and the incompleteness in this process seems to be in favor of the non-categoricity of phonological representation (Port & Leary, 2005), although other factors, such as phonetic reduction and formant alteration due to a neighboring segment, cannot be completely ruled out at this point. As

seen in the production experiments, the features of the deleted vowel can linger on the preserved vowel without changing the vowel category of the latter. In other words, a vowel can shift to another place in the vowel space according to the place of the deleted vowel, while its phonologically defined vowel category remains the same. This indicates that, at least for vowels, a segment can take a continuous space, possibly suggesting that phonological representation is gradient in nature. However, even though this dissertation provides some data and discussion, further evidence is needed before one can make a stronger claim. The big question of the nature of phonological representation is never easily answered and calls for more research into this matter.

Although the discussion of the three issues concerning vowel selection, corpus-based vowel selection patterns, vowel competition within the framework of vowel sonority, and the incomplete vowel deletion are distributed in three chapters, these issues are inherently connected. As the focus of the study, how the vowels survive or delete is the main concern of the analyses. The selection of the surviving vowel seems like a complete process that is conditioned by the vowel sonority hierarchy. However, even with the predictions of sonority, the deletion of the losing vowel is shown to be incomplete; the fact that the phonetics of the surviving vowel is different from its lexical counterpart is accountable by the fact that the deleted vowel is a schwa. This leads to the more essential question of how the phonetics and phonology interact, or are integrated (Ohala, 1990). This dissertation adds to this discussion by providing more evidence for phonology and phonetics to be integrated at least to some extent, as the phonologically defined process is shown to be phonetically complex, while such complexity cannot impede the process that is deeper in the phonology.

5.2 Remaining issues and future projects

Although I approached the phenomenon of syllable contraction from a variety of perspectives, including phonological analysis, experimental studies, and phonetic measurements, this dissertation still leaves many questions unanswered due to the limits on the scope and the limitations of each methodology. Here, I would like to point out or recapture some broader issues for the entire study and discuss some possible future projects.

First, to achieve some simplification of the analysis, all the analyses presented in this dissertation are on the segmental contraction of this process, which is only part of the complicated details of this phenomenon. The aspect of the tonal contraction while still an important component of syllable contraction was largely ignored. Based on preliminary observations on Rugao tonal contraction, the tones also contract during the contraction process with a similar Edge-In pattern but seem to have more freedom in violating the phonotactics of the language. A similar pattern is observed in the tonal contraction of other Chinese languages (Hsiao, 1995; Wee, 2014). Having said this, the issue of how tones actually contract and interact with the phonotactics of the language needs extensive research. Furthermore, if the tones of the precontraction form can affect the likelihood of contraction and the form of the contracted output, the investigation of tones may offer another perspective for explaining the cases that are not compatible with the current framework.

Second, even with segmental contraction, the dissertation focused on vocalic competition. As noted before, consonants are much more predictable because of the interplay between the basic CV(C) syllable structure of Rugao and the Edge-In, which basically mandates that the initial and final consonants must survive if present. However, whether a consonantal segment survives is not the whole picture of this phenomenon. A previous production study

using Taiwan Mandarin as a target language suggests that the inner consonants, especially nasals and obstruents, may impede or decrease the likelihood of contraction (Cheng & Xu, 2009); note, a similar medial consonant effect was not found in Rugao. In fact, abundant cases of contraction were found with words with medial nasals or glottal stops. This is likely due to the weakening of the Rugao nasals and glottal stops in the coda which leads to the disappearing of the coda and a nasal vowel or oral vowel, respectively. In addition, the production experiment in Chapter 4 offers primary evidence that the contracted consonant may also have discrete differences from the lexical consonants at least in the length. Many other aspects of the consonants can be examined, and these wait for careful research. Furthermore, with the previous finding in Mandarin (Kuo 2010), there should be perceptual distinctions between the fully contracted syllable and the lexical counterpart. If vowel quality difference can contribute to such distinction, will listeners make use of differences in consonantal phonetics to distinguish the two syllables as well? All these questions raised here need investigation on the consonants.

Third, this dissertation concerns mostly cases of full contraction, in which case only one vowel survives. The vowel competition and vowel deletion studies are specifically limited to fully contracted syllables. However, syllable contraction means much more than fully contracted syllables; there are many other cases of partial contraction. Will partial contraction pose challenges to the conclusions on vowel competition and incomplete neutralization? Are there any patterns regarding what kinds of vowel combinations are more or less likely to induce full/partial contraction? More broadly, is partial syllable contraction even true syllable contraction? All the questions are worth investigation, and the exploration of them will give us a better understanding of the nature of syllable contraction.

Finally, two major content chapters were experimental studies, which was in combination with a corpus study. Such a combination enables a more comprehensive investigation of this phenomenon from both macro and micro perspectives. As discussed in each chapter, the experiment designs were the best effort to test a small aspect of the phenomenon while controlling the others, but any experimental manipulation could cause loss of information. For example, the nonce words experiment avoided the complications of other non-phonological factors and focused on the vowels, but the other factors do contribute to this process. How to balance controlled experiment and the real word phenomena is still a broader question that one should ask for any study that relies on experimental techniques.

Syllable contraction in Rugao exhibits both consistent patterns in phonology and intricate details in phonetics. This dissertation serves as a first comprehensive look into the complicated phenomenon of syllable contraction using an understudied language and offers some baselines for future projects, including studies on other aspects of syllable contraction and deeper investigations on the topics that have been explored. The results of this dissertation also call for and facilitate studies that utilize the vast source of languages in the Chinese language family.

APPENDICES

Appendix A: Full list of questions used for corpus data collection

1. What do you like about our hometown?
2. What do you not like about our hometown?
3. Talk about the most memorable event in your life/of this year.
4. Talk about the most interesting event in your life/ of this year.
5. Talk about the most meaningful event in your life/ of this year.
6. Talk about your family.
7. Talk about your hobbies.
8. Talk about your favorite musician/ movie star.
9. Talk about your favorite song/movie/book.
10. How did you come to know your best friend/husband/wife/boyfriend/girlfriend? Any interesting story?

Appendix B: Full list of stimuli for practice session–forced choice contraction

Pre-contraction Word	Option 1	Option 2
pha + xən	phan	phən
pha+ in	phan	phɪn
pha+ xɛn	phan	phɛn

Appendix C: Full list of practice sentences, annotated in standard *pinyin*.

1. *Ni jintian xiawu shenme shihou lai banshi a?*

‘When in the afternoon will you come?’

2. *Wo mashang jiu lai, ni zai denghui zai.*

‘I will come right away; wait for me.’

3. *Rugao you naxie haoshuazi de difang a?*

‘What fun places are there in Rugao?’

Appendix D: Full list of *burn-in* sentences, annotated in standard *pinyin*. Contractible words bolded.

1. *Xiangxia de lao fangzi zhuqilai fan'er shufu.*

‘Old houses in the country are actually comfortable to live in.’

2. *Jinnian dongtian pade hou yishang buhao mai.*

‘Heavy clothes probably won’t sell well this winter.’

3. *Zhehui tuidiao shang na'er qu mai xinde a?*

‘Where can you buy a new one if you want to return it right now?’

Appendix E: Full list of test sentences. Tested contractible words bolded. Tested monosyllabic word underlined.

1. *Ziji xian yanjiu xiazi zai qu zhao renjia.*

‘Do some research yourself before asking for help from others.’

2. *Dongtian nuanhe de shihou you shi dao shi'er du.*

‘It can be 10-12 degrees when it is warmer in winter.’

3. *Mashang jiushi nongli zhengyue shi'er.*

‘It will be Lunar January 12 soon.’

4. *Zhengyue li chichi shuashua benlai manhaode.*

‘It is really nice to just eat and enjoy the Lunar January.’

5. *Shi'er hao shang yingye bu ban xinyong ka song bingxiang.*

‘(They’ll) give free fridge if you open a credit card in the Business Department (of the bank) on the 12th.’

6. **Yingye** bu song **bingxiang** wo bu xiangxin.

‘I don’t believe the Business Department gives free fridge.’

7. **Xinyong** ka taiduo dui **xinyong** jilu buhao.

‘Having too many credit cards isn’t good for your credit history.’

8. Na lipin **zhiyou** shang **yingye** bu.

‘You’ll need to go to the Business Department to get the free gift.’

9. Ta jingran shua **renjia** de **xinyong** ka.

‘He dares to steal others’ credit card.’

10. **Renjia** jige **pengyou** lian fangzi zong **yiyang** de.

‘Those friends even have the same house style.’

11. Gaozi **zhiyao** pianyi bu chou mei de ren mai.

‘You don’t need to worry about sales if things are cheap enough.’

12. **Bingxiang** mashang ye buhao mai le, fanzheng.

‘Fridge sale is going to tough, anyway.’

13. **Yiyang** de huanjing hai buru zhuzai xiangxia.

‘I’d rather live in the rural area if the house (condition) is the same.’

14. **Jingyan** jiushi zui pianyi de shihou shi **zhengyue**.

‘(My) experience is that (it is) cheapest in January.’

15. **Benlai** zhege shi wo jiu **zhiyou** jiaoxun meiyou **jingyan**.

‘For this matter, I only have lessons but no experience.’

16. **Zhiyou** zhehui’er cai nong de xin gaozi wo hai **buhui**.

‘The only things I can’t do are those that are made just now.’

17. **Buhui** shuohua de ren **jinhou** bu keneng you **yiyang** de chenggong.

‘Those who don’t know how to communicate well will find it impossible to have the same achievements.’

18. Fanzheng zhechang chehuo shuoming **jingyan** zhongyao le.

‘Anyway, this car accident proves the importance of (driving) experience.’

19. **Zhiyao** fangzi zhiliang hao zai xiangxia ye **bu’ai**.

‘If the house is good quality, it doesn’t matter if it is in the rural area.’

20. **Bu’ai bu’ai**, wo zhege ren jiushi xihuan chu **pengyou**.

‘Never mind, I like making friends anyway.’

21. **Pengyou** quan de ren zongzai fa bainian de zhaopian.

‘People on WeChat keep posting photos of them paying New Year visits.’

22. **Xinnei** jiu xiang a, zheyangzi fanzheng budui a.

‘I was like, this is not right anyway.’

23. **Jinhou** yinhang mima jibude ye **bu’ai**.

‘From now on, it won't matter if (you) forget your bank account passwords.’

24. Ni **xinnei buhui** zhende zheme xiang ba?

‘You don’t really think this way, do you?’

25. Shaonian er **zhiyao** haodian ba yanguang fang zai **jinhou**.

‘The young people only need to look well ahead.’

26. Zhege yishang **zhengyue** chuan **benlai** jiu xian hou.

‘This garment is too hot for (this) winter anyhow.’

27. Sou le bantian taobao ye buceng soudao yizhang hao piao.

‘After searching Taobao for half a day, (I still) didn’t find any good ticket.’

28. Que de jizhang piao fandao zongshi weizhi haode.

‘The several missing tickets have good locations.’

29. Fan jijie de yishang buyao bai waitou wang li bai.

‘Don't put the out-of-season clothes outside; move them inside.’

30. Yang ayi nage jian lingkou de yishang xiang chuande youdian fan.

‘Auntie Yang’s heart-shaped collar shirt looks inside out a bit.’

31. Pian wo kai yinhangka bu rongyi de a.

‘It is not easy to fool me into opening a credit card.’

32. Piao fanzi zong hen congming buhao pian.

‘The ticket dealers are all clever and won’t be easily fooled.’

33. Tui diao zhiqian you buceng xian zai wangshang haohao sousou.

‘You didn’t do a good online research before you returned it.’

34. Song dian wan a shaor de wo hai xiangxin.

‘I’d believe if they give free spoons or something.’

35. Shaor a wan de wojia you bu que.

‘We don’t need spoons or bowls.’

36. Wojia Li Yang hai qu tui a yitao wan he shaor.

‘My Li Yang just returned a spoon and bowl set.’

37. Bainian bei, ni bu gaoxing jiu buyao qu bai.

‘As for New Year visits, don’t go if you don’t want to.’

38. Re a huo hai bu xiaode **zhiyou** Li Yang.

‘Only Li Yang would stir things up and not realize it at all.’

39. Naxie jian sangzi de ren ni buyao re.

‘Stay away from those loud talkers.’

40. Nage guniang kanqilai jiu xiong.

‘That girl looks competent.’

41. Zai jiali shuohua shengyin buyao zheme jian.

‘Don’t talk so loud at home.’

42. Bao le jiu buyao chi le, wen xiazi cai ge neng tui.

‘Don't eat if you're full; ask the waiter if we can return it.’

43. Qiu ren zhe'a qiu ye tai re pa le.

‘It is really terrible to prank like this.’

44. Bai lianshang pian ren, ni jiu bupa ren qiu?

‘Don't you worry about revenge if you just fool people so apparently?’

45. Xiang shi xiang de, jiushi nadian ye cai pade chibubao.

‘It does smell good, but just those leaves are probably too little to feed (us).’

46. Ye cai bi rou duo jiushi zenme chi zong bu xiang.

‘It is just not tasty if there's more veggies than meat.’

47. Hou lianpi bao a hai chi de nageren xing Ye.

‘That person who overeats impudently is Ye.’

48. Xiong daoshi xiong de jiushi xian luosuo.

‘Competent as she is, she's just too talkative.’

Appendix F: Script of audio introduction to syllable contraction and task

接下来的部分需要用到一个概念，叫合音现象。合音现象就是比如说两个字的词用一个音节来说，三个字的词用两个音节来说。比如说，普通话里的“酱紫”其实是一个三个音节的词“这”“样”“子”。我们如皋话里其实也有这样的例子，比如说“zan.zi”，其实说的是“这”“样”“子”。我用几个例子说明一下。你问“你什么时候来啊？”，在说得快的情况下，可以说“nei.sen.sei.le.a”，仔细听的话，其实说的是“nei”“sen”“sei”“le.a”。“sen”说的是“sen.de”，“sei”说的是“si.hei”。如果你回答“我马上来啊”，说得快的话会说“ngo.man.le.a”。“man”这一个音节其实代表的是“ma.san”。还有比如说，“xian.ha”可以说成“xia”。

接下来你会看到一些词，有些事两个字的，有的是三个字的。如果是两个字的，想想怎么能把它合成一个音节，如果是三个字的，想想怎么能把它合成两个音节，但是意思都不变。合成的词可能听起来比较奇怪或者没有办法用一个字来表达，这些都没有关系。

‘In the next section, you will need a concept called syllable contraction. Syllable contraction means, for instance, a two-character word said in one syllable, and a three-character word said in two syllables. For example, in Mandarin, “jiang.zi” is actually a three-syllable word “zhe.yang.zi”. Our Rugao also has this kind of examples, such as “zan.zi”, which in fact is “za.yan.zi”. I’ll use an example as an illustration. When you ask “when will you come?” and speak fast, you can say “nei.sen.sei.le.a”. If you listen carefully, it is actually “nei” “sen” “sei” “le.a”. “sen” is “sen.de”, and “sei” is “si.hei”. If your answer is “I will come right away!” and you speak fast, you will say “ngo.man.le.a”. The one syllable “man” represents “ma.san” (‘right away’). Another example is that “xian.ha” can be pronounced “xia”.

Next, you will see some words, some of which are two-character words and some are three-character words. If it is a two-character word, think about how to contract it to one syllable. If it is a three-character word, think about how to contract it to two syllables. The meaning should not change. The contracted word may sound strange or cannot be represented by a character, but these won’t matter.’

Appendix G: Full list of words in the Rugao Syllable Contraction Corpus

Word (pinyin)	Chinese & English translation	Pre-contraction	Contracted
banyewu	办业务 (do business)	pjen + iʔ + vuu	pej + vuu
benlai	本来 (originally)	pəŋ + lɛ	pɛɛ
bijiao	比较 (comparatively)	pi + tɕjo	pjo
bingxiang	冰箱 (fridge)	pjəŋ + ɕjæn	pɛæn/pjæn
biru	比如 (for example)	pi + zu	piw

bixiazi	比下子 (compare a bit)	pi + xa + tsi	piə + tsi
biye	毕业 (graduate)	pjəʔ + iʔ	piʔ/pii
bu'ai	不挨(not be)	pəʔ + ɲɛ	pɛɛ
buceng	不曾 (did/has not)	pəʔ + tshəŋ	pəŋ
buhao	不好 (not good)	pəʔ + xo	po
buran	不然 (otherwise)	pəʔ + zɿn	pzɿn
bushi	不是 (not)	pəʔ + si	pɿi
buxuyao	不需要 (no need)	pəʔ + ɕy + jo	pəʔ + ɕjo
buyao	不要 (not)	pəʔ + jo	pjo
buye	不也 (also)	pəʔ + ja	pja
buyeyou	不也有 (also have)	pəʔ + ja + jow	pja + jow
buyiyang	不一样 (not same)	pəʔ + jəʔ + jæn	pɛæn/pjæn
chehuo	车祸 (car accident)	tsha + xow	tshaw
chishenme	吃什的 (eat what)	tɕjəʔ + səŋ + dej	tɕjən + nej
daikuan	贷款 (loan/mortgage)	thɛ + khun	thɛun
dajia	大家 (everyone)	ta + ka	taa
danshi	但是 (but)	tjɛŋ + si	tɛi
diyi	第一(first)	thi + jəʔ	thjəʔ/thjə
duochang	多长 (how long)	tow + tsæn	twæn
erqie	而且 (furthermore)	aɪ + tɕhja	aiə
ershíwu	二十五 (twenty-five)	aɪ + səʔ + vu	aiə + vu
fan'er	反而 (on the contrary)	fɛn + aɪ	fɛɪ
fangxiang	方向(direction)	fæn + ɕjæn	fæn
fangxue	放学 (after school)	fæn + xəʔ	fəʔ
fanyi	翻译 (translate)	fɛn + iʔ	fɛʔ
fanzheng	反正 (anyway)	fɛŋ + tsəŋ	fɛŋ
fuwuyuan	服务员 (servant)	fɔʔ + vu + jun	fɔv + jun
gaosu	告诉 (tell)	kɔ + sɔʔ	kɔɔ/kɔə/kɔɪ
gaosuniting	告诉你听 (tell you)	kɔ + sɔʔ + nej + thjəŋ	kɔn + thəŋ
gege	哥哥(elder brother)	kow + kow	kow
geye	个也/可也(isn' t it)	kow + ja	kɛa
haihao	还好 (OK)	xa + xo	xao
haishi	还是(still)	xa + si	xai
haoxiang	好像 (seems)	xo + tɕhjaen	xwæn
jiaode	教得 (teach)	kɔ + dəʔ	kɔə
jiaqu	家去 (go home)	ka + tɕhy	kay/kai
jide	记得 (remember)	tɕi + təʔ	tɕiəʔ
jingyan	经验 (experience)	tɕəŋ + in	tɕin
jinnianzi	今年子 (this year)	kəŋ + nin + tsi	kənnts

jiushi	就是 (be)	təhju + si	təhqi
juede	觉得 (think/feel)	kæʔ + tshɛ	kææ/kæɛ
kanjian	看见 (see)	khun + tein	khwin
keshi'a	个是啊 (is it?)	kow + si + a	ka
laoshi	老师 (teacher)	lo + si	loi
liuxialai	留下来 (stay)	lej + xa + le	lɛa + le
lihai	厉害 (good)	li + xɛ	lie
manhaochide	蛮好吃的 (pretty tasty)	mjen + xo + təhəʔ + tej	mɛo + təhət + tei
manhaoting	蛮好听 (pretty pleasant)	mjen + xo + tjən	mjeo + tjən
mashang	马上 (right away)	ma + sən	mæn
mouge	某个 (some CL)	mow + kow	mow
muogu	蘑菇 (mushroom)	muu + ku	muuu
nage	那个 (that CL)	now + kow	now or nuw
nage	那个 (that CL)	la + kow	law/low
nali	那块 (there)	low + kwɛ	lwe
naxie	那些 (those)	low + ɕja	lwa
naye	那也 (that also)	now + ja	naa/nwa
nimen	你徠 (you-PL)	nej + le	nɛɛ/njɛ
nizai	你在 (you at)	nej + təhɛ	nɛɛ
nuanhuo	暖和 (warm)	nun + xow	now
peiyang	培养 (foster)	phɛj + jæn	phjæn
pengyou	朋友 (friend)	phən + jow	phiw/phjow/
pengyou	朋友 (friend)	phən + jow	phɔw
pianyi	便宜 (cheap)	phin + i	phii
pigu	屁股 (bottom)	phi + kow	phiw
qiao'er	巧儿 (Qiao'er (name))	təhjo + ai	təhjoɪ
qiongru	穷如 (just like)	təhjoŋ + zu	təhjoɪ
qishi	其实 (actually)	təhi + səʔ	təhəʔ
quanbu	全部 (all)	təhun + phuu	təhum
quguo	去过 (have been)	təhy + kow	təhyw
ranhou	然后 (then)	zɪŋ + xej	zii/zɛj
renjia	人家 (they)	zən + ka	zɛa
renjia	你家 (your)	nej + ka	nja
Rudong	如东 (Rudong (city name))	zu + tən	zɔwən
Rugao	如皋 (Rugao (city name))	zu + kə	zɔwə
ruguo	如果 (if)	zu + kow	zuw
shenghuo	生活 (life)	sən + xuʔ	səw

shenmegaozi	什的稿子(what)	səŋ + tei + kɔ + tsi	sŋ + kɔ + tsi
shenxian	神仙(god/goddess)	sən + ein	sin
shi'er	十二(twelve)	səʔ + aɪ	sai
shihou	时候(time)	si + xej	sej
shiji	世纪(century)	si + tɛi	sii
shijian	时间(time)	si + tɛjən	sjən
shijie	世界(world)	si + tɛhjɛ	sɛ
shijie	世界(world)	si + kɛ	sɛ
shiyang	式样(style)	səʔ + jæn	sjæn
shiying	适应(adapt)	səʔ + jəŋ	siŋ
shouxian	首先(first)	sej + ein	sein/sin
shuibuzhao	睡不着(can't fall asleep)	ʂuj + pəʔ + tɛhæʔ	ʂuib + tɛhæʔ
shuju	数据(data)	suu + tɛu	suu
si'ren	死人(damn)	si + ɪəŋ	səŋ/siən
tahaizai	他还在(he is still)	tha + xa + tshɛ	thaa + tshɛ
tingbudong	听不懂(not understand)	thjəŋ + pəʔ + tɔŋ	thjəm + tɔŋ
tou Shang	头上(on the head)	thɛj + sən	thəŋ
weishenme	为什么(why)	vej + səŋ + məɪ	vejɪ + məɪ/vəmməɪ
wojia	我家(my)	ŋwa + ka	ŋwa
wolai	我们(we)	ŋow + lɛ	ŋwɛ
woye	我也(I also)	ŋow + ja	ŋwa
xiangxia	乡下(countryside)	ɛjæn + xa	ɛja/ɛaa
xiangxin	相信(believe)	ɛjæn + ɛjəŋ	ɛjəŋ
xianzai	现在(now)	ein + tshɛ	ɛjɛ/ɛje
xiaoxue	小学(primary school)	ɛjo + ɛjæʔ	ɛyæ
xihuan	喜欢(like)	ɛi + xun	ɛun
xinyong	信用(credit)	ɛjəŋ + jəŋ	ɛjəŋ
xixian	你先(you first)	nej + ein	nein/nen
xuexiao	学校(school)	ɛæʔ + ɛjo	ɛjo
yanjing	眼睛(eye)	ŋɛn + tɛjəŋ	ŋɛiəŋ
yihou	以后(later)	ji + xej	jij
yijing	已经(already)	i + tɛən	iən
yinggai	应该(should)	jəŋ + kɛ	jɛɛ
yingxiang	影响(affect)	jəŋ + ɛjæn	jæn
yingyebu	营业部(business depart.)	jən + iʔ + pɰu	jiʔ + pɰu
yinhang	银行(bank)	jən + xæn	jæn
yinxiang	印象(impression)	jən + ɛjæn	jæn
yiyang	一样(same)	jəʔ + jæn	jæn
youxiao	有效(effective)	jow + ɛjo	jwo

youxie	有些 (some)	jow + ɕja	ya
zenmeban	怎啊办 (how to do)	tsæn + a + pen	tsaa + pen
zenmeban	怎啊办 (how to do)	tsəŋ + a + pen	tsaa + pen
zenyang	怎样 (how)	tsəŋ + jæn	tsæn
zheli	这里 (here)	tʂa + khwɛ	tsæ/tse
zhege	这个 (this CL)	tʂa + kow	tsaw
zhengyue	正月 (lunar first month)	tsəŋ + ju?	tsju?
zhexiazi	这下子 (as a result)	tsəʔ + xa + tsi	tsaats/tsats
zhexie	这些 (these)	tsəʔ + xa	tsaa
zheyang	这样 (so)	tsəʔ + jæn	tsæn
zheyangzi	这样子 (this way)	tsəʔ + jæn + tsi	tsents
zheyeshi	这也是 (this is)	tsəʔ + ja + si	tʂea + si/tsaa + si
zhiyao	只要 (only)	tsəʔ + jo	tsjo
zhiyou	只有 (only)	tsəʔ + jow	tsjow/tsew
zhuyao	主要 (mostly)	tsu + jo	tswo
ziji	自己 (oneself)	tʂhi + tɕi	tʂhi
zijia	自家 (one's own)	tʂhi + ka	tʂhaa/tʂhia
ziyou	自由 (free)	tʂhi + jow	tʂhjow
zuoshenme	做什么的 (do what)	tsow + səŋ + dej	tsun + nej

REFERENCES

REFERENCES

- Audacity, T. (2008). Audacity (Version 1.3.4-beta). Retrieved from <http://audacityteam.org/>
- Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer. Retrieved from <http://www.praat.org/>
- Cai, H. (2011). *Yancheng fangyan yanjiu [A study on Yancheng dialect]*. Zhong Hua Shu Ju.
- Casali, R. F. (1998). *Resolving Hiatus*. Garland Publishing.
- Chao, Y.-R. (1927). Lia, sa, si'e, ba'e [Two, three, four and eight]. *Dongfang Zazhi [Oriental Magazine]*, 12.
- Chao, Y.-R. (1968). *A Grammar of Spoken Chinese*. University of California Press.
- Cheng, C., & Xu, Y. (2009). Extreme reductions: Contraction of disyllables into monosyllables in Taiwan Mandarin. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*, 456–459.
- Chung, K. S. (2006). Contraction and Backgrounding in Taiwan Mandarin. *Concentric: Studies in Linguistics*, 1(January), 69–88.
- Chung, R.-F. (1996). *The segmental phonology of Southern Min in Taiwan*. Crane Publishing.
- Chung, R.-F. (1997). Syllable contraction in Chinese. In F. Tsao & S. Wang (Eds.), *Chinese Languages and Linguistics III: Morphology and Lexicon* (pp. 199–235). Academia Sinica.
- Crosswhite, K. M. (2000). Sonority-driven Reduction. *Proceedings of the Twenty-Sixth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Aspect*, 77–88.
- Davidson, L. (2007). The relationship between the perception of non-native phonotactics and loanword adaptation. *Phonology*, 24(2), 261–286.
<https://doi.org/10.1017/S0952675707001200>
- de Lacy, P. (2010). The interaction of tone, sonority, and prosodic structure. In P. de Lacy (Ed.), *The Cambridge Handbook of Phonology* (pp. 281–308). Cambridge University Press.
<https://doi.org/10.1017/CBO9780511486371.013>
- Dmitrieva, O., Jongman, A., & Sereno, J. (2010). Phonological neutralization by native and non-native speakers: The case of Russian final devoicing. *Journal of Phonetics*, 38(3), 483–492.
<https://doi.org/10.1016/j.wocn.2010.06.001>
- Duanmu, S. (1990). *A formal study of syllable, tone, stress and domain in Chinese languages* [Unpublished doctoral dissertation]. Massachusetts Institute of Technology.

- Dupoux, E., Hirose, Y., Kakehi, K., Pallier, C., & Mehler, J. (1999). Epenthetic vowels in Japanese: A perceptual illusion? *Journal of Experimental Psychology: Human Perception and Performance*, 25(6), 1568–1578. <https://doi.org/10.1037/0096-1523.25.6.1568>
- Ellis, L., & Hardcastle, W. J. (2002). Categorical and gradient properties of assimilation in alveolar to velar sequences: Evidence from EPG and EMA data. *Journal of Phonetics*, 30(3), 373–396. <https://doi.org/10.1006/jpho.2001.0162>
- Gordon, M., Ghushchyan, E., McDonnell, B., Rosenblum, D., & Shaw, P. A. (2012). Sonority and central vowels: A cross-linguistic phonetic study. *The Sonority Controversy*, (Vaux 1998), 219–256. <https://doi.org/10.1515/9783110261523.219>
- Harford, C. (1997). When Two Vowels Go Walking: Vowel Coalescence in Shona. *Zambezia*, 24(1), 69–85. Retrieved from <http://pdfproc.lib.msu.edu/?file=/DMC/AfricanJournals/pdfs/Utafiti/s4NS/aejps004NS008.pdf>
- Hsiao, Y. E. (1995). *Violable phonotactics in syllable contraction : A corpus-based study* [Unpublished working paper].
- Hsu, H. C. (2003). A sonority model of syllable contraction in Taiwanese Southern Min. *Journal of East Asian Linguistics*, 12(4), 349–377. <https://doi.org/10.1023/A:1026108613211>
- Huang, T. (2011). Contextual and Pitch Range Effects on Tonal Realizations in. *Journal of Chinese Linguistics*, 149–176.
- Kager, R. (1999). *Optimality theory*. Cambridge University Press.
- Kenstowicz, M. J. (1994). *Phonology in generative grammar* (7th ed.). Blackwell Publishing.
- Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, 10(3), 421–441. <https://doi.org/10.2307/4178120>
- Kuo, G. C. (2010). Production and perception of Taiwan Mandarin syllable contraction. *The Journal of the Acoustical Society of America*, 127(3), 1954–1954. <https://doi.org/10.1121/1.3384952>
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Blackwell Publishing.
- Li, C. N., & Thompson, S. A. (1989). *Mandarin Chinese: A Functional Reference Grammar*. University of California Press.
- Li, D. C.-H. (2011). Vowel coupling in Mandarin syllable contraction. *Proceedings of ICPhS XVII*, 1206–1209.
- Li, R. (1989). Hanyu fangyan de fenqu [The dialect areas of Chinese]. *Fangyan [Dialect]*, 4.
- Lin, H. (1995). Yinjie de suojian? Suojian de yinjie? [Reduced syllable count? Reduced

- syllable?]. In *Proceedings of the second international linguistics symposium of Taiwan*. Institute of Linguistics of National Taiwan University.
- Lin, Y.-H. (2007). *The sounds of Chinese*. Cambridge University Press.
- Lü, S. (1984). Shi nin, an, zan, da, fulun men zi [Explanation of nin, an, zan, and men]. In *Hanyu yufa lunwenji [Papers of Chinese grammar]*. The Commercial Press.
- Manaster Ramer, A. (1996). A letter from an incompletely neutral phonologist. *Journal of Phonetics*, 24(4), 477–489. <https://doi.org/10.1006/jpho.1996.0026>
- Marslen-Wilson William. (1973). Linguistic Structure and Speech Shadowing at Very Short Latencies. *Nature*, (244), 522–523. Retrieved from <https://doi.org/10.1038/244522a0>
- McCarthy, J. J., & Prince, A. S. (1994). Prosodic Morphology I: Constraint Interaction and Satisfaction. *Yearbook of Morphology 1993*, (January), 79–153.
- McCollum, A. (2019). Gradient morphophonology: Evidence from Uyghur vowel harmony. *Proceedings of the Annual Meetings on Phonology*, 7, 1–12. <https://doi.org/10.3765/amp.v7i0.4565>
- Mermelstein, P. (1975). Automatic segmentation of speech into syllables. *Journal of the Acoustics Society of America*, 58(4), 880–883.
- Myers, J., & Li, Y. (2009). Lexical frequency effects in Taiwan Southern Min syllable contraction. *Journal of Phonetics*, 37(2), 212–230. <https://doi.org/10.1016/j.wocn.2009.02.002>
- Ohala, J. J. (1990). There is no interface between phonology and phonetics: a personal view. *Journal of Phonetics*, 18(2), 153–171. [https://doi.org/10.1016/s0095-4470\(19\)30399-7](https://doi.org/10.1016/s0095-4470(19)30399-7)
- Ohala, J. J., & Kawasaki-Fukumori, H. (2011). Alternatives to the sonority hierarchy for explaining segmental sequential constraints. In *Language and its Ecology: Essays in Memory of Einar Haugen* (pp. 343–365). <https://doi.org/10.1515/9783110805369.343>
- Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of Phonetics*, 36(1), 55–90. <https://doi.org/10.1016/j.wocn.2007.09.003>
- Parker, S. (2011). Sonority. In M. Van Oostendorp, C. J. Ewen, E. Hume, & K. Rice (Eds.), *The Blackwell Companion to Phonology*. Blackwell Publishing.
- Parker, S. (2012). *The Sonority Controversy*. (S. Parker, Ed.), *The Sonority Controversy*. De Gruyter. <https://doi.org/10.1515/9783110261523>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>

- Port, Robert F. & O'Dell, M. (1985). Neutralization of syllable-final voicing in German. *Journal of Phonetics*, 15, 455–471.
- Port, R. F., & Leary, A. P. (2005). Against formal phonology. *Language*, 927–964.
- Selkirk, E. (1984). On the major class features and syllable theory. In *Language sound structure* (Vol. 5079, pp. 107–136). <https://doi.org/10.1037//0003-066X.46.5.506>
- Shih, S. (2016). Sonority-Driven Stress does not Exist. *Proceedings of the Annual Meetings on Phonology*, 3, 1–11. <https://doi.org/10.3765/amp.v3i0.3666>
- Sun, H. (2014). Lun Hanyu Heyin Xianxiang Yanjiu (A study on Chinese syllable fusion research). *Journal of Southeast University (Philosophy and Social Science)*, 40(1).
- Ting, P.-H. (1966). Rugao fangyan de yinyun [Phonology of the Rugao dialect]. *Bulletin of the Institute of History and Philology, Academic Sinica*, 36, 573–633.
- Tseng, S.-C. (2005a). Contracted syllables in Mandarin: Evidence from spontaneous conversations. *Language and Linguistics*, 6(1), 153.
- Tseng, S.-C. (2005b). Syllable contractions in Mandarin Conversational Dialogue Corpus. *International Journal of Corpus Linguistics*, 10(1), 63–83.
- Urbanczyk, S. (2006). Reduplicative form and the Root-Affix Asymmetry. *Natural Language & Linguistic Theory*, 24(1), 179–240. <https://doi.org/10.1007/s11049-005-4373-x>
- Warner, N., Jongman, A., Sereno, J., & Kemps, R. (2004). Incomplete neutralization and other sub-phonemic durational differences in production and perception: Evidence from Dutch. *Journal of Phonetics*, 32(2), 251–276. [https://doi.org/10.1016/S0095-4470\(03\)00032-9](https://doi.org/10.1016/S0095-4470(03)00032-9)
- Wee, L. H. (2014). Casual speech elision and tone sandhi in Tianjin trisyllabic sequences. *International Journal of Chinese Linguistics*, 1(1), 71–95. <https://doi.org/10.1075/ijchl.1.1.03wee>
- Wong, W. Y. P. (2006). *Syllable fusion in Hong Kong Cantonese connected speech* [Unpublished doctoral dissertation]. The Ohio State University.
- Wu, F. (2006). *Rugao fangyan yanjiu [Study on Rugao dialect]*. Zhongguo Wenlian Chubanshe.
- Xu, C. (2014). An OT analysis on syllable contraction in Jianghuai Chinese [Unpublished working paper].
- Xu, C., Lin, Y.-H., & Durvasula, K. (2018). Sonority bias in Rugao di-syllabic syllable contraction. *Proceedings of the Linguistic Society of America*, 3(1), 28. <https://doi.org/10.3765/plsa.v3i1.4316>
- Xu, C., & Mao, L. (2017). The sociolinguistic meanings of syllable contraction in Chinese. *Asia-Pacific Language Variation*, 3(2), 160–199. <https://doi.org/10.1075/aplv.16004.xu>

- Yip, M. (1988). Template morphology and the direction of association. *Natural Language and Linguistic Theory*, 6(4), 551–577. <https://doi.org/10.1007/BF00134493>
- Zec, D. (1995). Sonority constraints on syllable structure. *Phonology*, 12(1), 85–129. <https://doi.org/10.1017/S0952675700002396>