VOWELS AND TONES IN MEIXIAN HAKKA: AN ACOUSTIC AND PERCEPTUAL STUDY

CHEUNG YUK MAN

DOCTOR OF PHILOSOPHY

CITY UNIVERSITY OF HONG KONG

APRIL 2011

CITY UNIVERSITY OF HONG KONG
香港城市大學

Vowels and Tones in Meixian Hakka: An Acoustic and Perceptual Study
梅縣客家話的元音及聲調: 聲學及感知研究

Submitted to
Department of Chinese, Translation and Linguistics
中文、翻譯及語言學系
in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy
哲學博士學位

by

Cheung Yuk Man
張玉敏

April 2011
二 零 一 一 年 四 月

ABSTRACT

This dissertation presents a phonetic description of acoustic and perceptual properties of tones as well as acoustic properties of vowels in Meixian Hakka. The vowel and tone systems of Meixian Hakka have been described in the past studies. However, the basis of such descriptions is mainly individual impression and there is a considerable discrepancy in transcription among linguists. Due to the lack of experimental data presented in the literature, it is worthwhile to investigate the sound system of Meixian Hakka in acoustic and perceptual perspectives.

The F_0 patterns of the citation tones in Meixian Hakka were analyzed based on the raw and normalized data. It shows that there are acoustically six citation tones in Meixian Hakka. Yin Ping is a mid level tone and is transcribed as [33]; Yang Ping is a slight falling tone at the low frequency range and is transcribed as [11]; Shang Sheng is a mid-high to low falling tone and is transcribed as [41]; Qu Sheng is a high to low falling tone with a delayed falling contour and is transcribed as [51]; Yin Ru is a mid-high to low falling short tone and is transcribed as [41]; Yang Ru is a high level short tone and is transcribed as [55 or 5]. In spite of duration, the overall contour shape of Shang Sheng resembles that of Yin Ru. The two tones have a simple falling tone with an onset at the middle frequency range and an offset at the low frequency range.

The comparison of F_0 patterns of the citation tones on vowels [i, a, u] indicates that high vowels such as [i] and [u] tend to have a higher intrinsic fundamental frequency (iF0) than low vowels such as [a]. The correlation between F_0 and vowel height is observed, not only in Meixian Hakka, but also in languages from different language families as described in the past studies and such correlation is held to be linguistic universal. There is a tendency for tones in CV(C) syllables to begin at a higher frequency than tones in V(C) syllables. Such difference in F_0 pattern between CV(C) syllables and V(C) syllables may be associated with the effect of neighboring consonants on the F_0 onsets of the following tones.

Tone sandhi rules in Meixian Hakka are formulated in this research. The rules are to be compared with the tone sandhi rules in the past studies. The demonstration of F_0 patterns of tones in two-tone combinations has shown that there are three sandhi rules in Meixian Hakka and they are as follows: (1) Yin Ping [33] is realized as a mid to high rising tone [35] when it is followed by [11, 41, 41]; (2) Shang Sheng [41] is realized as a mid level tone [33] when it is followed by [11, 41, 41]; and (3) Qu Sheng [51] is

realized as a high level tone [55] when the second syllable is superimposed by [11, 41, 51, 41]. Two sandhi forms from Yin Ping [33] and Qu Sheng [51] that are different from the six citation tones in Meixian Hakka are found. They are correspondingly the high rising tone [35] and the high level tone [55]. The tone sandhi patterns observed in this study are consistent with the patterns of Lü (2006) which were based on the acoustic data, but not with other past studies which were mainly based on individual impression.

In addition to the acoustic characteristics of tones, the perceptual properties of tones and the correlation between perceptual and acoustic characteristics of tones in Meixian Hakka are evaluated. The potency of tonal height, slope, F_0 onset and F_0 offset in the identification of the four long citation tones has been examined. It shows that there is a strong correlation between tone perception and the acoustic characteristics of tones in Meixian Hakka. Tonal height plays an important role in the identification of the two level tones [33, 11]. A high F_0 onset is an important cue for the identification of [51]. Slope is potent for the distinction between the three tones [51, 41, 11]. Stimuli with a steep falling contour are highly identified as [51], whereas stimuli with a less steep falling contour are identified as [41]. Stimuli are identified as [11] when the slope of falling is small. The perceptual results of this study show that not only slope but also tonal height are important for the distinction between [51] and [41].

The acoustic characteristics and temporal organization of monophthongs, diphthongs and triphthongs have been examined in Meixian Hakka. There are five Meixian Hakka vowels in the open syllable and they are transcribed as [i, e, a, o, u]. Results indicate that the relative distance between mid and high vowels differs with regard to vowel backness and speaker gender. The diphthongs in Meixian Hakka are suggested to be divided into two categories according to the degree of resemblance between formant patterns of the diphthong elements and formant patterns of their corresponding target vowels in the acoustic space. Such claim is supported by the temporal organization of diphthong. The results indicate that the dynamic nature of diphthong in terms of F_2 range of change is sufficient for the distinction among diphthongs in a language such that the dynamic properties might be language-universal. In consideration of the phonetics and phonology of Meixian Hakka, it is proposed that there are 9 Meixian Hakka diphthongs and their narrow transcriptions are [ie, ia, io, iu, aɪ, ɔɪ, ʊi, ao, ɛʊ]. Besides, it is suggested that there are two triphthongs in Meixian Hakka and they are [iʊ̈i] and [iao].

ACKNOWLEDGEMENTS

I owe my greatest thanks to all those who made it possible for me to undertake this research and complete my thesis. I would like to express my sincerest thanks to my supervisor, Professor Jonathan J. Webster, for his encouragement and continuous support in all possible ways. My grateful thanks also go to my examiners, Dr. Paul Law, Professor K. K. Luke and Professor Yi Xu, for their insightful comments and constructive suggestions. I am very thankful to all the teachers who gave me advice, encouragement and support during my post-graduate studies.

I would also like to take this opportunity to thank my dear colleagues and friends. I appreciate their continued friendship, assistance and support in the completion of my studies. Last but not least, I am extremely indebted to my family, especially my parents, who offered me endless support, caring, love and understanding throughout my doctoral journey. Without their unconditioned love and strong backup, I would have never been able to accomplish my PhD degree.

TABLE OF CONTENTS

TABLES

CHAPTER 1

CHAPTER 2

CHAPTER 4

FIGURES

CHAPTER 1

CHAPTER 2

CHAPTER 3

CHAPTER 4

Appendix A

Appendix B

CHAPTER ONE

INTRODUCTION

1.1 General information about the Hakka dialects

The Chinese language forms one of the two branches of the Sino-Tibetan family of languages which is the most extensive language family in East Asia (Gleason, 1961; Thurgood, 2003). Many languages in the Sino-Tibetan language family are tonal and Chinese dialects are well known for their complicated tone systems. There are seven main dialect groups that are traditionally recognized in the Chinese language and they are Mandarin (官话), Wu (吴语), Cantonese/Yue (粤语), Min (闽语), Xiang (湘语), Hakka (客家话) and Gan (赣语) (Yuan et al., 2001). The dialects are mutually unintelligible in general across dialect groups. Mandarin, one of the seven dialect groups, has been spoken predominantly as the mother tongue in northern China and it is widely spoken throughout the whole of China. The Beijing dialect is the most well-known among the dialects of Mandarin and has been typically viewed as a prime example. This may be due to the fact that Putonghua, which is the national standard language of China, is based on the phonology, lexicon and grammar of Beijing Mandarin. Figure 1a displays the geographic distribution of different dialect groups in China (A detailed dialect map can be found in Appendix A). As can be seen, the dialects of Hakka are spoken predominantly in southern China. Most Hakka live in southern China, especially in the provinces of Guangdong, southwestern Fujian, southern Jiangxi, part of Hunan, Sichuan, Guangxi and Taiwan, etc. (Yuan et al., 2001; Wen, 2006) In addition to the ethnic groups in China, Hakka dialects are also spoken in several Southeast Asian countries, such as Indonesia, Burma, Malaysia, Singapore, Thailand, Vietnam, Philippines and India, etc. (Norman, 1988; Wen, 2006) The speech data obtained in this study was produced by the native speakers of Meixian Hakka. Meixian

Figure 1a Geographic distribution of the dialect groups in China
Retrieved June 25, 2009 from http://en.wikipedia.org/wiki/Chinese_ languages

Hakka is the representative dialect of the dialect group of Hakka. Meizhou City (梅州市), located in the northeast of the Guangdong Province, is one of the largest agglomerations for Hakkas. The map of the Meizhou City is shown in Figure 1b. As can

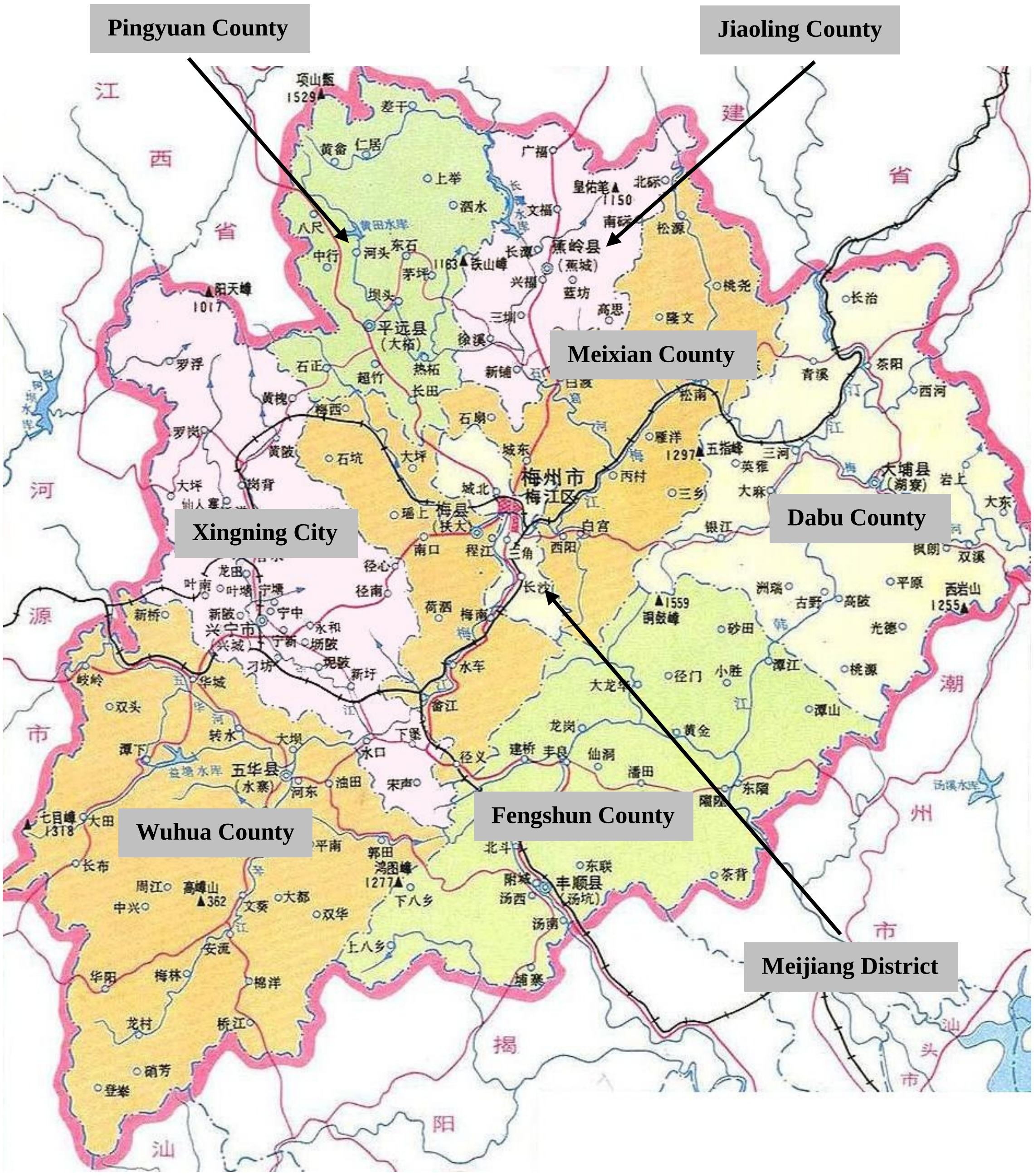


Figure 1b The map of the Meizhou City (梅州市). The counties of Meixian (梅县), Dabu (大埔县), Fengshun (丰顺县), Wuhua (五华县), Pingyuan (平远县), Jiaoling (蕉岭县), the city of Xingning (兴宁市) and the district of Meijiang (梅江区) were labelled.
Retrieved June 25, 2009 from http://www.lvyou114.com/UploadFile/200511017851588.jpg

be seen, the Meizhou City comprises Meijiang District (梅江区), Meixian County (梅县), Pingyuan County (平远县), Jiaoling County (蕉岭县), Dabu County (大埔县), Fengshun County (丰顺县), Wuhua County (五华县) and Xingning City (兴宁市). Meijiang District is located in the centre of the county of Meixian. Generally speaking, the dialect of Meixian Hakka is interpreted as the Hakka dialect spoken in Meijiang District and Meixian County.

1.2 The sound system of Meixian Hakka

The sound system of Meixian Hakka has been described extensively in the past studies (Hashimoto, 1973; Hanyu Fangyin Zihui, 1989; Huang, 1992; 1995; Chen, 1993; Xie, 1994; Yuan et al, 2001; Wang, 2006). The phonemic descriptions of place of articulation, manner of articulation and number of consonants in Meixian Hakka are rather consistent among the majority of the past studies. It was stated that there were 17 consonants in Meixian Hakka and they were [p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v, h, m, n, ŋ, l] including a zero initial consonant ø (Huang, 1992; 1995; Xie, 1994; Yuan et al., 2001; Wen, 2006). The zero initial consonant has been traditionally considered as one of the consonants by the Chinese linguists. There were 16 Meixian Hakka consonants in the study of Hashimoto (1973) and [v], which has been described as a consonant in other studies, was recognized as a glide or a vowel in his study. On the other hand, Wang (1956) declared that the palatal consonant [ɲ] should be considered as one of the phonemes in the sound system of Meixian Hakka and there were a total of 18 consonants in the language. Chen (1993) has found much more consonants in Meixian Hakka than other linguists. Regardless of the zero initial consonant, there were 20 Meixian Hakka consonants in Chen's study. In addition to the consonants [p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v, h, m, n, ŋ, l], which have been listed in most of the past studies, three additional alveolo-palatal initial consonants [ɕ, tɕ, $tɕ^h$] were included in Chen's study. Table 1.1 displays the phonological descriptions of syllable initial consonants in Meixian Hakka described in the past studies. The zero initial consonant ø was included and it was agreed among the past studies that there are six stops [p, p^h, t, t^h, k, k^h], three fricatives [f, s, h], two alveolar affricates [ts, ts^h], three nasals [m, n, ŋ], and one lateral approximant [l] in Meixian Hakka. As described in the past studies, the consonants in Meixian Hakka undergo the process of palatalization when they are followed by a high

	consonants in Meixian Hakka (number of consonants)
Chen (1993)	[p, p^h, t, t^h, ts, ts^h, ɕ, tɕ, $tɕ^h$, k, k^h, s, f, v, h, m, n, ŋ, ɲ, l, ø] (21)
Hahsimoto (1973)	[p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, h, m, n, ŋ, l, ø] (16)
Huang (1992; 1995)	[p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v(ʋ), h, m, n, ŋ, l, ø] (17)
Wang (1956)	[p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v, h, m, n, ŋ, ɲ, l, ø] (18)
Wen (2006)	[p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v, h, m, n, ŋ, l, ø] (17)
Xie (1994)	[p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v, h, m, n, ŋ, l, ø] (17)
Yuan et al. (2001)	[p, p^h, t, t^h, ts, ts^h, k, k^h, s, f, v, h, m, n, ŋ, l, ø] (17)

Table 1.1 The initial consonants in Meixian Hakka that listed in the past studies. The total number of consonants including the zero initial consonant ø was shown in parentheses.

<table>
<tr><th></th><th>bilabial</th><th>labio-
dental</th><th>dental</th><th>alveolar</th><th>palatal</th><th>velar</th><th>glottal</th></tr>
<tr><td>plosive</td><td>p p^h</td><td></td><td colspan="2">t t^h</td><td>(c) (c^h)</td><td>k k^h</td><td></td></tr>
<tr><td>nasal</td><td>m</td><td></td><td colspan="2">n</td><td>(ɲ)</td><td>ŋ</td><td></td></tr>
<tr><td>fricative</td><td></td><td>f v</td><td colspan="2">s</td><td>(ç)</td><td>(x)</td><td>h</td></tr>
<tr><td>affricative</td><td></td><td></td><td colspan="2">ts ts^h</td><td></td><td></td><td></td></tr>
<tr><td>approximant</td><td></td><td>(ʋ)</td><td colspan="2"></td><td></td><td></td><td></td></tr>
<tr><td>lateral approximant</td><td></td><td></td><td colspan="2">l</td><td></td><td></td><td></td></tr>
</table>

Table 1.2 The consonant chart of Meixian Hakka consonants that have been described in the past studies. The chart contains all phonetic representations of the consonants in Meixian Hakka. IPA in parenthesis represents the allophone.

front vowel [i]. It was stated that palatalization assimilated the velar consonants [k, k^h, ŋ] and glottal consonant [h] to their alveolo-palatal or palatal counterparts. Xie (1994) declared that a velar nasal [ŋ] was realized as its palatal counterpart [ɲ] while a voiceless glottal fricative [h] was realized as a velar fricative [x] when they were followed by a high front vowel [i]. Descriptions of palatalization were also found in the studies of Huang (1992; 1995), Yuan et al. (2001) and Wen (2006). In their studies, the palatal nasal [ɲ] was considered as an allophone of the velar nasal [ŋ] and occurred before [i]. Instead of the phonetic realization of the velar fricative [x] as stated in Xie's study (1994), they declared that the glottal fricative [h] was realized as [ç] when it was in the palatalization environment. Similarly, they claimed that voiceless velar stops [k, k^h] were realized as palatal stops [c, c^h]. In the studies of Huang (1992; 1995), she stated that the so-called [v] was voiced and frictionless and it would be rational to consider the sound as a labiodental approximant [ʋ] instead. The phonetic representations of Meixian Hakka consonants that have been described in the past studies are shown in Table 1.2.

Meixian Hakka has a rich inventory of rimes. The total number of rimes in Meixian Hakka was found vary in the past studies. The Meixian Hakka vowels occur in the V, VC, CV and CVC syllables. In the V and CV syllables, Meixian Hakka has monophthongs, diphthongs and triphthongs. Table 1.3 displays the possible combinations of syllable-initial consonant and rime which form various syllable types in Meixian Hakka. The possible Meixian Hakka rimes are listed according to the descriptions in the studies of Xie (1994), Huang (1995), Yuan et al. (2001) and Wen (2006). As shown in Table 1.3, there are 69 rimes in Meixian Hakka which were commonly found by the dialectologists. Rimes in shading characters indicate rimes that only listed in the study of Yuan et al. (2001). Those in bold letters indicate rimes that have been mentioned in the studies of Xie (1994), Huang (1995) and Wen (2006), but were not listed in the study of Yuan et al. (2001). In addition, rimes which were described in the study of Xie (1994), but were not listed in other studies are displayed in italics. Including two syllabic nasals, Xie (1994) and Yuan et al. (2001) declared that there were 74 rimes in Meixian Hakka, while Huang (1995) and Wen (2006) claimed that there were 73 rimes. The possible syllable-initial consonants of every rime listed in Table 1.3 are based on the studies of Huang (1995) and Wen (2006). Consonants in shading characters denote syllable initials of the rimes which have been described only in the study of Wen (2006), whereas consonants in bold letters indicate syllable initials which have been described only in the study of Huang (1995). It can be noticed that there is some discrepancy between the two studies in the possible syllable-initial consonants of the Meixian Hakka rimes. There are three separated tables in Table 1.3.

Vowel	Diphthong		Triphthong
[i] [p, pʰ, m, f, v, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]	[ie] [k, ŋ, ø]	[ia] [m, **v**, t, ts, tsʰ, s, k, kʰ, ŋ, **h**, ø]	[iui] [ø]
	[io] [ts, tsʰ, k, kʰ, ŋ, h, ø]	[iu] [ts, tsʰ, s, t, l, k, kʰ, ŋ, h, ø]	[iai] [k, ø]
			[iau] [p, pʰ, m, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]
[e] [p, pʰ, m, v, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]	[eu] [m, f, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]		*[ieu]* [k, kʰ]
[a] [p, pʰ, m, f, v, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]	[ai] [p, pʰ, m, f, v, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]	[au] [p, pʰ, m, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]	
[o] [p, pʰ, m, f, v, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ, h, ø]	[oi] [p, pʰ, m, f, v, ts, tsʰ, s, t, tʰ, l, k, kʰ, ŋ, h, ø]		
[u] [p, pʰ, m, f, v, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ]	[ui] [**m**, ts, tsʰ, s, t, tʰ, n, l, k, kʰ, ŋ]	[ue]	[uai] [k, kʰ]
	[uo] [k]	[ua] [k, **kʰ**]	
[ɿ] [ts, tsʰ, s]			

Vowel + Nasal			Vowel + Stop		
[im] [ts, ts^h, s, l, k, k^h, ŋ, h, ø]	[in] [p, p^h, m, f, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, ø]		[ip] [ts, ts^h, s, l, k, k^h, ŋ, h, ø]	[it] [p, p^h, f, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, ø]	
[em] [ts, ts^h, s, t, n, l, k, h, ø]	[en] [p, p^h, m, f, ts, ts^h, s, t, t^h, n, k, k^h, ŋ, h, ø]		[ep] [ts, s, t, **t^h**, n, l, **k**, k^h, **h**, ø]	[et] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, h, ø]	
[am] [f, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, ø]	[an] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k^h, ŋ, h, ø]	[aŋ] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, ŋ, h, ø]	[ap] [f, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, ø]	[at] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, **ø**]	[ak] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, h, ø]
	[on] [**p**, p^h, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, h, ø]	[oŋ] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, ø]		[ot] [p, ts, ts^h, s, t, t^h, l, k, h, ø]	[ok] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, ŋ, h, ø]
	[un] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h]	[uŋ] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h, h]		[ut] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, l, k, k^h, ŋ]	[uk] [p, p^h, m, f, v, ts, ts^h, s, t, t^h, n, l, k, k^h]
[əm] [ts, ts^h, s]	[ən] [ts, ts^h, s]		[əp] [ts, s]	[ət] [ts, ts^h, s]	

Diphthong + Nasal			Diphthong + Stop		
	[ien] [p, p^h, m, ts, ts^h, s, t, t^h, l]			**[iet]** [p, p^h, ts, ts^h, s, t, t^h, l]	
[iam] [ts, ts^h, s, t, t^h, l, k, k^h, ŋ, h, ø]	[ian] [k, k^h, ŋ, h, ø]	[iaŋ] [p, p^h, m, ts, ts^h, s, **t**, t^h, l, k, k^h, ŋ, h, ø]	[iap] [ts, ts^h, s, t, t^h, l, k, k^h, ŋ, h, ø]	[iat] [k, k^h, ŋ, h, ø]	[iak] [p, p^h, ts, ts^h, s, t, l, k, k^h, ŋ, h, ø]
	[ion] [ts^h, ŋ]	[ioŋ] [p, p^h, m, ts, ts^h, s, l, k, k^h, ŋ, h, ø]		[iot] [ts]	[iok] [p^h, ts, ts^h, s, **t**, l, k, k^h, ŋ, ø]
	[iun] [k, k^h, ŋ, h, ø]	[iuŋ] [ts, ts^h, s, l, k, k^h, ŋ, h, ø]		[iut] [k^h, ø]	[iuk] [ts, ts^h, s, l, k, k^h, ŋ, h, ø]
	[uen] [k]			[uet] [k]	
	[uan] [k, k^h]	[uaŋ] [k, **k^h**]		[uat] [k, k^h]	[uak] [k]
	[uon] [k]	[uoŋ] [k, **k^h**]		[uot]	[uok] [k]

Table 1.3 Possible combinations of syllable-initial consonant and rime forming the (C)V, (C)D, (C)T, (C)VN, (C)VS, (C)DN and (C)DS syllables in Meixian Hakka

The first one displays vowels, diphthongs and triphthongs in Meixian Hakka that occur in V and CV syllables. The temporal organizations and formant frequencies of the Meixain Hakka rimes that listed in the first table will be investigated in this study. As shown in the first table, there are 6 monophthongs, 12 diphthongs and 5 triphthongs in Meixian Hakka. Among all vowels, [ʅ] is preceded by a limited number of syllable-initial consonants. Similar cases can be found in diphthongs [ie, uo, ue, ua] and triphthongs [iai, iui, ieu, uai]. Like Cantonese and Min, Hakka is well-known for its syllable-final consonants. Meixian Hakka preserves the six syllable-final consonants [p, t, k, m, n, ŋ] found in Middle Chinese. The second and third tables in Table 1.3 display Meixian Hakka rimes, which are composed of vowels or diphthongs followed by syllable-final consonants, in VC and CVC syllables. The former one presents rimes that are composed of a vowel followed by a nasal or a stop, whereas the latter one displays rimes which are composed of a diphthong followed by a nasal or a stop.

Being one of the dialects of the Hakka group of the Chinese language, Meixian Hakka is tonal. That is, difference in tone indicates lexical contrast. Tones in Meixian Hakka have been extensively studied on their historical development and phonological perspective in the past studies (Huang, 1988; 1989; 1992; Xie, 1994; Zhou, 1998; Xie, 2003). Researchers of the past studies generally agreed that there were six citation tones in Meixian Hakka (Hashimoto, 1973; Norman, 1988; Huang, 1992; 1995; Xie, 1994; Yuan et al., 2001). Based on the Qieyun dictionary which was the authoritative source for reconstructing Middle Chinese, Karlgren (1954) declared that there were four tonal categories in Middle Chinese, namely Ping Sheng (平声), Shang Sheng (上声), Qu Sheng (去声), and Ru Sheng (入声). He believed that the traditional names of the four tonal categories in the Qieyun dictionary were descriptive terms. In terms of the contour shape, the four tonal categories were interpreted as the level, rising, departing and entering tones, respectively. Among the four tonal categories, the entering tone (Ru Sheng) was superimposed on syllables of which the syllable finals were [p], [t] or [k]. The four Middle Chinese tonal categories have been split into two registers, which were Yin (阴) and Yang (阳), forming eight possible tone types in modern Chinese. The eight tone types are Yin Ping (阴平), Yang Ping (阳平), Yin Shang (阴上), Yang Shang (阳上), Yin Qu (阴去), Yang Qu (阳去), Yin Ru (阴入) and Yang Ru (阳入). Yin was realized as the higher or upper register which was applied to syllables with Middle Chinese voiceless initials, and Yang was realized as the lower register which was applied to syllables with Middle Chinese voiced initials (Norman, 1988). However, in some Chinese dialects, the Yin register may exhibit a tone with a lower pitch, and conversely, the Yang register may exhibit a tone with a higher pitch. Ru Sheng in Hakka is one of the examples. Syllables that belong to the tone type of Yin Ru always have a lower pitch, whereas syllables that belong to Yang Ru have a higher pitch. Karlgren (1915-1926) stated that there was no split in Shang Sheng and Qu Sheng in Meixian Hakka. In this case, there were six citation tones in Meixian Hakka which were derived from the tonal categories of Middle Chinese. In his study, the six citation tones were transcribed as follows: (1) Yin Ping was a level tone at the middle range of pitch with a slight rise and non-abrupt, (2) Yang Ping was a level tone at the low pitch range and non-abrupt, (3) Shang Sheng was a falling tone from the half-low to low pitch range and non-abrupt, (4) Qu Sheng was a falling tone from the high to low pitch range and non-abrupt, (5) Yin Ru was a falling tone from the half-low to low pitch range and abrupt, and (6) Yang Ru was a level tone at the high pitch range and abrupt. In this view, the citation tones in Meixian Hakka were perceived as two level tones, two falling tones and two short tones, which were characterized by different registers. However, there is some discrepancy in the descriptions of contour and pitch value of the Meixian Hakka tones among the past studies. Such discrepancy will be illustrated in detail in the latter chapters. For example, many linguists stated that Qu Sheng in Meixian Hakka was a high falling tone, but Hashimoto (1973) held a different view. He claimed that Qu Sheng was a high level with a little rising legato instead. In order to give a detailed description of the citation tones in Meixian Hakka, acoustic data obtained from the real speech of the native speakers are indubitably necessary.

1.3 Purposes and scope of the research

Though the sound system in Meixian Hakka was extensively described in the past studies, the transcriptions of tones and vowels were phonological and mainly based on individual impression of the linguists. There is no experimental data presented in the past studies and little work has been done on the phonetic analysis of the sound system in Meixian Hakka. The present study aims at investigating the tones and vowels in Meixian Hakka from acoustic and perceptual perspectives. The acoustic characteristics of vowels and tones in Meixian Hakka are uncovered through scientific investigations. With regard to vowels, formant frequencies for the first three formants and durations of the steady states and intermediate transitions were measured. On the other hand, F_0 contours of the citation tones in monosyllabic words are investigated. F_0 values at different percentage points of a syllable were obtained and reconstructed to exhibit the F_0 contours of tones. In addition, the realizations of tones in different phonetic contexts and in bisyllabic tone sandhi are also discussed. All possible two-tone combinations were included in the word list and the patterns of tone sandhi in Meixian Hakka bisyllabic words are examined. Based on the acoustic properties of tones in this study, the relative potency of tonal height, slope of F_0 contour, F_0 onset and F_0 offset in tone perception are explored. A perceptual experiment on the four long tones in Meixian Hakka was designed and carried out to examine the correlation between perceptual responses and acoustic properties of tones in Meixian Hakka. The purposes of this study are (1) to analyze the patterns of F_0 contours of the citation tones in isolation and their variations in different contexts, (2) to review the pitch value of the citation tones based on the raw and normalized data, (3) to compare the tone sandhi patterns observed in this study with the descriptions in the past studies, (4) to examine the potency of various perceptual cues in tone perception in Meixian Hakka, (5) to evaluate the correlation between identification responses and acoustic properties of tones in Meixian Hakka, and (6) to investigate the formant frequencies and temporal organization of monophthongs, diphthongs and tripthongs in Meixian Hakka. It is believed that the acoustic data in this study may achieve objectivity and provide a reliable basis for a better understanding of acoustic properties of vowels and tones as well as the correlation between perceptual and acoustic characteristics of tones in Meixian Hakka.

1.4 Outline of the dissertation

This study consists of five chapters. Chapter one introduces background information of the dialect of Meixian Hakka as well as the purposes and scope of the research. The phonological descriptions and transcriptions of the sound system of Meixian Hakka in the past studies are to be displayed. Besides, the geographic distribution of the Hakka dialects in China is exhibited. As a result of the lack of experimental data in the past studies, the area of this research covers studies aiming to attain a better understanding of the acoustic and perceptual characteristics of tones and vowels in Meixian Hakka. Chapter two investigates the phonetic characteristics of the citation tones in isolation and in different contexts. The F_0 trajectories of the six citation tones are plotted based on F_0 values obtained from the speech data. To factor out between-speaker variation in physical-phonetic manifestation of speech signals and to draw out invariant correlates of the linguistic units, the absolute F_0 values of tones were normalized using log z-score.

There is great divergence in opinions from linguists in the descriptions of pitch value of the citation tones in Meixian Hakka. In view of the normalized data, the pitch values of the six citation tones are evaluated. The variations in F_0 trajectory of the citation tones on syllables with vowels [i, a, u] are examined to see whether intrinsic fundamental frequency (iF0) exists between high and low vowels. In addition, the citation tones in CV(C) syllables are compared with those in V(C) syllables and mean F_0 shapes of the citation tones for the two syllable types are discussed. The F_0 patterns of all possible two-tone combinations in Meixian Hakka are also analyzed. By examining the F_0 contours of the first- and second-syllable tones, tone sandhi rules in Meixian Hakka are formulated. The similarities and differences between the tone sandhi rules observed in this study and the sandhi rules described in the past studies are displayed. Chapter three examines the perceptual properties of the citation tones in Meixian Hakka and the correlation between perceptual and acoustic characteristics of tones. The methodology and test materials of the perceptual test are shown. The base syllable of stimuli was from real speech and stimuli in this study were composed of varying onset and offset values. Through the analysis of perceptual responses of the Hakka listeners, the potency of varying perceptual cues in tone perception is investigated. Chapter four examines the acoustic characteristics of the Meixian Hakka vowels. Formant frequency values and durations of monphthongs, diphthongs and triphthongs in Meixian Hakka are investigated. Based on the acoustic data in this study, the phonetic transcription of the vowel system of Meixian Hakka has been proposed. The phonetic transcription in this study is compared with the phonological transcriptions in the past studies. Chapter five concludes the findings and discussions of this study. Some generalizations regarding the F_0 patterns of tone and formant patterns of vowel are stated.

CHAPTER TWO

ACOUSTIC ANALYSIS OF TONES IN ISOLATION AND IN CONTEXT

This chapter aims at investigating the acoustic properties of tones in Meixian Hakka. The chapter is divided into two main sections, namely citation tones on monosyllabic words and citation tones in bisyllabic tone sandhi. Literature reviews, the methodology of measurement and results are described and illustrated in the subsections of each section. In section 2.1, the focus is on tone patterns of the citation tones in Meixian Hakka. A normalization strategy was chosen and applied to the obtained F_0 data. The realization of F_0 contours of the citation tones on monosyllabic words is presented and discussed in section 2.1.3.1. In addition, the influence of vowel quality and syllable type on the F_0 contour of the citation tones are analyzed respectively in sections 2.1.3.2 and 2.1.3.3. In section 2.2, tonal alternation patterns in context are investigated. The section focuses on F_0 patterns of the citation tones in tone sandhi. All the possible combinations of tones in bisyllabic words were tested to examine the tone sandhi patterns in Meixian Hakka. Through the acoustic analysis of tones on monosyllabic words and in tone sandhi, the results of this chapter provide a phonetic basis for the analysis of the tone inventory of Meixian Hakka.

2.1 Citation Tones on Monosyllabic words

2.1.1 Introduction

The Meixian (Meizhou) dialect of Hakka is tonal which means differences in pitch and pitch contour indicate lexical contrast. The historical development of tone system in Meixian Hakka has been extensively studied in the past (Huang, 1988; 1989; 1992; Xie, 1994; Zhou, 1998; Xie, 2003). Hashimoto (1973) stated that there was a tonal development of Shang Sheng (上声) in Middle Chinese which was found in all Hakka dialects but not in other Chinese dialects. To be more specific, part of the Shang Sheng (上声) of the Hakka language in Middle Chinese merged into Yin Ping (阴平) in modern Chinese. The tonal development of Shang Sheng was unique to the Hakka dialects and it distinguished the dialects of the Hakka group from other Chinese dialects, i.e. Mandarin, Gan, Min, Wu, Xiang, and Cantonese. In contrast, the tonal development of Ping (平) tone and Ru (入) tone of the Hakka dialects was similar to other Chinese dialects (Guo, 1993; Huang, 1988; 1989). Depending on voicing of syllable-initial consonants in Middle Chinese, the tones split into yin (阴) and yang (阳) in modern Chinese. Ping and Ru tones manifested as yin (阴) in syllables that had voiceless initials in Middle Chinese, while they manifested as yang (阳) in syllables that had voiced initials. It has been described in the past studies that there are six citation tones in Meixian Hakka and they are two level tones, two falling tones and two entering tones (Hanyu Fangyan Zihui, 1989; Huang, 1992; Xie, 1994; Yuan et al., 2001). According to the analysis of Chinese Phonology, the tone system of Middle Chinese consisted of four tone categories, which were Ping (平), Shang (上), Qu (去) and Ru (入). The historical process of yin-yang dichotomy in Ping and Ru tones has given rise to the six categories of the tone system in modern Meixian Hakka. The six tone categories are Ying Ping (阴平), Yang Ping (阳平), Shang Sheng (上声), Qu Sheng (去声), Ying Ru (阴入) and Yang Ru (阳入). As mentioned in the studies of Hanyu Fangyan Zihui (1989) and Yuan et al. (2001), Ying Ping in Meixian Hakka corresponds to a high level tone; Yang Ping corresponds to a low level tone; Shang Sheng corresponds to a mid-to-low falling tone; Qu Sheng corresponds to a high-to-low falling tone; Yin Ru corresponds to a low entering tone; and Yang Ru corresponds to a high entering tone. In terms of pitch value, there are discrepancies among the past studies in the descriptions of the six citation tones in Meixian Hakka. It is assumed that the discrepancies may be due to either the differences in authors' individual impression on tone contour and tone level or the differences in language background of speakers, that is, the differences in hometown of speakers. The city of Meizhou is considered the center for standard Hakka dialect. It has jurisdiction over seven counties and one district. The seven counties are Meixian (梅县), Dubu (大埔), Fengshun (丰顺), Wuhua (五华), Jiaoling (蕉岭), Pingyuan (平远) and Xingning (兴宁) while the one district refers to the district of Meijiang (梅江区), which is located in the center of the county of Meixian (梅县). Though people in the city of Meizhou speak in Hakka dialects, Meixian Hakka that is spoken in the county of Meixian is linguistically different from the Hakka dialects spoken in the other six counties.

It was generally acknowledged that F_0 variation is influenced by neighbouring consonants and vowels. The F_0 perturbation by consonants has been investigated in the past studies (Lehiste & Peterson, 1961; Mohr, 1971; Gandour, 1974; Xu & Xu, 2003; Francis et al., 2006; Wong & Xu, 2007; Hanson, 2009). Regarding the influence of consonant voicing on F_0, Gandour (1974) and Hombert (1978) found in both tonal and non-tonal languages that F_0 at the onset of vowel after voiceless obstruent was relatively high and there was a falling transition between consonant and vowel. In contrast, F_0 at the onset of vowel after voiced obstruent was relatively low and there was a rising consonant-vowel transition. It was observed in the study of Hombert (1978) that the perturbatory effect of consonant voicing on tones was greater in non-tonal languages than that in tonal languages. In addition to the perturbation by prevocalic voicing, consonant aspiration also affects F_0 at the onset of the following vowel. F_0 at the vowel onset after a voiceless unaspirated stop was found to be higher than that after a voiceless aspirated stop (Gandour, 1974; Xu & Xu, 2003; Francis et al., 2006). However, conflicting results were also found in other studies (Han & Weitzman, 1970; Wong & Xu, 2007) and results showed that F_0 of the following vowel after voiceless aspirated stops is higher than that after voiceless unaspirated stops. As this study mainly concerns the overall patterns of F_0 contours of the citation tones in Meixian Hakka, F_0 variations at the vowel onset was not taken into account in the analysis.

The correlation between vowel and F_0 has been observed in the past studies (Ladd & Silverman, 1984; Whalen & Levitt, 1995; Connell, 2002; Gonzales, 2009) and such correlation is described as the presence of vowel intrinsic fundamental frequency (iF0). IF0 was verified in languages of different language families (Whalen & Levitt, 1995; Connell, 2002) and it is considered to be universal. IF0 is shown to be related mainly to vowel height rather than vowel backness. High vowels have a tendency to exhibit higher fundamental frequency than low vowels (Zhu, 1994; Whalen & Levitt, 1995; Connell, 2002; Gonzales, 2009). With respect to frequency range, F_0 differences between high and low vowels were frequently diminished when the tone was in a low frequency range as opposed to a high frequency range. Ladd and Silverman (1984) also found iF0 in connect speech as well as in controlled environments. In comparison to iF0 in controlled environments, Ladd and Silverman displayed that differences between high and low vowels in connected speech were smaller. In this chapter, the correlation between vowel and F_0 in Meixian Hakka is evaluated based on the obtained speech data and the corresponding results are presented in section 2.1.3.2.

No acoustic data of the citation tones in Meixian Hakka has been presented in the past and it is worthwhile to investigate the pitch value and tone contour of the six citation tones in Meixian Hakka from an objective perspective. It is believed that the results of the acoustic data in this study can clarify the discrepancies in pitch value among the past studies. The focus of this chapter is particularly drawn onto the acoustic analysis of the six citation tones in Meixian Hakka. The purposes of this study are (1) to investigate F_0 patterns of the six citation tones, (2) to evaluate the interaction between F_0 and vowel quality, and (3) to examine the influence of syllable type on F_0 contour, in Meixian Hakka. The acoustic results obtained in this study have been normalized using log z-score transform and the results of normalization have been used to determine the pitch values of the six citation tones in Meixian Hakka. The obtained pitch values in this study have been compared with the pitch values described in the past study that mainly based on the authors' impression

2.1.2 Methodology

a. Speakers

The speech data were provided by 5 male and 5 female speakers. All of them were native speakers of Meixian Hakka who were born and grew up in the district of Meijiang (梅江区). The district of Meijiang is located in the centre of the county of Meixian (梅县) and the county of Meixian is under the jurisdiction of Meizhou City (梅州市). The speakers were in the old age group and they were at the age between fifty and seventy with no history of speech and hearing disorders.

b. Test Materials

The word list consists of Chinese characters exemplifying the six citation tones in Meixian Hakka on the monosyllabic words. Table 2.1 displays the historical tone categories, the suggested pitch values or tone marks and all test monosyllabic words used in the study. The test words were V(C) syllables that included vowel [i] and CV(C) syllables that included vowel [i], [a], or [u]. In Meixian Hakka, meaningful

Historical tone categories	Suggested pitch values / tone marks	Test monosyllabic words			
		V(C)	CV(C)		
		[i]	[i]	[a]	[u]
Ying Ping (阴平)	[33] / [˧]	[i33] 医 'to cure'	[ts^{h}i33] 妻 'wife'	[ts^{h}a33] 车 'car'	[ts^{h}u33] 抽 'to take out'
Yang Ping (阳平)	[11] / [˩]	[i11] 余 'remaining'	[ts^{h}i11] 糍 'glutinous rice cake'	[ts^{h}a11] 茶 'tea'	[ts^{h}u11] 除 'to remove'
Qu Sheng (去声)	[51] / [˥˩]	[i51] 意 'desire'	[ts^{h}i51] 趣 'interest'	[ts^{h}a51] 岔 'to branch off'	[ts^{h}u 51] 臭 'stinking'
Shang Sheng (上声)	[41] / [˦˩]	[i41] 雨 'rain'	[ts^{h}i41] 取 'to take'	[ts^{h}a41] 扯 'to pull'	[ts^{h}u 41] 暑 'hot weather'
Yin Ru (阴入)	[$\underline{41}$] / [˦˩]	[it$\underline{41}$] 一 'one'	[ts^{h}it$\underline{41}$] 七 'seven'	[sat$\underline{41}$] 杀 'to kill'	[ts^{h}ut$\underline{41}$] 出 'to come out'
Yang Ru (阳入)	[$\underline{55}$] / [˥]	[it$\underline{55}$] 翼 'wing'	[ts^{h}it$\underline{55}$] 疾 'disease'	[sat$\underline{55}$] 舌 'tongue'	[ts^{h}ut$\underline{55}$] 捽 'to wipe clean'

Table 2.1 Historical tone categories and suggested pitch values or tone marks for the six citation tones in Meixian Hakka, and the test monosyllabic words used in the study

monosyllabic words can be produced by all possible combinations of vowel [i] and each of the six citation tones without any syllable-initial consonant. However, it is not the case for other vowels. To examine the effect of vowel quality on tones, CV(C) syllables that composed of the six citation tones in Meixian Hakka, vowels [i, a, u], and/or syllable-initial consonants [ts^h, s] were included in the word list. Voiceless aspirated obstruents were chosen as the syllable-initial consonants of the test words to control the consonantal effect on the voiced part of the syllable within and between speakers. The vast majority of the test syllables had a voiceless aspirated alveolar affricate [ts^h] as the syllable initial, while only minority of which had a voiceless alveolar fricative [s] as the syllable initial. For syllables that were superimposed by the entering tones, which were [$\underline{41}$] and [$\underline{55}$], each of the test monosyllabic words had an unreleased voiceless unaspirated alveolar stop [t˺] as syllable final. Each of the test words was embedded in a carrier sentence [ŋai21 t^huk˺55 ____ pun33 m̩11 t^haŋ33] "I read ___ for you (to) listen.". The order of the test monosyllabic words was randomized and four readings of the word list were recorded. The speakers were instructed to read the word list at a normal rate of speech and audio recordings were conducted in a quiet room.

c. Data Analysis

F_0 values and overall duration of the F_0 contours for each of the six citation tones in Meixian Hakka were captured using Praat 4.3.19 speech analysis software. For the six citation tones in Meixian Hakka, which are Yin Ping, Yang Ping, Qu Sheng, Shang Sheng, Yin Ru and Yang Ru, 11 percentage points were sampled from the overall duration of every F_0 contour (0% - 100%). The measurement procedures used in this study were similar to those described in the studies of Rose (1982, 1990) and Zhu (1999). The onset of the F_0 contour was interpreted as onset of the voicing part, that is, the vocalic nucleus, of a syllable. The voicing onset was located at the first periodicity of the voicing part of the syllable. In many cases, the first glottal pulse was not included as the first pulse was not necessary to be regular and the amplitude of the pulse was too weak to be audible (Lisker & Abramson, 1964; Rose, 1982; Zhu, 1999). The onset of the F_0 contour was the first data point (0%) and the subsequent points were sampled at every 10% of the overall duration of the F_0 contour. The mean F_0 values of each of the 11 percentage points and mean duration of the F_0 contour for the citation tones were measured and plotted on a figure for comparison in this chapter.

Figure 1a shows the waveform and F_0 contour of a Yin Ping tone in Meixian Hakka which were displayed in the Praat window. As can be seen, the total duration of the voicing part of the syllable was divided into ten equal portions and the F_0 values of 11

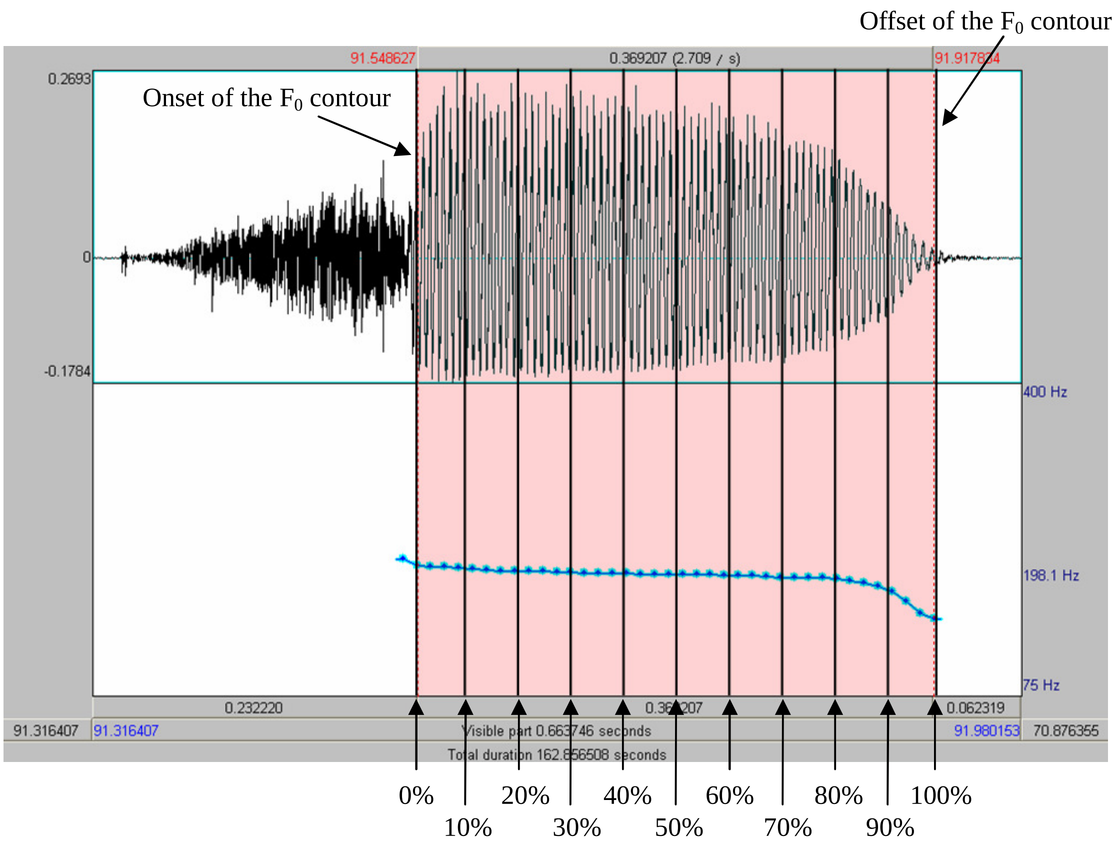


Figure 1a Waveform and raw F_0 contour of a Yin Ping tone in Meixian Hakka showed in the window of Praat. The overall duration of the F_0 contour was divided into 10 equal portions and F_0 values of 11 percentage points were extracted.

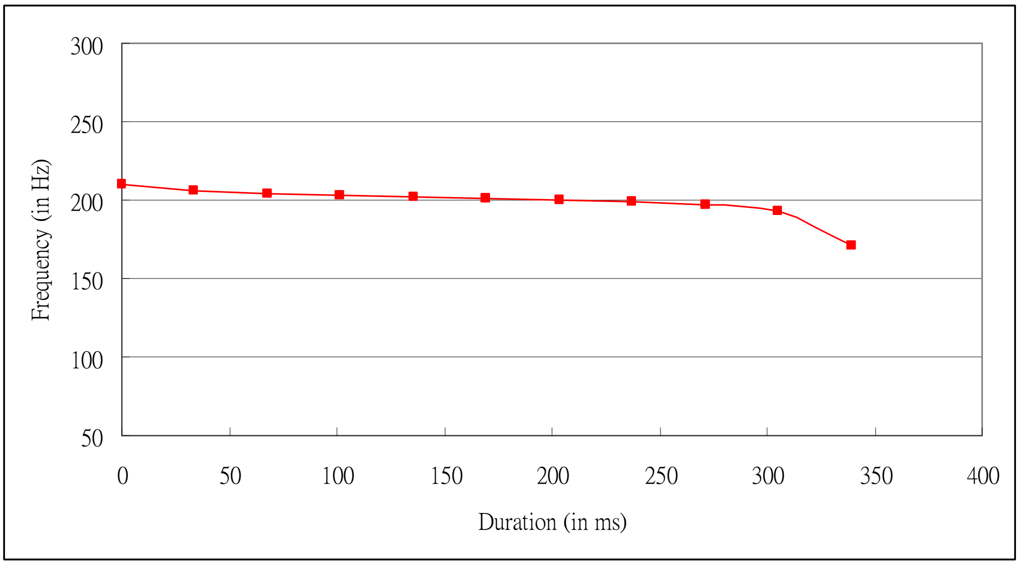


Figure 1b F_0 contour of a Yin Ping tone in Meixian Hakka plotted based on the data of 11 percentage points.

data points were measured. The extracted data points of each citation tone in all repetitions were then averaged and plotted on figures. Figure 1b shows the F_0 contour of the Yin Ping tone based on the mean F_0 values of the 11 data points extracted from

Figure 1a. The shape of the extracted F_0 curve as shown in Figure 1b resembles the calculated F_0 curve as shown in Figure 1a. The syllable duration, in millisecond (ms), is on the x-axis of the graph in Figure 1b and the fundamental frequency, in Hertz (Hz), is on the y-axis of the graph. In the following sections, only the F_0 curves will be shown on the graphs, while the solid squares representing the 11 data points in Figure 1b will not be shown.

d. Normalization of the F_0 contours

According to Peterson and Barney (1952) and Rose (1987), the same phonemic sound produced by different speakers of a language resulted in different acoustic outputs. It was stated that the between-speaker differences in acoustic outputs were caused by the differences in speakers' vocal tract size and size of the vocal cords. Yang and Kasuya (1996) measured the anatomical differences in vowel tract shapes between male and female speakers by magnetic resonance images. They demonstrated that the whole vocal tract length of male speakers was relatively longer than that of female speakers. Male speakers tend to have relatively larger laryngeal cavities and proportionally shorter oral cavities than female speakers. Simpson (2009) pointed out that difference in length and mass of vocal folds was found between male and female speakers. Male speakers tend to have longer and thicker vocal folds than female speakers, thus causing corresponding F_0 values to be lower. This is consistent with the observations of Rose (1987). Besides the physiological factors that have been mentioned, the between-speaker variation can also result from non-physiological factors such as dialect background, emotion, speaking tempo and speaker idiosyncracies (Rose, 2002; Simpson, 2009). To extract the invariant acoustic correlates of linguistic sounds produced by different speakers, normalization was proposed in the studies of Nearey (1978) and Disner (1980) on vowels and Rose (1987) on tones as a mathematical procedure to transform absolute F_0

values into normalized values. Maddieson (1978) also claimed that tone normalization was necessary for cross-language comparison. As the recorded data in this study were uttered by multiple informants of both sexes, normalization is necessary to factor out non-linguistic individual differences and to draw out invariant linguistic parameters of the citation tones of Meixian Hakka. Normalized F_0 values for the six citation tones in Meixian Hakka are calculated based on the obtained absolute F_0 values. It is believed that a more scientific representation and description of the six citation tones of Meixian Hakka can be achieved through the normalization process.

Two normalization strategies which were generally adopted in the past studies (Rose, 1987; Zhu, 1999; Ishihara, 2004) are considered in this study. They are linear z-score normalization strategy (also known as z-score transform) and log z-score normalization strategy (also known as log z-score transform). In the study of Zhu (1999), the log z-score transform was preferable over the z-score transform in consideration of three factors which are (i) the normalization index (NI) of different normalization strategies, (ii) the statistical advantage that reflects the production mechanism and (iii) the relation between F_0 values and perceived pitch. In order to determine the most appropriate strategy for the reduction of between-speaker differences in F_0 realization of tones in Meixian Hakka, the selection of normalization strategy in this study takes the three factors into account. As proposed by Rose (1987), normalization index (NI), which is the ratio of dispersion coefficient (DC) of raw data to the DC of normalized data, is a mean of comparing results obtained from different normalization strategies. On one hand, the value of DC reflects the degree to which the F_0 values of different speakers cluster together. A lower value of DC indicates a higher degree of speakers' F_0 values cluster after normalization. On the other hand, a higher value of NI indicates a higher degree of reduction of between-speaker variation. The NI obtained from the log z-score transform is compared with the NI obtained from the z-score transform. The normalization result shows that log z-score normalization strategy reduces the amount of between-speaker variance by a factor of 24.9. The DC value calculated on the normalized values of log z-score transform reduces to 0.031, as compared to the DC value calculated on the absolute frequency values of 0.774. In contrast, the NI of z-score transform is slightly better than that of log-z-score transform. The between-speaker variance reduces by a factor of 26.6 after applying linear z-score normalization strategy to the raw F_0 data and the DC value calculated on the normalized values of z-score transform reduces to 0.029. The effect of normalization on the raw F_0 data indicated by NI demonstrates that the z-score transform seems to perform better than the log z-score transform. However, as mentioned above, NI is not the only criterion for selecting the normalization strategy and there are two other factors which need to be considered.

Regarding the statistical advantage, Zhu (1999: 53) remarked that the logarithmically converted F_0 values had a more normal distribution than the absolute F_0 values and the log F_0 values were statistically advantageous for normalization. He claimed that the positive skewness factors for the absolute F_0 values indicated the unevenness of F_0 distribution and the lower part of the F_0 range was more exploited in contrasting tones. As stated in the studies of Zhu (1999) and Hsieh (2007), the positive skewness has been widely attested in non-tonal languages and tonal languages. Zhu (1999) and Hsieh (2007) suggested that the unevenness of F_0 distribution may reflect the logarithmic nature of F_0 variations with respect to the production mechanism. They agreed with the hypothesis of Fujisaki (1983) regarding the linear correlation between the change of logarithmic fundamental frequency and the displacement of a point in the laryngeal structure. It was supposed that a smaller F_0 interval at the lower part and a larger F_0 interval at the higher part of the F_0 range would appear as the same distance on the logarithmic scale (Zhu, 1999). Table 2.2 presents the skewness factors of F_0 distributions for the absolute F_0 values in Hz and those for the log F_0 values in this study. The corresponding factors are graphed for the ten speakers in Figure 1c. Result shows that the skewness factors for raw F_0 in Hz are positively skewed and this probably implies the logarithmic nature of the F_0 change as mentioned above. Table 2.2 shows that the skewness factors has been reduced by a mean factor of 0.009 (=0.231-0.222). Though the distinct statistical advantage of log F_0 displayed in Zhu's study (1999) cannot be found in this study, the relatively small mean factor for log F_0 indicate that the log F_0 values have a more normal distribution than the absolute F_0 values in Hz and F_0 data in logarithmic scale are assumed to have an advantage for statistical manipulation.

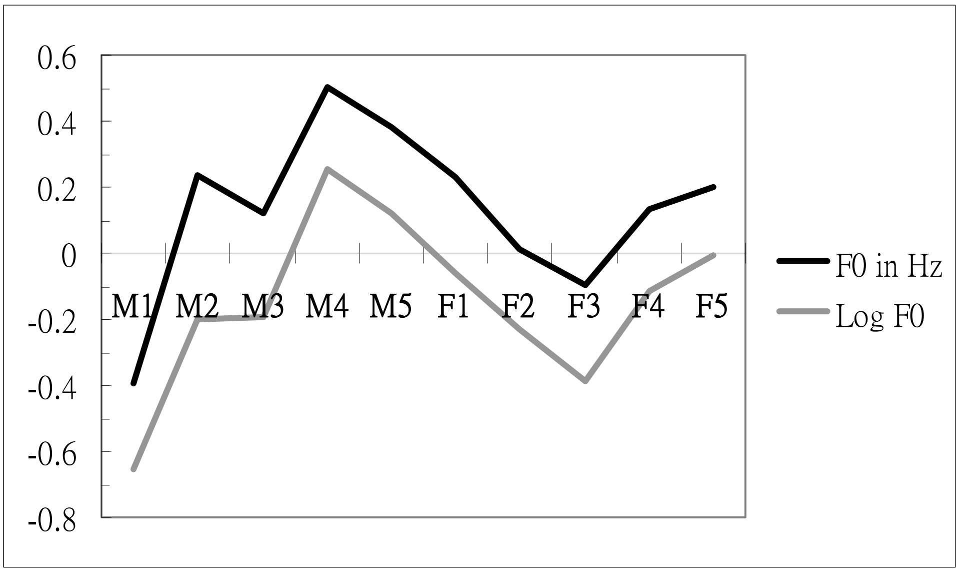


Figure 1c Skewness factors indicating the distributions of speakers' raw F_0 and those of speakers' log F_0 in line presentation

	M1	M2	M3	M4	M5	F1	F2	F3	F4	F5	Mean
Hz	-.396	.237	.119	.505	.381	.231	.010	-.097	.134	.197	.231
Log (Hz)	-.652	-.198	-.193	.254	.119	-.059	-.228	-.388	-.118	-.008	.222

Table 2.2 Skewness factors and their absolute means indicating the distributions of raw F_0 and those of log F_0

Mel scale is a unit for measuring subjective pitch (Stevens et al., 1937; Stevens & Volkmann, 1940). Pairs of sounds that have equal perceptual distance in pitch are separated by an equal number of mels. For frequencies above 1 kHz, it is well-known that the relation between perceived pitch and frequency is logarithmically correlated. However, it is controversial whether such relation is logarithmic or linear for frequencies below 1 kHz. Here we emphasize on the relation between perceived pitch in mel scale and physical frequencies in Hz that below 500 Hz. Frequencies below 500 Hz are concerned as they contain those harmonics which tend to play a major role in pitch perception (Moore, 1997). Stevens and Volkmann (1940) stated that the relation between perceived pitch and physical frequencies in low frequency domain was approximately linear. But Zhu (1999) held a different view. He pointed out that the relation between perceived pitch and frequency was logarithmic even in the low frequency domain due to the reason that the mel scale of equal perceived pitch below 500 Hz is not necessary to be linear and the quasi-linear relation below 500 Hz can be explained by the nature of logarithmic algorithms (1999: 54-55). From a perceptual point of view, it seems fair to believe that the quasi-linear relation between perceived pitch and frequencies below 500 Hz would be an approximation of the logarithmic algorithms and logarithmic algorithms rather than linear algorithms may be more appropriate for the F_0 transform of frequency.

Zhu (1999) and Ishihara (2004) showed that in terms of theoretical consideration and statistical performance, log z-score transform was preferable than other normalization strategies for the linguistic tones. Similarly, due to the logarithmic nature of frequency change related to tone production and perception, log z-score transform was adopted as the normalization strategy in this study. In the study of Jassem (1975), it has been proposed that the heights of tones for each speaker were relative to individual's range of tones and the realization of tones should be speaker-dependent. Based on this assumption, the z-score transform is expressed as 'an observed F_0 value as a multiple of a measure of dispersion away from a mean F_0 value' (Rose, 1987:347). The formula of z-score transform is as follows:

$$z_i = \frac{x_i - m}{s}$$

where x_i is an observed F_0 value, m is the mean value, and s is the standard deviation of a given sample. The formula of log z-score transform is similar to that of z-score transform, but all F_0 values are in logarithmic scale. The formula of log z-score that used in this study is as follows:

$$z'_i = \frac{y_i - m_y}{s_y} = \frac{\log_{10} x_i - \frac{1}{n}\sum_{i=1}^{n} \log_{10} x_i}{\sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(\log_{10} x_i - \frac{1}{n}\sum_{i=1}^{n}\log_{10} x_i\right)^2}}$$

where y_i is an logarithmically converted F_0 value, m_y is the logarithmic mean of raw F_0 values, and s_y is the standard deviation of a given sample.

The selection of normalization points in this study is similar to that in the study Zhu (1999). In order to maximize the between-speaker comparability of tones, the F_0 values obtained at 0% of the duration for all citation tones, and those at 100% of the duration of Yang Ru and Yin Ru were excluded from the normalization. The F_0 values at the onset time point (0%) for all tones were excluded as they were attributed to individual speaker and were affected by the syllable-initial consonant (Rose, 1987; Zhu, 1999). The F_0 values at the offset time point (100%) for Yang Ru and Yin Ru were also excluded because they were considered as an acoustic reflection of the closing portion of the syllable-final stop (Zee & Maddieson, 1980; Rose, 1987).

In addition to the fundamental frequency, absolute duration and normalized duration of tones are presented in this study. In contrast to the absolute values of duration, it is supposed that the normalized values show a better understanding of the perceptual process of a speaker. Normalized duration is computed in order to reduce between-speaker and within-speaker variations across tones (Zhu, 1999). Studies of different languages including English (Hillenbrand et al., 1995; Simpson & Ericsdotter, 2003), Swedish (Ericsdotter & Ericsson, 2001; Simpson & Ericsdotter, 2003) and Creek (Johnson & Martin, 2001) showed consistent differences in duration between male and female speakers. There were two main gender differences with regard to duration. First, female speakers were found in general to produce longer duration than male speakers. Second, durational differences between long and short vowel pairs of a language for female speakers were greater than for male speakers. In the study of Zhu (1999), cross-tone mean duration yielded the highest NI. The duration normalization method manipulated in this study takes individuals' cross-tone mean duration as reference and the average duration of all tones was used as the base to calculate the normalized duration of tones for each speaker.

2.1.3 Results

F_0 realizations of the citation tones in Meixian Hakka have been investigated in this study. The citation tones on vowels [i, a, u] and those in V(C) and CV(C) syllables were analyzed to discuss the influence of vowel quality and syllable type on tone. Figures are used to present average F_0 contours of the citation tones based on the raw data of five male and five female speakers. The results of this chapter are divided into three parts. The first part discusses the acoustic characteristics of F_0 contours of the citation tones in Meixian Hakka. The F_0 contours of each tone category in Meixian Hakka are described in terms of the absolute and normalized fundamental frequency. The shapes and values of F_0 contours of each of the tone categories in Meixian Hakka, namely, Yin Ping, Yang Ping, Shang Sheng, Qu Sheng, Yin Ru and Yang Ru, are examined and compared with each other. Statistical tests were used in the analysis of the acoustic data. The results of this study are further compared with the traditional descriptions of the citation tones in the past studies. The second part compares the F_0 contours of the citation tones on vowels [i, a, u] and examines the influence of vowel quality on tone. In addition, the influence of syllable type on tone is investigated in the third part by comparing the F_0 contours of the citation tones between V(C) syllables and CV(C) syllables. In addition to the fundamental frequency, the durations of F_0 contours of the citation tones were normalized and both the absolute values as well as the normalized values of the durations are displayed. It is expected that the durations of F_0 contours of the two Ru tones in closed syllables ending in a stop consonant, /-p/, /-t/, or /k/, would be shorter than those of the non-Ru tones in open syllables because closed syllables were perceived by Chinese phonologists as having concomitant shorter tones. However, results of this study demonstrate that though, in most cases, the two Ru tones are shorter in duration than the non-Ru tones in Meixian Hakka, it is not necessary to be the case. Instead, between-speaker variation is observed with respect to the duration for the Ru tones. As the raw data of Yang Ru tone for Female Speaker 3, Male Speaker 1 and Male Speaker 3 were corrupted, they were not included in the analysis. This chapter describes the acoustic characteristics of the citation tones in Meixian Hakka in detail from a phonetic basis. The obtained acoustic data were then used in the latter chapters as the basis for the analysis of tones in context and the synthesis of the non-Ru tones for the experiments on tone perception.

2.1.3.1 F_0 contours of the citation tones in Meixian Hakka

The Meixian Hakka system consists of six tone categories, namely Yin Ping, Yang Ping, Shang Sheng, Qu Sheng, Yang Ru and Yin Ru. Past studies showed that there are six citation tones in Meixian Hakka (Hashimoto, 1973; Huang, 1992; Xie, 1994; Yuan et al., 2001). Each of the six citation tones belongs to one tone category. In this section, the acoustic data of the six citation tones in Meixian Hakka are presented and discussed. The F_0 contours of the citation tones that occur in CV(C) syllables were analyzed. Results of this study show that there are three types of F_0 contours, which are level, simple falling and delayed falling tones, in Meixian Hakka. Level tone is realized on syllables associated with Yin Ping and Yang Ru. Simple falling tone is realized on syllables associated with Yang Ping, Shang Sheng and Yin Ru, whereas delayed falling tone is realized on syllables associated with Qu Sheng. The shapes of F_0 contours of each of the six citation tones for the male and female speakers are in general similar though inter-speaker variation in F_0 contour was found in the data. The acoustic characteristics of the citation tones are presented separately according to the tone categories. The raw F_0 data produced by the ten native speakers were normalized and are expressed in terms of the five-point scale proposed by Chao (1930). Normalization is necessary for the analysis as pitch, in physical term, varies in a wide range in speech and the unrelated linguistically significant variation have to be abstracted away to achieve the linguistic representation of tone (Anderson, 1978). Pitch values for each of the six citation tones in Meixian Hakka were assigned based on the the normalized F_0 contours. The assigned pitch values for the six citation tones are then be compared with the descriptions of the Meixian Hakka citation tones in the past studies.

a. Yin Ping and Yang Ping

Figures 2.1 and 2.2 exhibit the mean F_0 contours of Yin Ping and Yang Ping in Meixian Hakka produced by five male and five female speakers, respectively. A comparison of the F_0 contours of Yin Ping shows that for all male and female speakers, the shapes of the F_0 contours are similar. The F_0 contours of Yin Ping for the majority of speakers are basically level and remain stable throughout the overall duration of a syllable. The F_0 contours of Yin Ping for F1, F2 and M2 display that there can be a slight drop at the offset of the contour. Regardless of the data of M2, the F_0 contours of Yin Ping of the male speakers fall between 80 Hz and 140 Hz as shown in Figure 2.1. In comparison, irrespective of the data of F4, the F_0 contours of Yin Ping of the female speakers fall between 130 Hz and 200 Hz as shown in Figure 2.2. A 2-tailed t-test has been run for each time point throughout the overall duration of Yin Ping to examine whether there is any significant difference in F_0 values between male and female speakers. The results display that the difference in F_0 values between the two groups is significant at .05 level ($p < .000$, df = 6) when the data of M2 and F4 were excluded. That is, the mean F_0

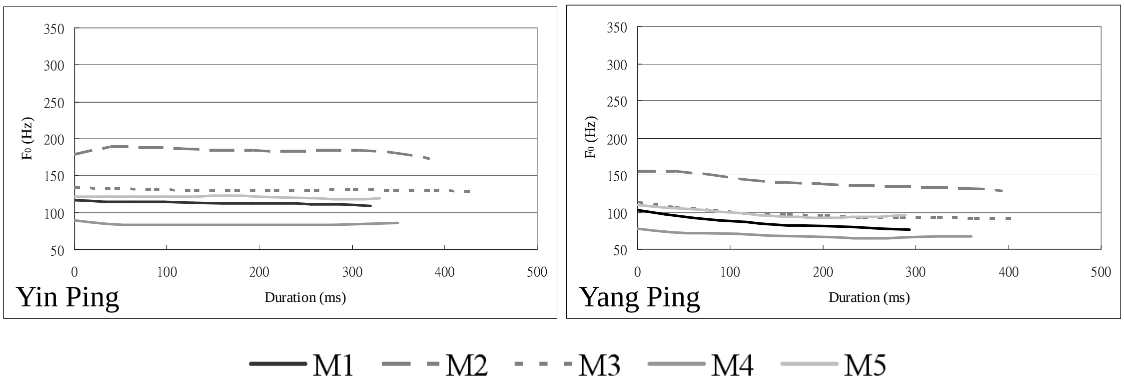


Figure 2.1 Mean F_0 contours of Yin Ping and Yang Ping in Meixian Hakka produced by five male speakers (in absolute duration).

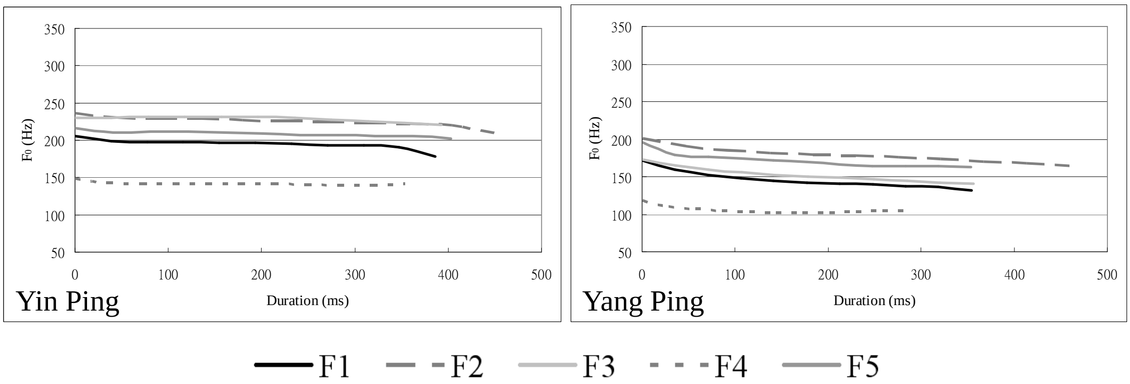


Figure 2.2 Mean F_0 contours of Yin Ping and Yang Ping in Meixian Hakka produced by five female speakers (in absolute duration).

values of Yin Ping of the female speakers are significantly higher than those of the male speaker for all time points. It should be noted that the comparatively high F_0 contour of M2 and the comparatively low F_0 contour of F4 can not be explained in this study. The observation in this study regarding the significant difference in mean F_0 values between male and female speakers is consistent with the conclusion of Traunmüller and Eriksson (1995). Based on the F_0 data from the previous studies investigating various languages, Traunmüller and Eriksson (1995) compared the mean F_0 of speakers who were in different sex and age. They concluded that in most studies the mean F_0 was higher and the F_0 range in Hz was wider for women than for men. Moreover, they declared that F_0 varied with the change of age and the changes in elderly voice may be resulted from the change in testosterone-oestrogen ratio of an individual. They found that F_0 for male speakers began to increase when men were at around 55 years old, while F_0 for female speakers lowered when women were at around 70 years old. However, such aged-related differences in F_0 do not provide an adequate explanation for the relatively high F_0 contour for M2 and relatively low F_0 contour for F4 as most subjects involved in this study were at their 60s at the time of recording and it is supposed that the changes in F_0 as a function of changes in age would at least apply to most of the subjects but not only to M2 and M4.

The invariant acoustic properties of Yin Ping are clearly displayed when expressed in normalized values. The normalized F_0 contours of Yin Ping for the male and female speakers are shown in Figure 2.3. The normalized values were calculated using logarithmic z-score transform. As Figure 2.3 shows, the individual F_0 contours of Yin Ping cluster together more than those in Figures 2.1 and 2.2. The shapes of the contours indicate that Yin Ping is a level tone. The contour of Yin Ping remains stable from the 10% to 90% time points for all speakers. The F_0 contours lie around the log z-score of 0

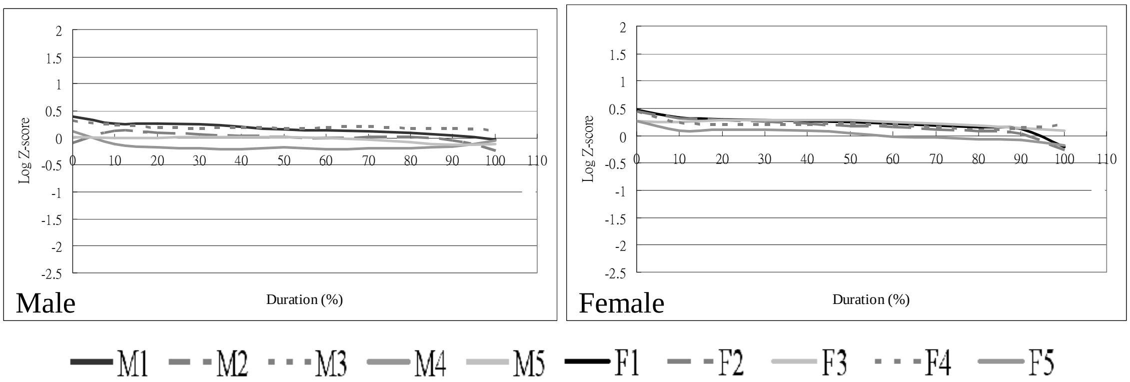


Figure 2.3 Normalized F_0 contours of Yin Ping in Meixian Hakka produced by five male and five female speakers.

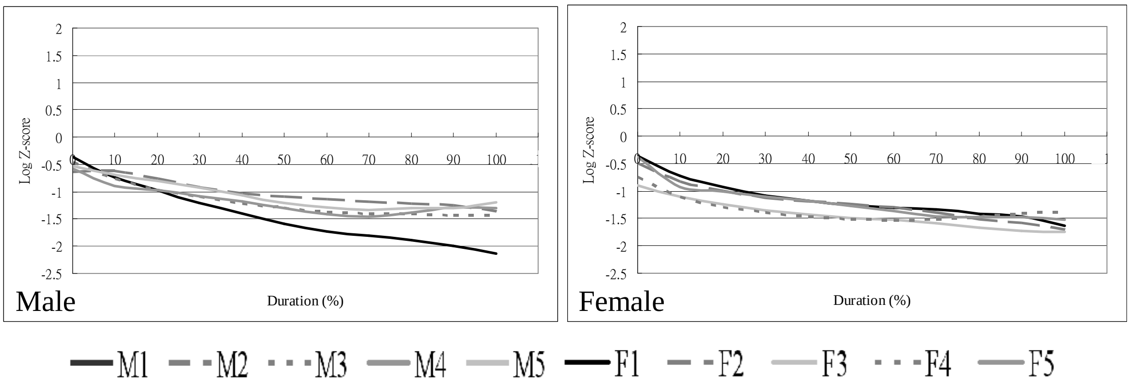


Figure 2.4 Normalized F_0 contours of Yang Ping in Meixian Hakka produced by five male and five female speakers.

which represents the mean of all data points of each individual speaker. Though Yin Ping and Yang Ping are classified as two registers of the same tone category, the acoustic realization of Yang Ping is different from that of Ying Ping in modern Meixian Hakka. As illustrated in Figures 2.1 and 2.2, Yang Ping has a slight fall throughout the overall duration of the syllable for both male and female speakers. It can be noticed that the F_0 values of M2's Yin Ping and Yang Ping are higher than those of other male speakers', whereas the F_0 values of F4's Yin Ping and Yang Ping are lower than those of other female speakers'. Regardless of the data of M2 and F4, the onsets of Yang Ping range from 77 Hz to 113 Hz when the tone bearing units were produced by the male speakers. The onsets range from 172 Hz to 200 Hz when the tone bearing units were produced by the female speakers. The offsets of Yang Ping range from 67 Hz to 95 Hz with respect to the male voices, whereas the offsets range from 131 Hz to 164 Hz with respect to the female voices. As shown in Figure 2.4, the normalized F_0 contours of Yang Ping cluster together as in the case of Yin Ping. In terms of the normalized F_0 values, the onsets of Yang Ping lie at around the log value of -0.5 and the offsets lie at around the log value of -1.5 for most speakers. Independent-samples t-test was used to see whether there is any significant difference in onset and offset values of Yang Ping between male and female speakers. The results indicate that there is no significant difference in onset value between the two groups at .05 level, $t(6) = 0.505$, $p < .633$. In contrast, the offset values of all male speakers except M1 tend to be higher than those of the female speakers. Regardless of the data of M1, the difference between male and female speakers is shown to be significant at .05 level [2-tailed t-test, $t(7) = 3.28$, $p < .013$]. The low offset of Yang Ping of M1 may due to individual difference and it can not be explained in this study. As the number of subjects seems not to be large enough to make definitive conclusions, further investigations with more subjects may be necessary. Based on the normalized data of Yang Ping in this study, male and female speakers have a similar onset value and the difference between the two sex groups is the offset value. As female speakers tend to have lower offset values than male speakers, Yang Ping in female voice tends to have a relatively steeper F_0 contour.

In summary, Yin Ping and Yang Ping differ in F_0 height and contour shape as displayed in Figures 2.3 and 2.4. A paired-samples t-test reveals significant differences in mean F_0 between Yin Ping and Yang Ping for all time points [2-tailed t-test, $p < .000$, $df = 9$]. The mean F_0 of Yin Ping for each time point is significantly higher than that of Yang Ping. Regarding the contour shape, the shapes of Yin Ping are substantially kept unchanged for the overall duration, whereas Yang Ping is characterized by a slight drop in F_0 at the low frequency range. The findings of this study indicate that the shapes of F_0 contour of male speakers and female speakers look similar in most cases. A paired-samples t-test displays a significant difference in duration between Yin Ping and Yang Ping at 0.5 level, $t(9) = 2.887$, $p < .018$. This shows that the average duration of Yin Ping ($m = 379.7$) is significantly longer than the average duration of Yang Ping ($m = 355.1$). Moreover, the positive correlation is significant ($r = .88$. $p < .001$) indicating that speakers have a long Ying Ping also tend to have a long Yang Ping.

b. Shang Sheng and Qu Sheng

In addition to Yin Ping and Yang Ping, Shang Sheng and Qu Sheng are considered as the long tones in Meixian Hakka. Figures 2.5 and 2.6 display the mean F_0 contours of Shang Sheng and Qu Sheng produced by five female and five male speakers, respectively. Similar to the case found in the analysis of Yin Ping and Yang Ping, tone contours produced by M2 are at a higher F_0 range than those produced by other male speakers, while tone contours produced by F4 are at a lower range than those produced by other female speakers. This also holds true in the analysis of Yin Ru and Yang Ru, of which the results are exhibited in the later section. As shown in the figures, Shang Sheng and Qu Sheng are realized by noticeable falling contours and the results are consistent for both male and female speakers. In terms of fundamental frequency, Shang Sheng and Qu Sheng differ in the contour shape and tone height. Shang Sheng has a falling contour that drops steadily from the onset towards the offset. Regardless of the data of M2 and F4, the onsets range from 99 Hz to 143 Hz for the male speakers, while the values range from 210 Hz to 242 Hz for the female speakers. The offsets range from 63 Hz to 85 Hz for the male speakers, while they range from 140 Hz to 159 Hz for the female speakers. The smallest F_0 range of Shang Sheng is 36 Hz, which was produced by M4, and the largest F_0 range is 109 Hz, which was produced by M2. In contrast, two types of F_0 contours for Qu Sheng can be seen in Figures 2.5 and 2.6. The F_0 contours of Qu Sheng for M2, M3, M5, F2, F3 and F5 are delayed falling tones. That is, the

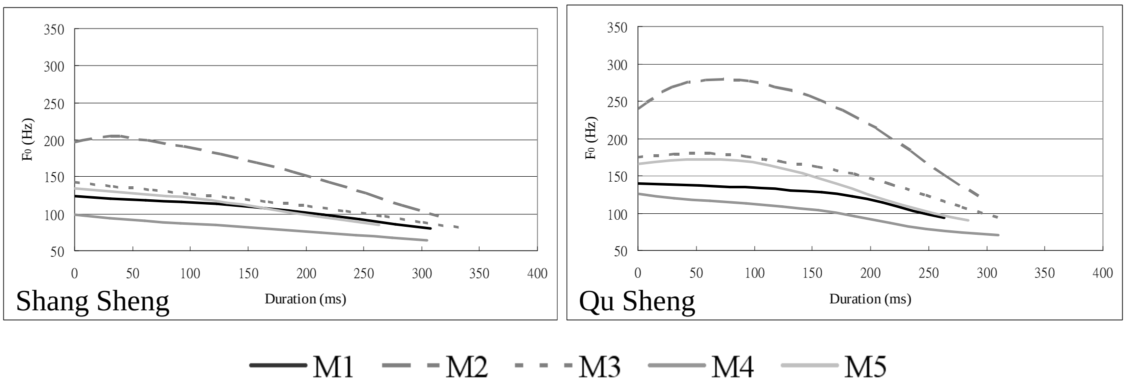


Figure 2.5 Mean F_0 contours of Shang Sheng and Qu Sheng in Meixian Hakka produced by five male speakers (in absolute duration).

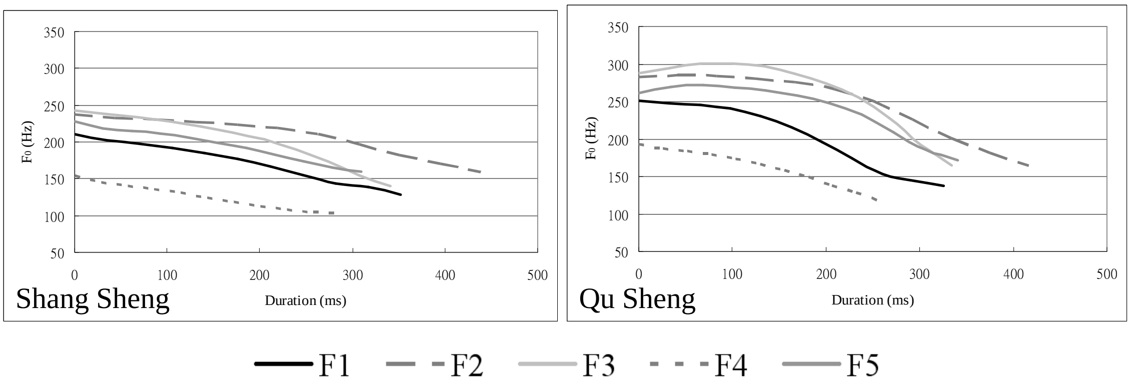


Figure 2.6 Mean F_0 contours of Shang Sheng and Qu Sheng in Meixian Hakka produced by five female speakers (in absolute duration).

contour of Qu Sheng has a rising or steady onset and a plateau at the early part of the duration. On the other hand, the F_0 contours of Qu Sheng for M1, M4, F1, and F4 are simple falling tones which are similar in shape to those of Shang Sheng. The two F_0 contours of Qu Sheng in this study are similar to the results of Tone 1 mentioned in the study of Zhu (1994). In his study of the Shanghai tones, there were also two types of F_0 contours for Tone 1 which were the ‘straight falling’ tone and the ‘delayed falling’ tone. Statistical analysis was conducted and he concluded that the two contours of Tone 1 were associated with sex. However, the results of this study seem not to support the conclusion of Zhu (1994) as nearly half of the speakers (M1, M4, F1 and F4) in this study in both sex groups have a simple falling contour for Qu Sheng. It is not conspicuous whether a simple falling contour for Qu Sheng is favorable for male or female speakers in Meixian Hakka. Further large-scale experiments are necessary to reach a conclusion.

Paired-samples t-test has been conducted to compare the F_0 values at each time point of Shang Sheng and Qu Sheng. Results reveal a significant difference in F_0 value between Shang Sheng and Qu Sheng at 0.5 level [2-tailed t-test, $p < .001$, $df = 9$] throughout the whole duration of a syllable. This indicates that the overall F_0 contour of Qu Sheng is significantly higher than that of Shang Sheng. Nonetheless, the differences in F_0 value between Shang Sheng and Qu Sheng do not remain constant over time. As shown in Figures 2.5 and 2.6, Shang Sheng is realized as a simple falling tone whereas the F_0 contour of Qu Sheng has an initial rise at the early part of the duration and falls towards the offset value after the point of inflexion. The difference in mean F_0 between Shang Sheng and Qu Sheng is 35 Hz at the onset. Such difference increases from the 0% to 40% time points. The difference is largest (50.5 Hz) and becomes stable at the 30% time point. It becomes smaller at the later part of the duration from the 40% to 100% time points. The difference in mean F_0 between the two tones is smallest at the offset (13 Hz).

The normalized F_0 contours of Shang Sheng and Qu Sheng are shown in Figures 2.7 and 2.8. The results of the independent-samples t-test indicates that there is no significant difference in log value between male and female speakers for all time points of Shang Sheng [$p < .99$, $df = 8$] and Qu Sheng [$p < .90$, $df = 8$], respectively. In terms of the normalized F_0 values, the onsets of Shang Sheng for both male and female speakers lie at around the log value of 0.5 ($m = 0.54$) and the offsets for all speakers except M2 lie between the log value of -1.47 and -2.1 ($m = -1.83$). In comparison to Shang Sheng, the onsets of Qu Sheng are higher and the log values of the 0% ($m = 1.47$) to 20% ($m = 1.52$) of the time points are around 1.5. The offset of Qu Sheng is the

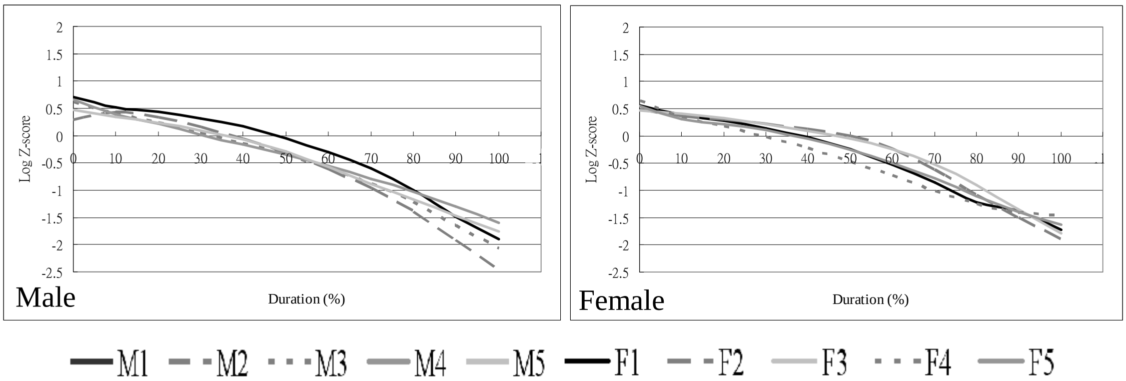


Figure 2.7 Normalized F_0 contours of Shang Sheng in Meixian Hakka produced by five male and five female speakers.

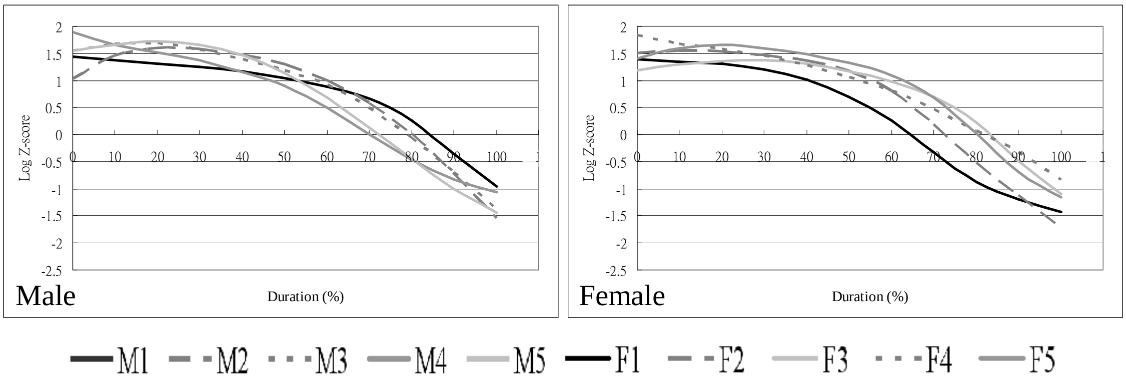


Figure 2.8 Normalized F_0 contours of Qu Sheng in Meixian Hakka produced by five male and five female speakers.

lowest point of the F_0 contour. Figure 2.8 shows that the offsets of Qu Sheng lay between the log value of -0.8 and -1.7. The average slopes of the two falling tones have been computed and compared in this study. The average slope of F_0 contour between the highest point and the lowest point, which was the offset in this study, was defined as the ratio of the change in frequency to the change in duration. A paired-samples t-test exhibits that Qu Sheng has a greater average slope than Shang Sheng and it is statistical significant [2-tailed t-test, $p < .001$, df = 9]. In respect of absolute duration, there is no significant difference between Shang Sheng, which has an average duration of 325.2 ms, and Qu Sheng, which has an average duration of 313.3 ms, in Meixian Hakka [2-tailed t-test, $p < .148$, df = 9]. In summary, the two falling tones in Meixian Hakka differ in tone height, contour shape and slope. Shang Sheng is a simple falling tone which drops from the mid F_0 range to the low F_0 range, while Qu Sheng is a delayed falling tone with a plateau at the early part of the F_0 contour and it drops from the high F_0 range to the low F_0 range. The slope of the F_0 contour of Qu Sheng is greater and the F_0 contour of Qu Sheng is also higher over time than those of Shang Sheng.

c. Yin Ru and Yang Ru

There are two short tones in Meixian Hakka and they are Yin Ru and Yang Ru. The mean F_0 contours of the two short tones produced by five female and five male speakers, respectively, are exhibited in Figures 2.9 and 2.10. Similar to the case of the four long tones, acoustic results of this study show that M2 has higher F_0 contours for Yin Ru and Yang Ru than other male speakers do, and F4 has lower F_0 contours than other female speakers do. In spite of interspeaker variability in F_0, the shapes of F_0 contours for both male and female speakers are similar. As shown in the figures, the shape of the F_0 contour of Yin Ru is different from that of Yang Ru. Yin Ru is a simple falling tone, of which the F_0 value at the onset is the highest and the value at the offset is the lowest for most speakers. In terms of the F_0 range, the onset of Yin Ru lies between 109 Hz and 145 Hz and the offset lies between 72 Hz and 91 Hz for most male speakers except M2, whereas the onset lies between 214 Hz and 250 Hz and the offset lies between 152 Hz and 173 Hz for the female speakers except F4. On the other hand, Yang Ru is a level tone of which the onset falls between 108 Hz and 167 Hz for the male speakers except M2 and between 234 Hz and 292 Hz for the female speakers except F4. The F_0 contour of Yang Ru remains in the high frequency region stable throughout the overall duration.

A paired-samples t-test reveals a significant difference in average F_0 value between Yin Ru and Yang Ru for each time point [2-tailed t-test, $p < .000$, $df = 9$]. This indicates that the F_0 contour of Yang Ru is significantly higher than that of Yin Ru over time. With respect to average duration, Yin Ru and Yang Ru are shorter as compared to other

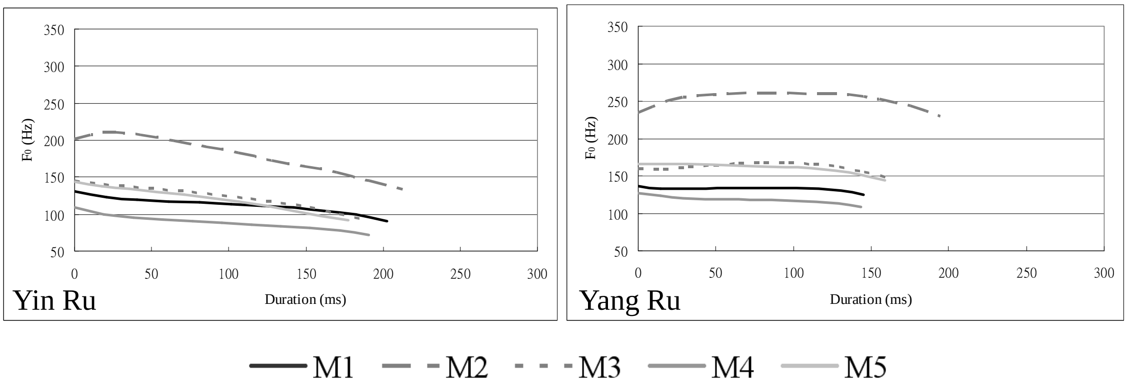


Figure 2.9 Mean F_0 contours of Yang Ru and Yin Ru in Meixian Hakka produced by five male speakers (in absolute duration).

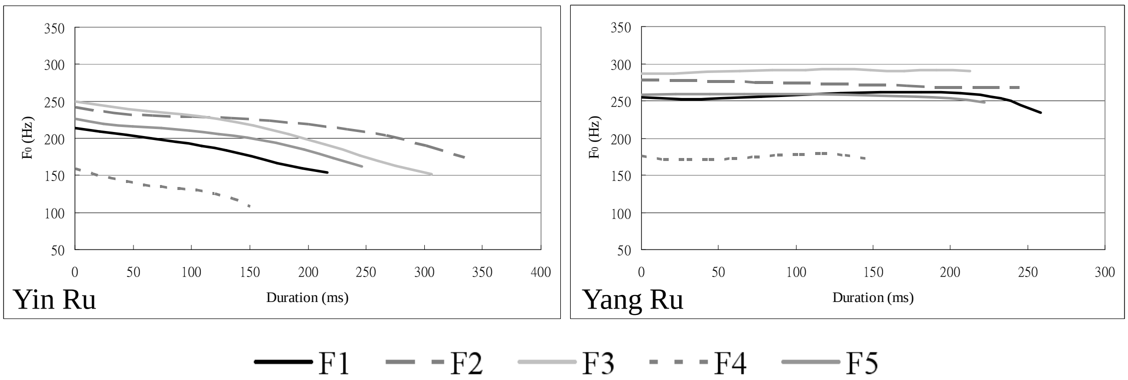


Figure 2.10 Mean F_0 contours of Yang Ru and Yin Ru in Meixian Hakka produced by five female speakers (in absolute duration).

tone categories in Meixian Hakka. Yin Ru has an average duration of 223 ms which is significantly longer than Yang Ru of which its average duration is 189 ms [paired-samples t-test, t (9) = -2.673, p < .025]. An independent t-test was also conducted to determine whether there is significant difference in F_0 value of the two Ru tones between different sex groups. Statistical analysis shows that the F_0 values of Yin Ru and Yang Ru for the female speakers are significantly higher than those for the male speakers at every time point of the F_0 contour even when the data of M2 and F4 were included [p < .033, df = 8]. An independent t-test further reveals that the difference in average duration between male and female speakers is significant at .05 level for the Yang Ru tone [t (8) = -2.616, p < .031], but not for the Yin Ru tone [p < .16]. The result indicates that the average duration of Yang Ru for the female speakers (m = 216.7) is significantly greater than that for the male speakers (m = 160.7). Though the average

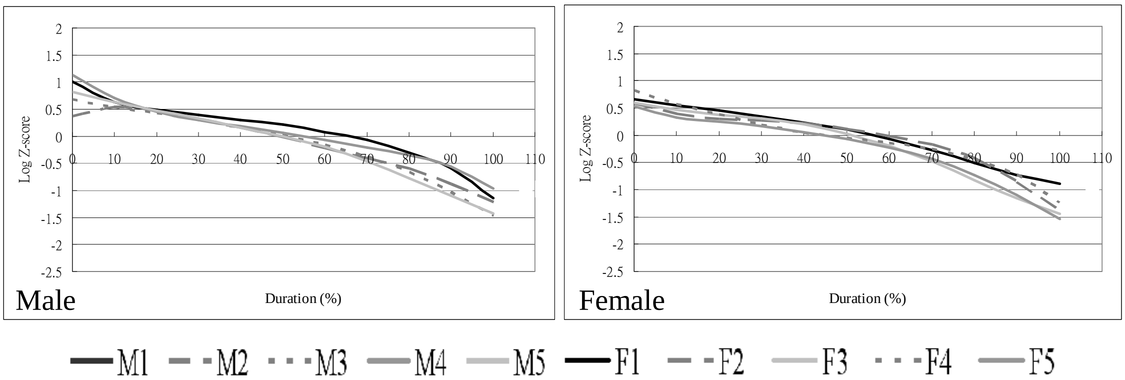


Figure 2.11 Normalized F_0 contours of Yin Ru in Meixian Hakka produced by five male and five female speakers.

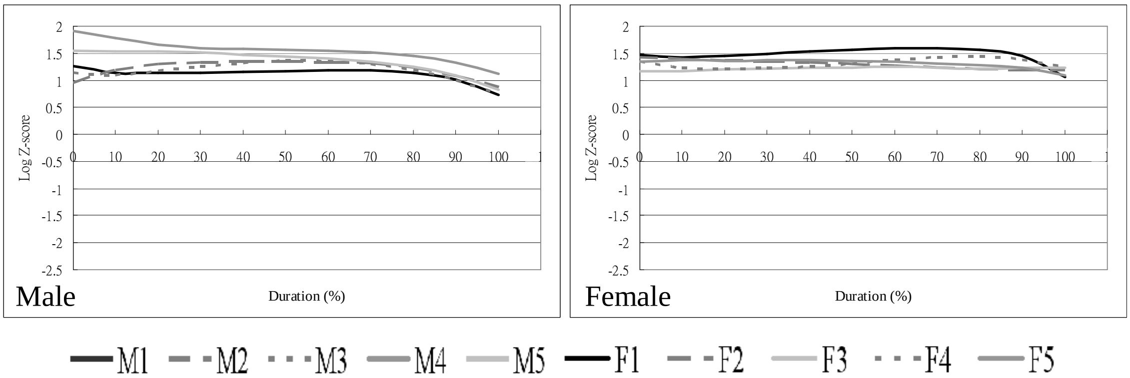


Figure 2.12 Normalized F_0 contours of Yang Ru in Meixian Hakka produced by five male and five female speakers.

duration of Yin Ru for the female speakers (m = 250.6) appears to be greater than that fro the male speakers (m = 194.7), this is not a statistically significant difference.

The normalized F_0 data of Yin Ru and Yang Ru are plotted on Figures 2.11 and 2.12. It is seen that the F_0 patterns shown for the male and female speakers are similar after normalization. The clustered F_0 contours display that Yin Ru is a falling tone of which the onset ranges from the log value of 0.36 to 1.13 and the offset ranges from the log value of -0.88 to -1.54, while Yin Ru is a level tone of which the F_0 contour remains stable between the log value of 0.73 and 1.91. In other words, the F_0 contour of Yin Ru starts at the middle region of an individual speaker's F_0 range and finishes at the low region, whereas the level contour of Yang Ru is at the high frequency region.

d. Tone values of the six citation tones

According to the findings of 2.3.1.1 - 3, the six Hakka tones can be classified into three F_0 patterns which are level, simple falling and delayed falling. Qu Sheng is the citation tone in Meixian Hakka which has a delayed falling contour. Yin Ping and Yang Ru have a level contour, whereas Yang Ping, Shang Sheng, and Yin Ru have a simple falling contour. In this section, the normalized F_0 contours of the six citation tones in Meixian Hakka are compared. A tone value is assigned for each of the six citation tones based on the normalized F_0 pattern. Figure 2.13 displays the mean normalized F_0 value as function of normalized duration for the six citation tones that produced by the ten speakers in this study. As shown in the figure, the F_0 contours of the six citation tones occur in the range between the log value of 1.52 and -1.83, and the overall range of the normalized F_0 values is 3.356 units. The mean onset value of Qu Sheng is the highest point of the range, whereas the mean offset value of Shang Sheng is the lowest point. It can be noticed that the onsets of Yang Ru and Qu Sheng are at the high frequency range; the onsets of Yin Ping, Shang Sheng and Yin Ru at the middle frequency range; and the onset of Yang Ping at the low-to-mid frequency range. That is, there are three citation tones in Meixian Hakka of which the F_0 contour starts at the middle frequency range. With regard to the location of the F_0 offset, the offsets of most tones in this study, which are Yang Ping, Shang Sheng, Qu Sheng and Yin Ru, occur at the low frequency range. Whereas Yang Ru and Yin Ping, which have a level F_0 contour, have an offset at the high frequency range and at the middle frequency range, respectively.

A one-way ANOVA is conducted to examine differences in absolute duration between tones. The statistical significance in duration is achieved at .05 level [$F(5, 45) = 91.82$, $p < .000$] The post hoc tests reveal significant differences in duration between the non-Ru tones and the Ru tones at .05 level and the means of the six tones are

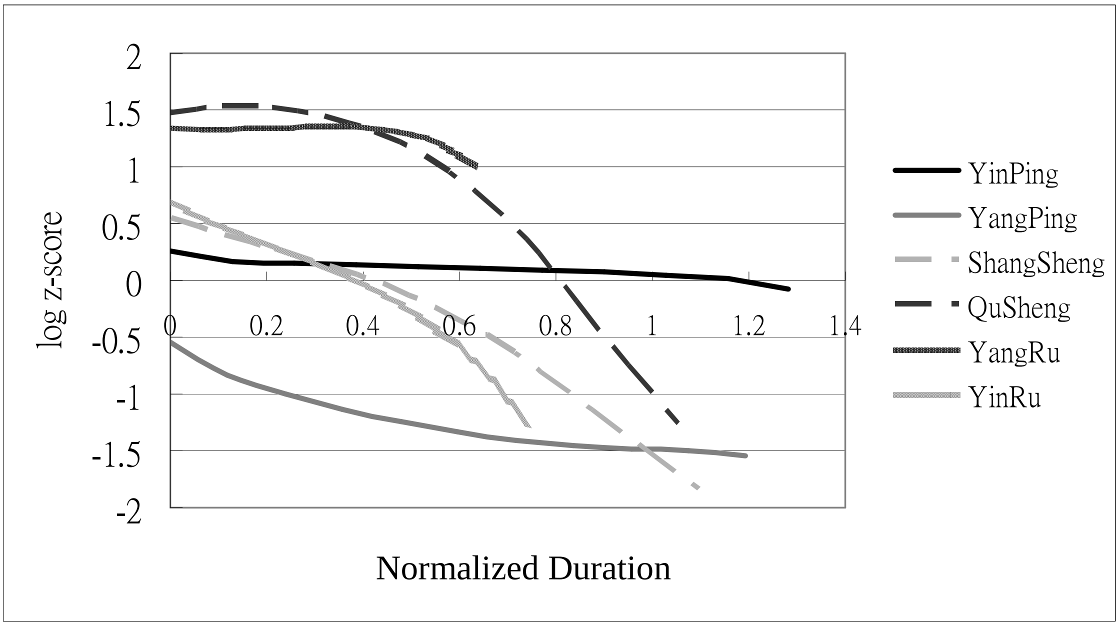


Figure 2.13 Mean normalized F_0 contours of the six citation tones in Meixian Hakka based on the data of ten speakers.

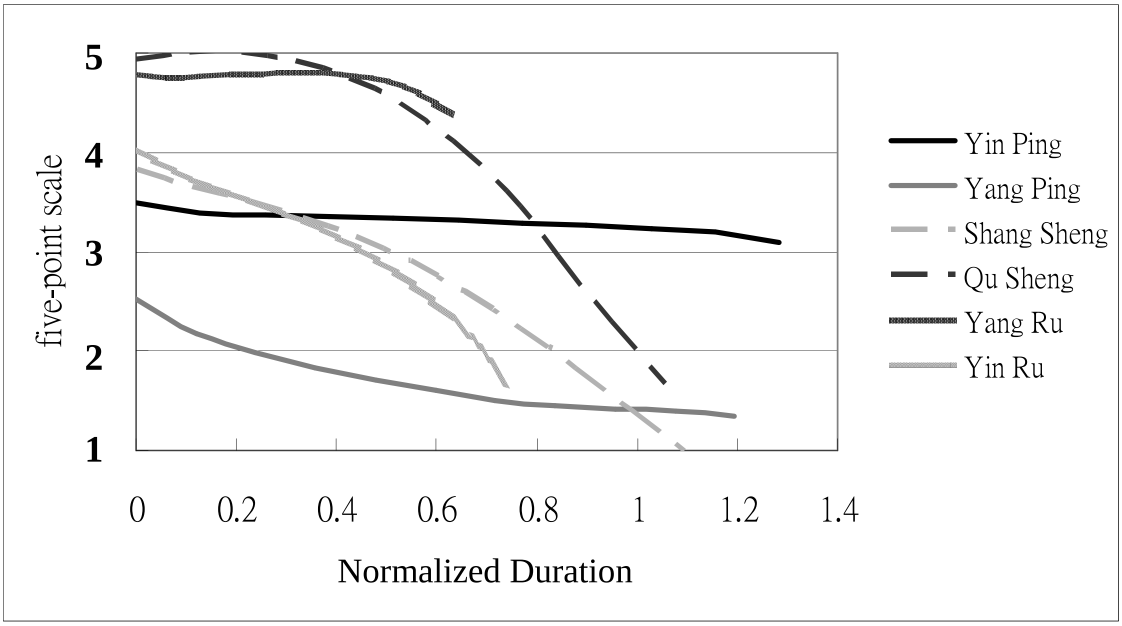


Figure 2.14 Mean normalized F_0 contours of the six citation tones in Meixian Hakka in terms of five-point scale.

significantly different from each other in general. Thus, the duration of Yin Ping (m =379.7), Yang Ping (m = 355.1), Shang Sheng (m = 325.2) and Qu Sheng (m = 313.3) is significantly greater than that of Yin Ru (m = 188.7) and Yang Ru (m = 222.7). Among the non-Ru tones, Yin Ping is longest in duration and its duration is significantly greater than that of Shang Sheng [$p < .004$] and Qu Sheng [$p < .001$]. The duration of Yang

Ping is significantly greater than that of Qu Sheng [$p < .042$], but is not significantly different from that of Yin Ping [$p < .270$] and Shang Sheng [$p < .224$]. It should be noted that there is no significant difference in duration between Shang Sheng and Qu Sheng [$p < 1.00$]. Regarding the Ru tones, the difference in duration between Yin Ru and Yang Ru is not significant [$p < .382$]. It can be concluded that, in descending order, Yin Ping is the longest tone, followed by Yang Ping, Shang Sheng, Qu Sheng, Yang Ru and Yin Ru. Citation tones that belong to the tone category of Ping are showed to have a longer F_0 contour than those that belong to other categories do, whereas citation tones that belong to the tone category of Ru have a shorter F_0 contour in comparison. This finding can be observed in Figure 2.13 which exhibits the normalized duration of the six citation tones. This may suggest that, in addition to the difference in contour shape between tones, duration may also play a role in the perception of tones in Meixian Hakka. However, it should be noted that the temporal pattern of tones is not constant for all speakers in this study and interspeaker variation in duration was observed in the acoustic data. For example, Yin Ping is shown to be the longest tone for the majority of speakers, but not for all. In addition, there are two types of temporal patterns for Yang Ping. In terms of absolute duration, on one hand, the mean duration of Yang Ping is similar to that of Yin Ping for some speakers (F2, M2, M3 and M4) and the two Ping tones are longer than the non-Ping tones. On the other hand, Yang Ping is similar in duration to Shang Sheng and/or Qu Sheng for other speakers and it is smaller in duration than Yin Ping.

Figure 2.14 is derived from Figure 2.13 in which the mean normalized F_0 contours of the six citation tones are shown. Instead of the calculated log z-score, the six citation tones in Meixian Hakka are displayed in terms of five-point scale in Figure 2.14. It can be seen that the overall range of the log values is divided into four equal intervals. The number of 5 is assigned to the highest level of the frequency range; the number of 4 to the mid-high level; the number of 3 to the middle level; the number of 2 to the mid-low level; and the number of 1 is assigned to the lowest level. The difference between every two levels is 0.839 units. As shown in Figure 2.14, Yin Ping occupies the middle frequency range of an individual speaker. It has a level contour which starts at the mid point between 3 and 4 and remains fairly stable with a slight fall towards its end. Yin Ping goes into the mid register and should be considered as a mid-level tone. The contour of Yin Ping is represented as [33]. In contrast, the F_0 contour of Yang Ping occupies the low frequency range. It begins at the mid-low point between 2 and 3 and falls gently towards the low point between 1 and 2. Its contour is falling and should be represented as [21] based on the phonetic data. As stated in the study of Yip (2001), the lowest tone in the system was always realized as a slight falling contour in natural speech. Yip (2007) claimed that the falling portion of a low tone was transitional which allowed the F_0 values to drop to the lowest level of the range. In consideration of phonological tone distribution, it was suggested that the phonetically falling tone, as in [21], should be considered phonologically as level tone [11] with a slight fall (Chen, 2000; Yip, 2001; 2007). This study follows the view of Yip (2001; 2007). Yang Ping, which is the lowest tone in Meixian Hakka, is represented as low-level tone [11]. Shang Sheng has a falling contour that starts at a point which is slightly lower than the mid-high point of 4. The F_0 contour falls all the way to the bottom of speaker's frequency range. The contour of Shang Sheng is transcribed as the tone letter of [41]. Yin Ru has a falling F_0 contour which starts at the mid-high point of 4 and finishes at the low point between 1 and 2. A comparison of F_0 contour between Yin Ru and Shang Sheng displays that the F_0 contour of Yin Ru is significantly shorter in duration than that of Shang Sheng. This can be explained by the difference in syllable structure between the two tones. Ru tones in Meixian Hakka only occur in closed syllable, whereas Shang Sheng occur in open syllable and syllable with nasal endings. The shape and direction of movement of the contour of Yin Ru before its 70% time point sounds very much like those of the contour of Shang Sheng before its 50% time point as shown in Figure 2.14. The difference in F_0 contour between the two tones is displayed in the second half of the overall duration, for instance, the location of F_0 offset. The offset of Yin Ru, which is near the low point between 1 and 2, is phonetically higher than that of Shang Sheng, which is at the bottom of the frequency range. Nonetheless, it can be observed that Yin Ru ends at nearly the same low point which is between 1 and 2 as Yang Ping. The contour of Yin Ru may be phonologically transcribed as [<u>41</u>] which is a short tone. Qu Sheng is a high to low falling tone as shown in Figure 2.14. The tone contour of Qu Sheng starts at the highest level of the frequency range and it remains fairly level before the 20% time point of the overall duration. It then falls sharply toward its end. The offset of Qu Sheng is at the same low point as Yin Ru. The contour of Qu Sheng is represented as [51]. Besides Yin Ru, Yang Ru is another short tone in Meixian Hakka and it is the shortest one in the system for all speakers. Yang Ping is basically a level tone. The tone contour of Yang Ru starts at a point that is slightly lower than the high point of 5. It remains stable throughout the overall duration and there is a slight drop toward its end at the 80% time point. The overall contour occupies the top of the high register and its contour is represented as [<u>55</u>] or simply [5].

Table 2.3 displays the tone letter assigned for each citation tone in this study based on the normalized F_0 contours and the corresponding tone letter or descriptions described in the past studies. Yin Ping was described as [44] in the past studies and its contour is represented as the tone letter of [33] in this study based on the acoustic data. The transcriptions both display that Yin Ping in Meixian Hakka occupies the mid

Historical Tonal Category	This study	Norman (1988)	Huang (1992)	Xie (1994)	Yuan et al. (2001)	Hashimoto (1973)
陰平 Yin Ping	[33]	[44]	[44]	[44]	[44]	mid level legato
陽平 Yang Ping	[11]	[12]	[11]	[11]	[11]	low level (a little falling) legato
上聲 Shang Sheng	[41]	[31]	[31]	[31]	[31]	mid falling legato
去聲 Qu Sheng	[51]	[42]	[53]	[52]	[52] or [42]	high level (a little rising) legato
陽入 Yang Ru	[55] or [5]	[44]	[5]	[5]	[4] or [5]	high rising staccato
陰入 Yin Ru	[41]	[21]	[1]	[1]	[21] or [1]	mid falling staccato

Table 2.3 Comparison of the tone letters assigned to the six citation tones in Meixian Hakka between this study on the basis of normalized F_0 data and the descriptions of the past studies.

register and appears to be phonologically level. The transcription of Yin Ping suggested in this study agrees with the description in Hashimoto's study (1973). In the study of Hashimoto, acoustic data of the six tones in Hakka were obtained through "pitch-stress recorder". Based on the data of a single speaker, Yin Ping was transcribed as a mid level legato tone (1973:104). In contrast, it was transcribed as [44] in the studies of Huang (1992), Xie (1994) and Yuan et al. (2001). The difference in transcription between studies may be resulted from the lack of acoustic data as reference and the subjectivity of authors' perceptual impression. It may also due to the effects of phonetic context on tone production. Taking these factors into consideration, the one digit difference in tone letter between this study and other past studies may not be phonologically relevant. Yang Ping is phonetically realized as a low tone with a slight drop. It is suggested that the slight fall of the tone may not be phonologically relevant and Yang Ping is represented as [11] in this study. In comparison, the results of this study do not agree with the transcription of Yang Ping in the study of Norman (1988). In his study, Yang Ping was transcribed as a low rising tone [12]. The phonetic results for all speakers in this study display that Yang Ping has either a level contour or a slight falling contour. Nonetheless, no rising contour is found in the acoustic data. Despite the difference in transcription between this study and the study of Norman, it is generally accepted that Yang Ping is a tone in the low register and most of the studies agree that Yang Ping is phonologically a level tone.

The transcription of Shang Sheng in this study is nearly identical to that of the past studies. Shang Sheng was transcribed as the mid to low falling tone in Meixian Hakka. The tone was represented as [31] in the past studies, whereas it is transcribed as [41] in this study. It can be noticed that the difference in transcription between this study and the past studies is the starting pitch. As shown in Figure 2.14, the starting point of the F_0 contour of Shang Sheng is significantly higher than that of Yin Ping [paired-samples t-test, $t(9) = -6.028$, $p < .000$]. This holds true for the absolute frequency value of the two tones, $t(9) = -6.116$, $p < .000$. If the starting points of the two tones are phonologically different, it would be more reasonable to represent Yin Ping and Shang Sheng as the tone letter of [33] and [41], respectively. Otherwise, Yin Ping and Shang Sheng would be transcribed as [44] and [41], respectively. Nonetheless, the transcription of Shang Sheng and Yin Ping described in the past studies is not consistent with the obtained phonetic data in this study. It was believed in the past studies that the starting point of Shang Sheng was lower than that of Yin Ping. There is general consensus that Qu Sheng in Meixian Hakka has a high falling contour. Most of the studies considered Qu Sheng the falling tone that starts at the high register and ends at the low register. The tone was transcribed as either [52] or [42] by many linguists as shown in Table 2.3 (Norman, 1988; Xie, 1994; Yuan et al., 2001). The transcription of this study and that of Huang

(1992) are different in the ending pitch. Huang suggested that Qu Sheng falls from the highest point of the frequency range to a point at the mid register and it was represented as [53]. In contrast, the tone is transcribed as a high-to-low falling tone [51] in this study. It can be seen in Figure 2.14 that the onset of F_0 contour of Qu Sheng is always at the highest frequency range and the tone value of [5] is assigned to the starting pitch. On the other hand, the offset of F_0 contour of Qu Sheng is at the low register and it is at almost the same tone level as that of Yang Ping and Yin Ru. The lowest value of a five-point scale, i.e. 1, is assigned to the ending point. It can be noticed that Qu Sheng was transcribed as high level legato tone by Hashimoto (1973). The transcription in Hashimoto's study is different from that in other studies. The former considered Qu Sheng as phonologically level, whereas the latter considered the tone as phonologically falling. The discrepancy in tonal shape between studies may be explained by the tone sandhi phenomenon in Meixian Hakka. In two-syllable structure, the citation form of Qu Sheng, which is the high falling tone, is realized as its sandhi form, which is the high level tone, when it is followed by Yang Ping, Shang Sheng, Qu Sheng or Yin Ru (Cheung, 2006; Lü, 2006). In the study of Hashimoto, the six tones were pronounced one after another in individual utterances. As Yin Ru was preceded by Qu Sheng in the sequence, the sandhi environment of Qu Sheng seemed to be achieved and Qu Sheng would be realized as its sandhi form, i.e., high level tone [55].

The transcription of Yang Ru in this study is consistent with that of the past studies. The phonetic data in this study displays that Yang Ru is realized as a level tone at the high register. It is considered as short tone as the duration of Yang Ru is shorter than that of the non-Ru tones. The tone is represented as the tone letter of [<u>55</u>] or [5]. Norman (1988) suggested that Yang Ru and Yin Ping were transcribed as mid-high level tone [44]. The transcription proposed by Norman (1988) is not supported by the phonetic data in this study. It can be seen in Figure 2.14 that the level tone of Yang Ru occupies the high frequency range and it is 1.5 digits higher than the level tone of Yin Ping, which remains stable within the middle range, on the five-point scale. That is, Yang Ru is supposed to be phonological higher in tone level than Yin Ping in Meixian Hakka. The transcriptions of Yin Ru in the past studies indicated that Yin Ru was a short tone with a slight fall at the low register. The contour of Yin Ru was represented as either [21] or [1] by most linguists (Huang, 1992; Norman, 1988; Xie, 1994; Yuan et al., 2001). The transcriptions in these studies are different from the transcriptions in this study and in the study of Hashimoto (1973). Hashimoto stated that Yin Ru was a mid falling staccato of which the F_0 contour falls form a point at the mid register to a point at the low register. It showed that Yin Ru was similar to Shang Sheng in starting and ending pitch but the two tones were different in duration. The results of this study are in general consistent with the descriptions in Hashimoto's study (1973). As shown in Figure 2.14, Yin Ru is shorter in duration that Shang Sheng which is a non-Ru tone. Both Yin Ru and Shang Sheng start at the mid-high level and the tone letter of [4] is assigned to the starting points. The ending point of F_0 contour of Yin Ru is phonetically higher than that of Shang Sheng, but such difference may not be phonologically relevant as the ending points of the two tones are at the lowest frequency range. The offset of F_0 contour of Yin Ru is therefore represented as [1] and the contour of Yin Ru is transcribed as [41]. The phonetic data in this study demonstrated that it would be more rational to represent Yin Ru using 2 digits, one for the starting pitch and one for the ending pitch as the falling contour of Yin Ru is not restricted to the low register but drops from a mid-high point to a low point of the frequency range. In conclusion, there are historically six tonal categories in Meixian Hakka, which are Yin Ping, Yang Ping, Shang Sheng, Qu Sheng, Yang Ru and Yin Ru. A comparison of the transcription of this study based on the phonetic data with the transcriptions of the past studies indicates that the tone notations of Yin Ping, Yang Ping, Qu Sheng and Yang Ru are generally similar in different studies. The results of this study exhibit that the starting point of Shang Sheng is phonetically higher than that of Yin Ping and the difference in starting point between Yin Ping and Shang Sheng should be shown in the first digit of the tone letter. In addition, the results of this study demonstrate that the phonemic transcriptions of Yin Ru in the past studies do not correspond to the phonetic realization of the tone. It is suggested that two digits instead of one should be used to express the mid to low falling of the tone. Based on the phonetic data in this study, the six citation tones in Meixian Hakka are realized as (1) mid level tone [33], (2) low level tone with a slight drop [11], (3) mid-high to low falling tone [41], (4) high to low falling tone [51], (5) high level short tone [<u>55</u> or 5] and (6) mid-high to low falling short tone [<u>41</u>].

2.1.3.2 Comparison of F_0 contours of the citation tones on vowels /i/, /a/ and /u/

The phonetic data presented in the previous section consisted of average F_0 contours which computed based on tones that superimposed on syllables in different phonetic contexts. In this section, the variation in F_0 contour resulted from the difference between vowels are evaluated and the citation tones that superimposed on vowels /i/, /a/, and /u in CV(C) syllables are analyzed separately. A one-way ANOVA is conducted to determine whether there is significant difference in duration between tones that superimposed on vowels /i, a, u/. The statistical results show no significant difference in absolute duration between tones that superimposed on different vowels and this holds true for all six tones [$p > .05$]. Figures 2.15 – 2.17 display the average F_0 contours of the six citation tones which superimposed on syllables with vowels /i/, /a/ and /u/. The x-axis of the diagram displays the fundamental frequency values in Hertz while the y-axis

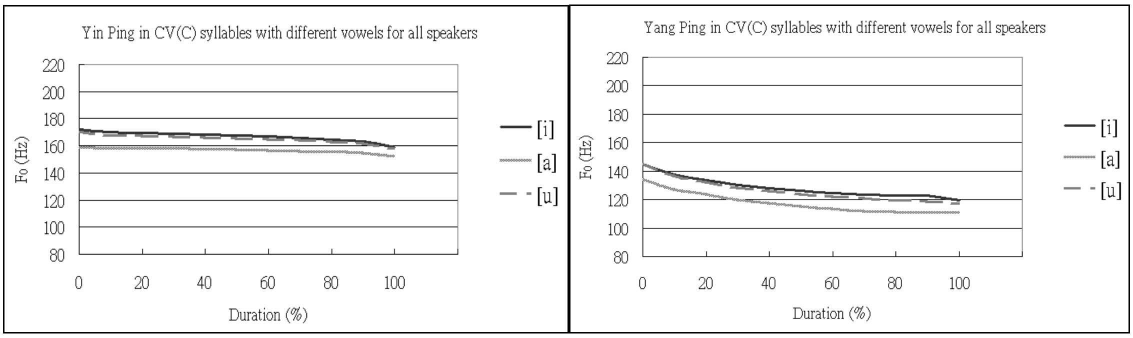


Figure 2.15 Mean F_0 contours of Yin Ping (left) and Yang Ping (right) which superimposed on syllables with vowels [i, a, u] produced by ten Meixian Hakka native speakers

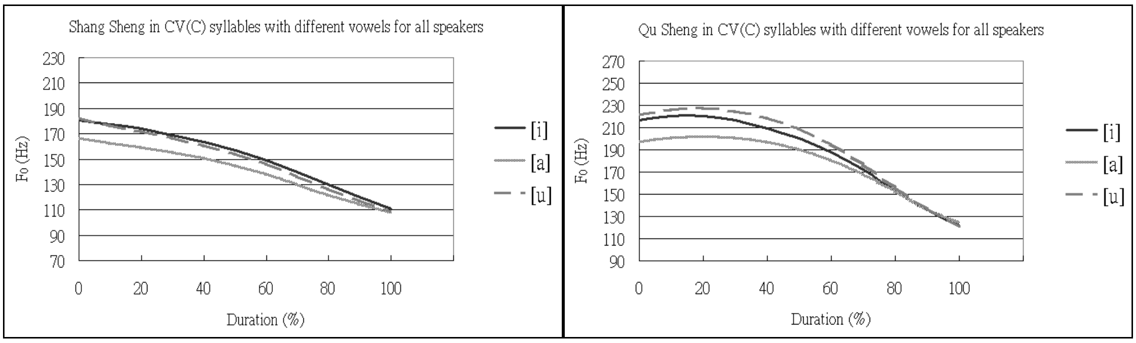


Figure 2.16 Mean F_0 contours of Shang Sheng (left) and Qu Sheng (right) which superimposed on syllables with vowels [i, a, u] produced by ten Meixian Hakka native speakers

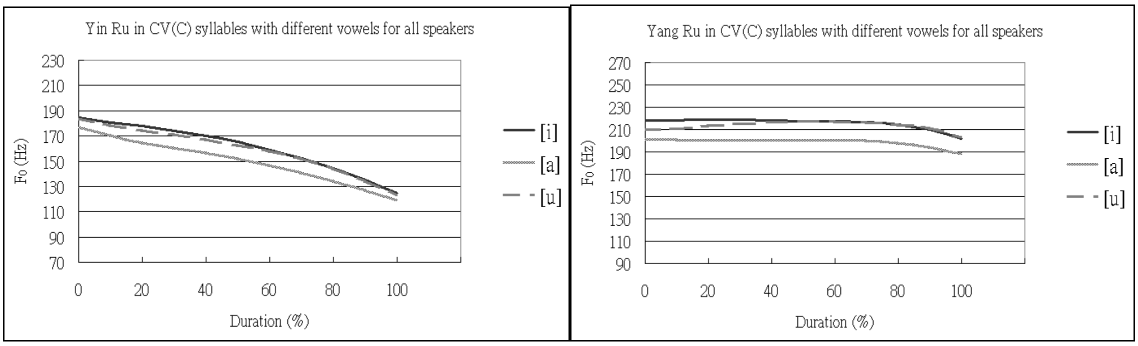


Figure 2.17 Mean F_0 contours of Yin Ru (left) and Yang Ru (right) which superimposed on syllables with vowels [i, a, u] produced by ten Meixian Hakka native speakers

displays the relative duration in percentage. The average F_0 contours of the citation tones on different vowels, i.e., /i, a, u/, are represented in different curves on the diagram. As shown in the figures, the average F_0 contours of each of the six citation tones on different vowels in CV(C) syllables are similar in shape. The similarity of F_0 contour shape between different vowels is also found across speakers as shown in

Appendix B. Appendix B exhibits the average F_0 contours of the six citation tones in different phonetic contexts which consist of (a) tones on vowel /i/ in V(C) syllables, (b) tones on vowel /i/ in CV(C) syllables, (c) tones on vowel /a/ in CV(C) syllables and (d) tones on vowel /u/ in CV(C) syllables, for each individual speaker. Regardless of individual differences, it is conspicuous that the F_0 contour on vowel /i/ generally overlaps with the F_0 contour on vowel /u/ and the F_0 contours of the six citation tones on vowels /i, u/ are higher than those on vowel /a/. In other words, the F_0 contours of the citation tones are lower on low vowel than on high vowels and such interaction between F_0 and vowel quality is related to vowel height. The results of a series of paired-samples t-tests for the six citation tones on the high vowels against the low vowels are shown in Table 2.4. Results show that the vowel intrinsic fundamental frequency (iF0) is found across tones but the iF0 difference varies throughout the overall duration depending on the contour shape of the tone.

Figure 2.15 shows the relative constant iF0 difference for Yin Ping and Yang Ping throughout the overall duration of the tone. Yin Ping is a mid level tone [33] while Yang Ping is a low level tone [11] in the phonological tone system of Meixian Hakka. It can be seen that the average F_0 curves of /i/ and /u/ overlap each other from the beginning to the end and the two curves are higher than the F_0 curves of /a/. The difference in F_0 between the high vowels and the low vowel is around 10 Hz and such difference remains stable throughout the tone. The statistical results in Table 2.4 show that the mean F_0 values of the low vowel /a/ are significantly different from those of the high vowel /i/ and the combined high vowels /i, u/ at nearly all time points of Yin Ping and Yang Ping except at the end point of Yang Ping. The significant differences in F_0 between /i/ and /a/ are found from the 0% to 60% points for Yin Ping and there is no F_0 difference between /u/ and /a/ at most of the points for Yang Ping. The iF0 differences,

YinPing [33]	0%	10%	20%	30%	40%	50%	60%	70%	80%	90%	100%	df
/i/ vs /a/	***	***	**	**	**	**	**	*	*	**	*	9
/u/ vs /a/	***	**	*	*	*	*	*	~	~	~	~	9
/i, u/ vs /a/	***	***	**	**	**	**	**	*	*	*	*	9
Yang Ping [11]												
/i/ vs /a/	*	*	*	*	**	**	**	**	**	**	*	9
/u/ vs /a/	*	*	~	~	~	*	~	*	~	~	~	9
/i, u/ vs /a/	**	**	**	**	**	**	**	**	**	**	~	9
Shang Sheng [41]												
/i/ vs /a/	**	**	**	*	*	*	*	*	~	~	~	9
/u/ vs /a/	**	**	**	**	**	*	*	~	~	~	~	9
/i, u/ vs /a/	**	**	**	**	*	*	*	*	~	~	~	9
Qu Sheng [51]												
/i/ vs /a/	***	***	***	***	**	*	*	~	~	~	~	9
/u/ vs /a/	***	***	***	***	**	**	*	~	~	~	~	9
/i, u/ vs /a/	***	***	***	***	**	**	*	~	~	~	~	9
Yin Ru [41]												
/i/ vs /a/	*	*	**	**	**	***	***	***	***	**	~	9
/u/ vs /a/	*	*	*	**	**	**	**	**	**	*	~	9
/i, u/ vs /a/	*	*	**	**	**	***	**	**	**	*	~	9
Yang Ru [55]												
/i/ vs /a/	~	~	*	*	*	~	~	~	~	~	~	6
/u/ vs /a/	*	**	**	**	**	**	**	**	***	***	**	9
/i, u/ vs /a/	**	***	***	***	***	***	***	***	***	***	***	9

Table 2.4 Results of paired-samples t-tests comparing tones in different vowel contexts

however, occur at all time point data in Ying Ping and Yang Ping when the data of high vowels /i, u/ are combined to test against the low vowel /a/. The F_0 curve of /a/ is consistently lower than the curves of /i, u/ throughout the tone for Yin Ping and from the 0% to 90% points for Yang Ping.

In contrast, Shang Sheng and Qu Shang in Meixian Hakka are contrastive tones of the same direction and the two tones have a falling F_0 contour. As shown in Figure 2.16, high iF0 differences are observed throughout the tone, but with few exceptions from 80% point for Sheng Sheng and from the 70% point for Qu Sheng. In Shang Sheng, the F_0 curve of /a/ is lower than those of /i, u/ from the 0% to 80% time pionts. The F_0

differences between high and low vowels are greatest at the beginning (the 0% point) of the tone. The differences become smaller throughout the overall duration and the curves merge from the 90% point. The t-test results in Table 2.4 indicate that Shang Sheng on the high vowels is significantly higher than the tone on the low vowel from the 0% to 70% points. The F_0 differences between high and low vowels are significant at .01 level over the early part of the tone (from the 0% to 30% points) and at .05 level over the middle part (from the 40% to 70% points). There are no significant differences between high and low vowels from the 80% point at .05 level. The iF0 differences between high and low vowels in Qu Sheng are similar to the differences in Shang Sheng. Figure 2.16 displays that the F_0 contour of the tone on /i, u/ start around 20 Hz higher than the curve of /a/. The F_0 curve of Qu Sheng on /i/ at 0% is between that on /u/ and /a/. The F_0 curves on the three vowels then rise over the early part of Qu Sheng (from the 0% to 20% points) and there is a sharp fall over the remaining part of the tone (from the 20% to 100% points). The iF0 differences between high and low vowels remain relatively constant over the early part of Qu Sheng. The F_0 curves of /i/ and /u/ approach the curve of /a/ gradually from the 20% to 70% points and they merge at 70% point. The t-test results exhibit that high iF0 differences are found between the high vowels /i, u/ and the low vowel /a/ in Qu Sheng from the 0% to 30% points and the differences are significant at .001 level. The iF0 differences become smaller after the inflexion point at the 20% point. The tone on low vowel is significantly lower than that on high vowels at .05 level from the 40% to 60% point, but no iF0 differences are found between either /i/ or /u/ and /a/ from the 70% point.

Yin Ru and Yang Ru are short tones in Meixian Hakka. Yin Ru is acoustically realized as a mid-to-low falling tone while Yang Ru is a high level tone. Figure 2.17 shows that the average F_0 contour of Yin Ru on high vowel /i/ start at nearly the the same point as that on high vowel /u/ and the F_0 curves of Yin Ru on /i/ and /u/ overlap each other over most of the time points. The tone on the low vowel /a/ starts 7 Hz lower than that on the high vowels /i, u/. The difference in F_0 between high and low vowels increases at the 20% point and it is approximately 11 Hz from the 20% to 80% points. The differences are slightly smaller (around 5 Hz) at the offset of the tone. Table 2.4 displays that the iF0 differences between high and low vowels for Yin Ru are significant at .05 level throughout the tone, with one exception at the end of the tone (the 100% point). It indicates that Yin Ru on the high vowels /i/ and /u/ has an F_0 contour which is significantly higher than that on the low vowel /a/ from the beginning to the end of the tone. In contrast, the F_0 curve of Yang Ru on the vowel /u/ has an onset which is between that on /i/ and /a/. The curve on /u/ rises and merges with the curve on /i/ at the 30% point. The curves on /i, u/ is higher than that on /a/ from the 30% point. The iF0 differences between the high vowel /i/ and low vowel /a/ are around 17 Hz throughout the tone (from the 0% to 90% points). The differences between the high vowel /u/ and the low vowel /a/ are less than 13 Hz over the early part of the tone (from the 0% to 20% points) and around 17 Hz from the 30% to 90% points. The t-test results in Table 2.4 show significant iF0 differences between high and low vowels for all time points at .01 level when the average F_0 values for either /u/ or combined /i, u/ are tested against the corresponding values for /a/. The iF0 differences between /i/ and /a/ are not significant at .05 level throughout the overall duration, with the exceptions of the 20% to 40% points. It differs from the results of other tones which exhibit the significant iF0 differences between /i/ and /a/.

Results of this study confirm the existence of F_0 differences between high and low vowels in tone languages. The acoustic data and statistical results indicate that tones on high vowels such as /i/ and /u/ tend to have higher F_0 than those on the low vowels such as /a/. The iF0 differences are found in all Meixian Hakka tones which differ from each other in tone height, contour shape and duration. The correlation between F_0 and vowel height is not only limited to contour tone languages such as Meixian Hakka and Shanghai Chinese (Zhu, 1999), but has also been attested in various languages including register tone languages, for example, Yoruba, and non-tone languages such as English and Danish (Whalen & Levitt, 1995; Connell, 2002; Gonzales, 2009). The difference in F_0 as a function of vowel height is considered to be universal according to the data of various languages from different language families as stated in the past studies (Whalen & Levitt, 1995; Connell, 2002). Whalen and Levitt (1995) declared that significant iF0 differences between high and low vowels were shown up in the high frequency range of an individual, but not in the low frequency range. IF0 differences were found to diminish in the lower part of the frequency range. The results of this study exhibit no evidence of the disappearance of iF0 differences in low tone. On one hand, Yin Ping [33] in Meixian Hakka is a mid tone in the higher frequency range while Yang Ping [11] is a low tone in the low frequency range. If the observation of Whalen and Levitt (1995) is true, iF0 differences would disappear in Yang Ping. Nonetheless, the results of measurements and statistical results display that iF0 differences are found even in low tone and the same degree of iF0 differences occur throughout the whole tone for Yin Ping [33] and Yang Ping [11]. On the other hand, the falling tones in this study which are Shang Sheng and Qu Sheng show that iF0 differences between high and low vowels decrease over time when the F_0 contour approaches the end of the tone. Significant high iF0 differences occur over the first half of the tones, that is, from 0% to 70% for Shang Sheng [41] and from 0% to 60% for Qu Sheng [51]. There are no significant iF0 differences over the final half of the tones. It should be noted that though iF0 differences decrease and finally diminish when the F_0 contour moves towards the lower frequencies, it is not determined by the frequency range. As can be seen in Figure 2.16, the curves on vowels /i, u/ merge with the curve on vowel /a/ at around 160 Hz and the iF0 differences disappear over the final part of the tone. However, significant iF0 differences between high and low vowels are found in Shang Sheng at the same frequency range. In this study, it is suggested that the variation of iF0 difference is conditioned by the contour shape and direction of the tone. For the tone that has a falling contour, iF0 differences between high and low vowels are found in the first half of the F_0 contour. The iF0 differences decrease and may finally diminish over the final portion of the tone which is in a relatively low frequency range. In contrast, for the tone that has a level contour, iF0 differences exist throughout the whole tone. Similar findings can also be observed in the acoustic data of Shanghai tones which were investigated the study of Zhu (1999).

2.1.3.3 Comparison of F_0 contours of the citation tones in V(C) syllables and CV(C) syllables

With respect to syllable typology, Blevins (1995) discussed cross-linguistic variation in syllables types with the data of eleven languages such as English, Finnish and Totonac. All of the languages have CV syllables and it was suggested that CV is universal in the languages of the world. Besides, the syllable type of CVC was found in most of the languages as compared to other syllable types such as V and VC. Hamdi and his colleagues conducted a comparative investigation of syllable types based on the data of Arabic dialects (Hamdi et al., 2005). The frequency of occurrence of different syllable types was computed and results showed that CV and CVC syllables were the most frequently used syllable types. Consonant clusters are not permitted in syllable initial and final positions in Meixian Hakka. Thus, Meixian Hakka has four syllable types which are V, CV, CVC and VC. It can be observed in the past studies that CV and CVC are predominant syllable types in Meixian Hakka and they are more frequently used than the syllable types of V and VC (Huang, 1995; Xie, 1994; Yuan et al., 2001). The previous sections of this chapter mainly concerned the patterns of F_0 variation of the citation tones in CV(C) syllables. In this section, the patterns of F_0 variation in different syllable types are examined and F_0 patterns of the citation tones in CV(C) syllables are compared with those in V(C) syllables despite influence of vowels.

In Meixian Hakka, lexical gaps are found in the combination of tone and vowel for V(C) syllables. All six citation tones can be associated with vowel [i] in V(C) syllables, but not vowels [a, u]. At least one meaningful monosyllabic word is found for each combination of the citation tone and vowel [i]. In contrast, no word of V(C) structure consists of any of the six citation tones and vowel [u]. The case of vowel [a] is in between these two extremes of [i] and [u]. It was observed in the past studies (Huang, 1995; Xie, 1994) that Yin Ping [33] and Shang Sheng [41] can be associated with [a] in V syllables, whereas Yin Ru [41] can be associated with vowel [a] in VC syllables. Due to the existence of lexical gaps, that is, not every citation tone on vowels /i, a, u/ in CV(C) syllables has a direct corresponding tone in V(C) syllables, it is not possible to have a direct comparison between mean F_0 contours of the citation tones in CV(C) syllables that computed based on the tones on vowels /i, a, u/ and mean F_0 contours of the citation tones in V(C) syllables. Alternatively, mean F_0 contours of the six citation tones that associated with vowel [i] in V(C) syllables are compared with the corresponding F_0 contours that computed from vowels [i, a, u] in CV(C) syllables. Table 2.5 exhibits mean duration of the F_0 contour of each of the six citation tones in Meixian Hakka produced by five male and five female speakers. Results display that the duration of the citation tones in CV(C) syllables tend to be shorter than those in V(C) syllables. The difference in duration between tones in CV(C) syllables and those in V(C) syllables

		Duration of F_0 contour (ms) – Tones in CV(C) syllable / Tones in V(C) syllable					
Tonal Category		Yin Ping [33]	Yang Ping [11]	Shang Sheng [41]	Qu Sheng [51]	Yang Ru [55]	Yin Ru [41]
Female Speakers	F1	386 / 499	353 / 514	352 / 513	325 / 513	258 / 345	216 / 427
	F2	449 / 521	459 / 570	438 / 512	416 / 503	244 / 312	327 / 496
	F3	392 / 425	355 / 487	341 / 376	334 / 422	212 / 277	305 / 402
	F4	353 / 459	287 / 466	286 / 445	257 / 431	146 / 201	150 / 234
	F5	402 / 500	353 / 400	309 / 420	340 / 475	221 / 361	246 / 399
Male Speakers	M1	319 / 418	293 / 427	307 / 439	263 / 408	145 / 230	202 / 261
	M2	383 / 448	394 / 549	316 / 447	292 / 412	194 / 272	212 / 346
	M3	430 / 491	404 / 556	334 / 455	309 / 378	161 / 237	190 / 302
	M4	348 / 450	360 / 504	304 / 493	309 / 479	143 / 223	190 / 324
	M5	329 / 461	288 / 418	263 / 370	283 / 436	159 / 307	177 / 310

Table 2.5 Mean duration of F_0 contour of the six citation tones in Meixian Hakka that superimposed on V(C) syllables and CV(C) syllables for five female and five male speakers

ranges from 30ms – 190ms. A t-test shows that the difference in duration between tones in the two syllable types is statistically significant [t (59) = -21.66, p < .000]. This indicates that the mean duration of the citation tones in V(C) syllables (m = 413.17) is markedly greater than the mean duration in CV(C) syllables (m = 297.45). As shown in Table 2.5, this finding holds true for all tones and for all males and females. The results of t-test show that there is no significant difference in duration between male and female speakers [p > .05, df = 58].

Figures 2.18a – 2.18j display F_0 contours of the six citation tones in V(C) syllables and CV(C) syllables for each of the ten speakers. The x-axis of the diagram shows F_0 values in Hertz while the y-axis shows the relative duration in percentage. For the majority of the speakers, Yin Ping [33] in CV(C) syllables has similar tone height and contour shape as the tone in V(C) syllables. Despite the data of M1, M2 and M4, the statistical results of t-test show no significant difference in F_0 contour between tone [33] in CV(C) syllables and that in V(C) syllables [p > .05, df = 6] throughout the 10% to

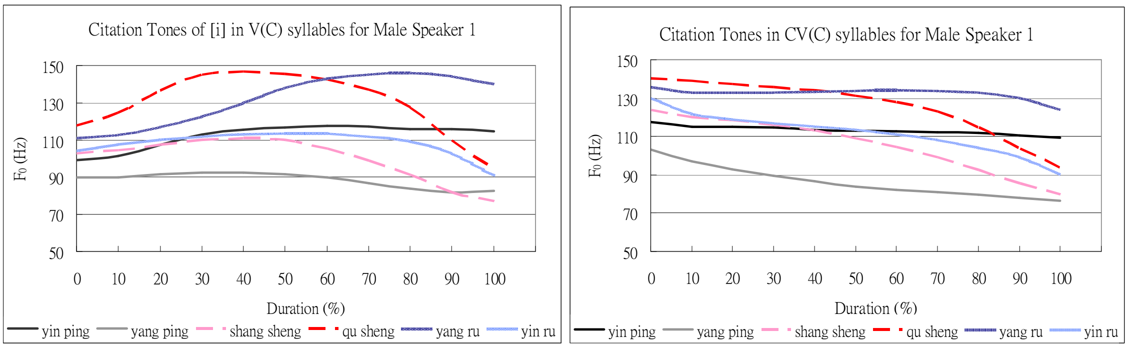


Figure 2.18a Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by M1

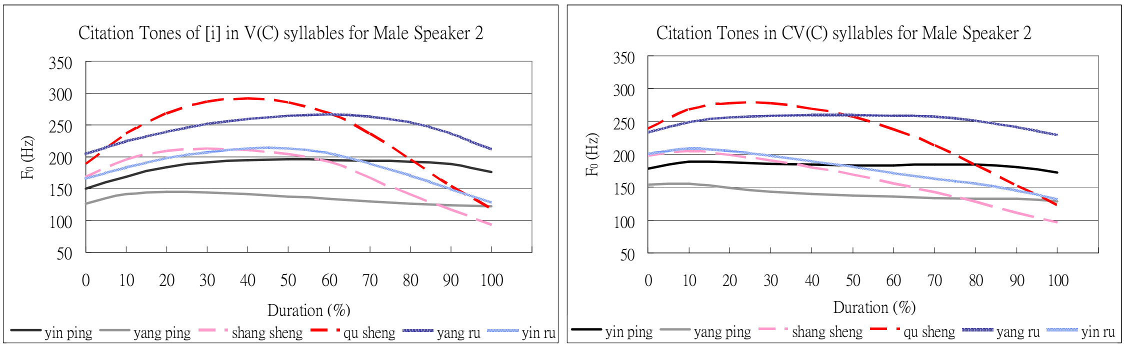


Figure 2.18b Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by M2

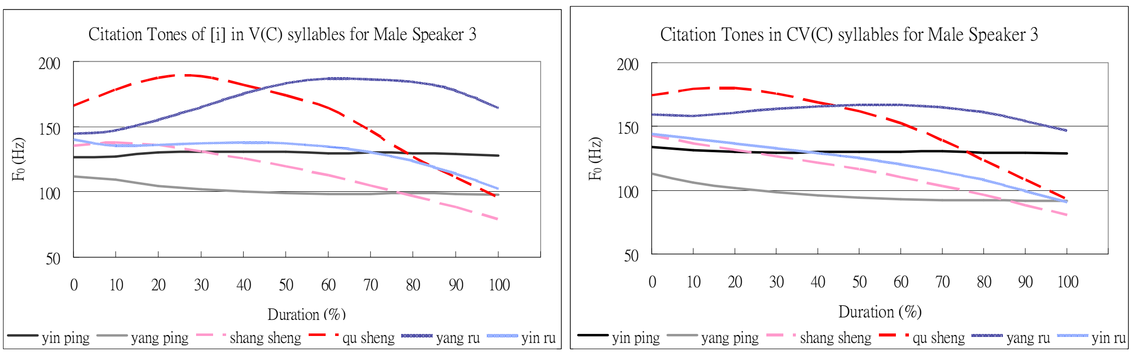


Figure 2.18c Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by M3

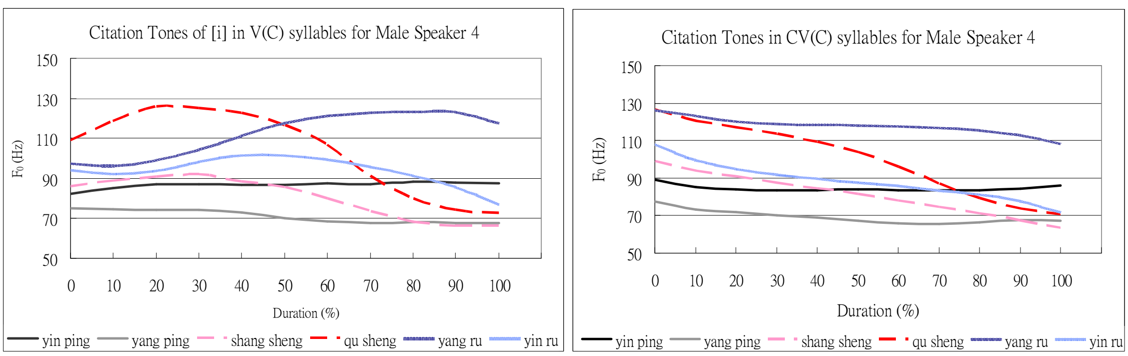


Figure 2.18d Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by M4

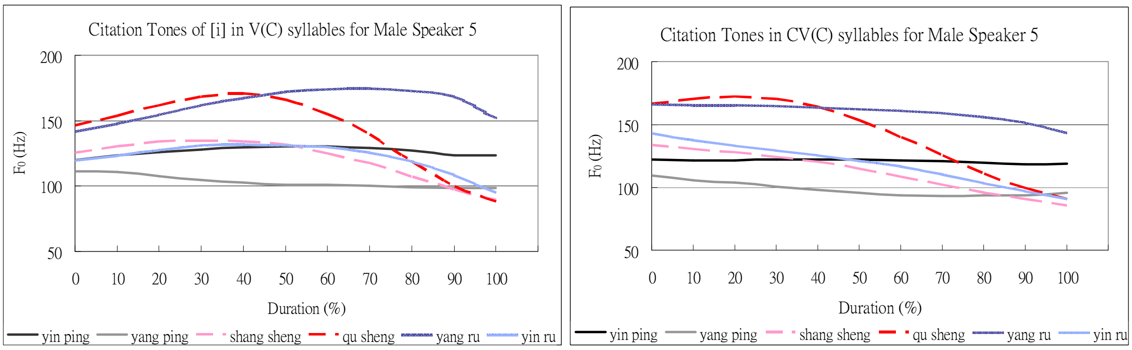


Figure 2.18e Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by M5

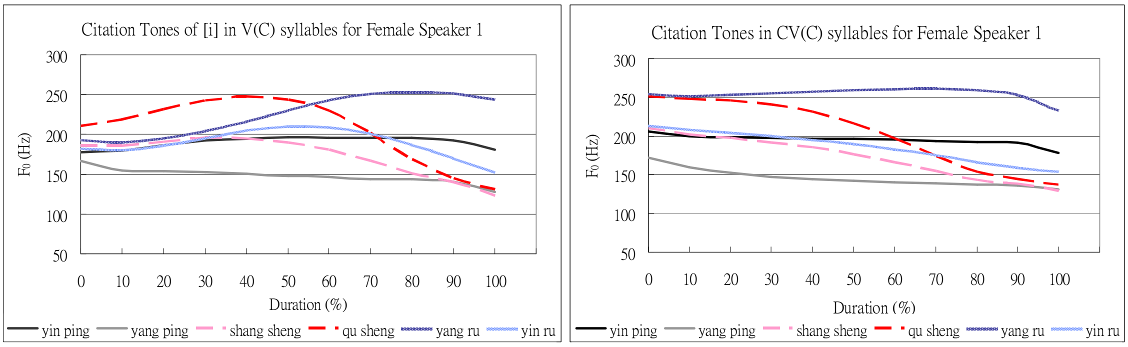


Figure 2.18f Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by F1

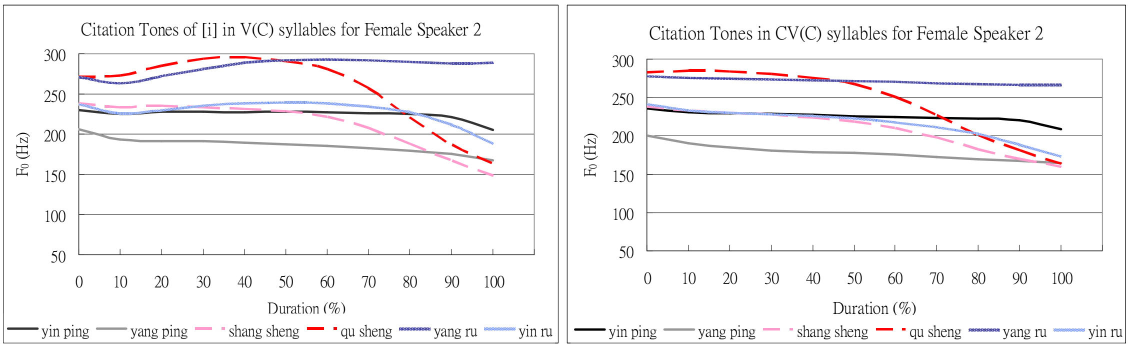


Figure 2.18g Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by F2

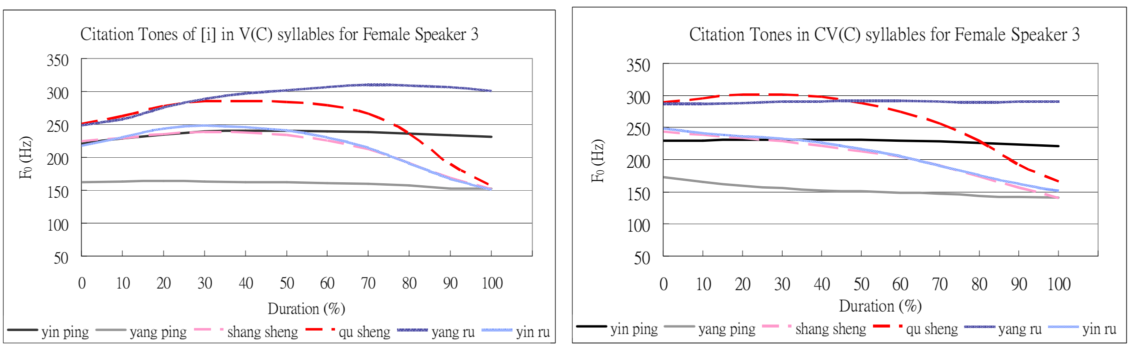


Figure 2.18h Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by F3

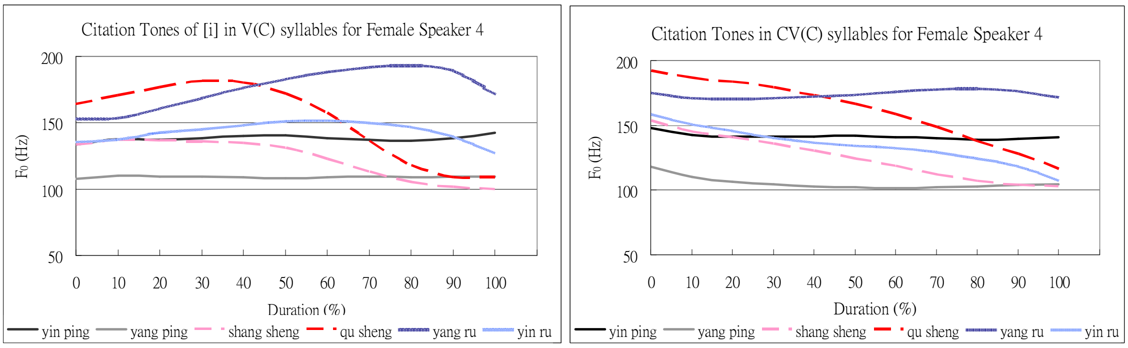


Figure 2.18i Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by F4

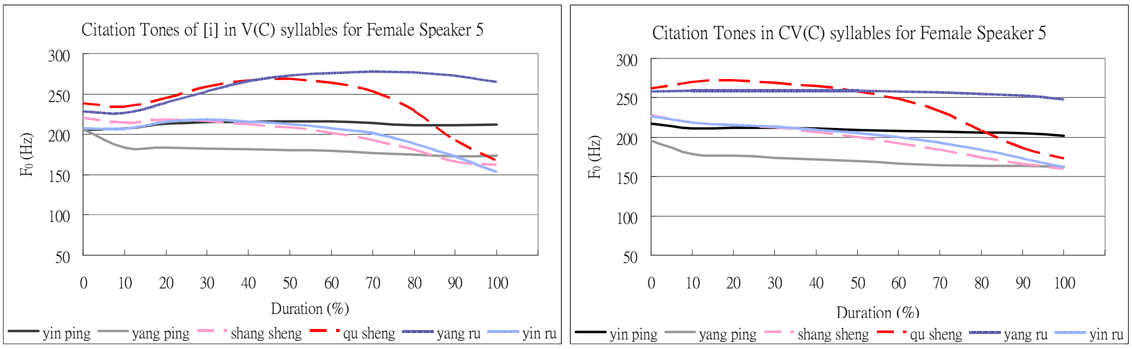


Figure 2.18j Mean F_0 contours of the six citation tones in V(C) syllables (left) and CV(C) syllables (right), respectively, uttered by F5

100% time points. As can be seen, the F_0 contours of Yin Ping in CV(C) and V(C) syllables for most speakers are basically level throughout the overall duration. In contrast, the F_0 patterns of Yin Ping for M1, M2 and M4 are different from the F_0 patterns for other speakers. Figures 2.18a and 2.18b show that the F_0 contour of Yin Ping in V(C) syllables starts at the mid-low region of the frequency range and there is a rising transition from the onset towards the tonal target of Yin Ping at the 30% time point. The remaining part of the tone after the 30% time point has a level contour and the tone height of Yin Ping in V(C) syllables is similar to that in CV(C) syllables. The F_0 contours of Yin Ping in CV(C) syllables and that in V(C) syllables are generally level for M4. Results show that the F_0 contours of Yin Ping in the two syllable types remain steady from the 20% to 100% time point. The tone in CV(C) syllables are different from that in V(C) syllables in that the latter has a slight rise starting from the onset of F_0 contour before the 20% time point, whereas the former may have a slight fall at the beginning of the F_0 contour.

As similar to the case of Yin Ping [33] in CV(C) syllables, Yang Ping [11] in V(C) syllables for most of the speakers has a level contour and there may be a slight fall at the onset. Despite the onset value, the F_0 range of tone [11] in V(C) syllables is relatively small which is less than 12 Hz for all speakers except M2, F1 and F2. In comparison, the difference in F_0 between the highest and lowest point of F_0 contour in CV(C) syllables ranges between 12 Hz and 41 Hz. The F_0 contours of tone [11] in CV(C) syllables fall gradually from the onset to at least the 30% time point for all speakers and the F_0 contours remain steady in the following portion of duration. It can be observed that the distance of initial fall in pitch in CV(C) syllables is longer than the possible initial fall in V(C) syllables. This may be explained by the existence of syllable initial consonant in CV(C) syllables. It is well-known that the F_0 contour of a following vowel is affected by the neighboring consonant. The F_0 onset immediately after voiceless consonants is high in pitch and the F_0 contour tends to be falling (Gandour, 1974; Xu & Xu, 2003). In this study, the CV(C) syllables start with a voiceless consonant. It is assumed that the F_0 onset of tone [11] in CV(C) syllables is higher than that in V(C) syllables, and a transition to the low tonal target of tone [11] tends to be longer in CV(C) syllables. The statistical results of t-test display no significant difference in F_0 contour of tone [11] between the two syllable types from the onset of the F_0 contour to the 20% time point, as well as at the offset. However, the results reveal that the F_0 contour of tone [11] in CV(C) syllables between the 30% time point and the 90% time point is significantly lower than that in V(C) syllables [$p < .05$, df = 9].

The F_0 contour of Shang Sheng [41] in CV(C) syllables is falling for all speakers as shown in Figures 2.18a – 2.18j. The fall of F_0 contour is relatively gentle and it starts from the onset of duration. The F_0 contours of tone [41] in CV(C) syllables and those in

V(C) syllables are different in the first half of duration of the tone. As exhibited in the figures, the portion of F_0 contour in V(C) syllables from the onset to around the 40% time point is basically level with a very gentle rise for all speakers except M3. Figure 2.18c shows that the steady state of F_0 contour for M3 ends at the 20% time point. After the steady state of F_0 contour, the F_0 contour of tone [41] falls from the highest point of F_0 contour towards the end point. Despite the onset value of F_0 contour in CV(C) syllables which tends to be influenced by the preceding consonant, no significant difference in F_0 is found between the 10% time point of F_0 contour in CV(C) syllables and the 30% time point, which would be the highest point, of F_0 contour in V(C) syllables [$t(9) = 1.791$, $p < .107$]. In addition, results of t-test also exhibit that there is no significant difference between the offset of F_0 contour in CV(C) syllables and that in V(C) syllables [$t(9) = -2.124$, $p < .063$]. In summary, results show that a steady state tends to occur in the first half of F_0 contour of tone [41] in V(C) syllables, but not in CV(C) syllables. It can be observed that the shape of the remaining half of F_0 contour of tone [41] in V(C) syllables after the 30% time point resembles the overall contour shape of tone [41] in CV(C) syllables as the highest point and end point of the F_0 contour of tone [41] between the two syllable types are similar.

The difference in F_0 contour of Qu Sheng [51] between CV(C) syllables and V(C) syllables is markedly the contour shape. As shown in the figures, the F_0 contour of tone [51] in V(C) syllables is convex in shape. The first half of F_0 contour from the onset to around the 40% time point exhibits a rising contour which begins at a point in the mid-high region of an individual's F_0 range and ends at the highest point of the F_0 contour. There is a fall of F_0 contour in the second half of duration and the F_0 values decreases from the highest point towards the end point of F_0 contour. In contrast, tone [51] in CV(C) syllables has either a simple falling contour (as shown in Figures 2.18a, 2.18d and 2.18i) of which the onset is the highest point of F_0 contour, or a falling contour with an initial plateau (as shown in other figures). There is a plateau at the beginning portion of F_0 contour of tone [51] for most of the speakers and the plateau generally finishes at around the 30% time point of duration. The F_0 contour then falls towards the end point which is the lowest point of the F_0 contour. Results indicate that a longer rise in pitch tend to occur at the beginning of F_0 contour of tone [51] in V(C) syllables, whereas there is no or little transition from the onset of F_0 contour in CV(C) syllables when tone [51] occurs after a voiceless obstruent which may result in a relatively high F_0 onset of the following vowel.

Yin Ru [41] in CV(C) syllables has a simple falling contour for all speakers and the F_0 contour of tone [41] falls from the mid-high region of an individual's F_0 range to the low region. The F_0 contour of tone [41] in V(C) syllables is different from that in CV(C) syllables and two F_0 patterns are found in the former. Yin Ru [41] in V(C) syllables has either a basically level F_0 contour for over 50% of the duration with a fall at the final part of F_0 contour (as shown in Figures 2.18c, 2.18g, 2.18i, and 2.18j) or a rising-falling contour that starts with a slight rise from the onset to around the 50% time point and falls towards the end point throughout the rest of F_0 contour (as shown in the data of other speakers). The F_0 contour of Yang Ru [55] in CV(C) syllables is basically level throughout the overall duration and it remains steady at the high frequency range. In contrast, Yang Ru in V(C) syllables has a rising contour. The F_0 contour starts at the mid-high region of the frequency range and rises to the high frequency range, which is the tonal target of a high level tone, at around the 60% time point. There is a plateau in the final portion of the F_0 contour. The statistical results of t-test display that the F_0 contour of tone [51] after the 60% time point in V(C) syllables is significantly higher than that in CV(C) syllables [$p < .05$, $df = 9$].

In summary, results of this study demonstrate that the citation tones in CV(C) syllables and those in V(C) syllables differ markedly in duration as well as F_0 contour. The mean durations of the six citation tones in V(C) syllables are significantly greater than those in CV(C) syllables for both male and female speakers. In addition, the citation tones in V(C) syllables generally start at a frequency lower than the citation tones in CV(C) syllables and it may be explained by the raising effect of voiceless initial obstruent on F_0 onset of the following vowel. There is a rise in pitch after the onset for high tones, namely, Qu Sheng [51] and Yang Ru [55] in V(C) syllables, whereas there is a plateau or gentle fall in pitch after the onset for the corresponding tones in CV(C) syllables. Besides, a plateau or gentle rise in pitch tends to occur after the onset for mid tones, namely, Yin Ping [33], Shang Sheng [41] and Yin Ru [41] in V(C) syllables, whereas a fall in pitch tend to occur for the corresponding tones in CV(C) syllables. The F_0 contour in V(C) syllables immediately after the onset for the low tone, namely, Yang Ping [11], may be level, slightly rising or slightly falling. In contrast, a fall in pitch tends to occur at the beginning of F_0 contour of the low tone in CV(C) syllables. Acoustic measurements and statistical analysis display that, except for Yin Ping [33] which has no significant difference between the two syllable types throughout most of the duration, Yang Ru [55] and other citation tones in V(C) syllables have a lower contour than the corresponding tones in CV(C) syllables during about the first half and about the first one-third of the duration of the tones, respectively. The F_0 difference between the two syllables becomes smaller over time. For the remaining half of the duration of Yang Ru [55] and the remaining two-third of the duration of other citation tones, the F_0 contours of the citation tones in V(C) syllables are higher than those in CV(C) syllables.

2.2 Citation Tones in Tone Sandhi

2.2.1 Introduction

It is well-known for tone languages that tones interact with each other in combination and such process is called tone sandhi. Tonal alternations occur depending on the nature of surrounding tones. Tone sandhi rules in Meixian Hakka have been described in the past studies based on mainly impressionistic descriptions (Hashimoto, 1973; Huang, 1992; 1995; Yuan et al., 2001). It was generally agreed that in Meixian Hakka, tone sandhi occurs in two-syllable structure. The tone pattern of the citation tone on the first syllable of a bisyllabic word deviates from its form in isolation when it occurs in the sandhi environment. Huang (1992; 1995) and Yuan et al. (2001) declared that tone sandhi is found only in Yin Ping and Qu Sheng, which corresponds to the mid level tone and the high falling tone, respectively, but not in other citation tones. Two sandhi forms, which are the high rising tone, and the high level tone or the high to mid-high falling tone, are derived from the citation forms of Yin Ping and Qu Sheng, respectively. Nonetheless, discrepancy in sandhi environment is found between the descriptions of the past studies. In the study of Huang (1995), the citation form of Yin Ping is the mid level tone [44]. It changes to the high rising tone [35] when the sandhi environment is fulfilled. The citation form of Qu Sheng is the high falling tone [51] and its sandhi form is the high level tone [55]. Both Yin Ping and Qu Sheng change to their sandhi forms when they are followed by Yang Ping [11], Shang Sheng [31], Qu Sheng [53] and Yin Ru [1]. In other words, Yin Ping and Qu Sheng have different sandhi forms but they share the same sandhi environment. In the study of Yuan et al. (2001), Yin Ping [44] changes from its citation form to the mid-high to high rising tone [45] when it precedes Yang Ping [11], Shang Sheng [31] and Yin Ru [21], whereas Qu Sheng changes from the high falling tone [52] to the slight falling tone [54] or [53] when it precedes Yin Ping

[44], Yang Ping [11], Shang Sheng [31] and Qu Sheng [52]. The descriptions of tone sandhi in the study of Yuan et al. (2001) showed that the sandhi forms of Yin Ping and Qu Sheng are triggered by different sandhi environments. It can be seen that, as the past studies were merely based on the impression of the linguist, the descriptions of tone sandhi rules between the past studies are inevitably different in terms of the pitch value of the citation form of the tone and its corresponding sandhi form, as well as the sandhi environment. In contrast, the description of the tone sandhi rules in Hashimoto's study (1973) was based on the acoustic data from real speech and it was supposed to be more scientific. Nonetheless, the tone sandhi rules that formulated in the study of Hahsimoto are shown to be different from the descriptions in the studies of Huang (1995) and Yuan et al. (2001). Hahsimoto claimed that tone sandhi occurs in Yin Ping, Yang Ping and Shang Sheng. The citation form of Yin Ping, the mid level legato, changes to a rising legato when it precedes Shang Sheng, the mid falling legato, and Yin Ru, the mid falling staccato, whereas Yang Ping changes from the low level legato into a slight rising legato when it is followed by Shang Sheng, the mid falling legato, Qu Sheng, the high level legato, and Yin Ru, the mid falling staccato. Besides, the citation form of Shang Sheng, the mid falling legato, is flatten when it is followed by any of the citation tones but Yang Ping, the low level legato. Lü (2006) formulated the tone sandhi rules in Meixian Hakka based on the acoustic data. He concluded that tone sandhi occurs when Yin Ping, Qu Sheng and Shang Sheng appear as the first tone in a bisyllabic tonal combination, but not the case of Yang Ping, Yin Ru and Yang Ru. Three tone sandhi rules were formulated in the study of Lü (2006): Yin Ping [33] is realized as the rising tone [35] when it is followed by Yang Ping [21], Shang Sheng [31] and Yin Ru [<u>31</u>]; Qu Sheng [52] is realized as the high level tone [55] when it precedes Yang Ping, Shang Sheng, Qu Sheng and Yin Ru; and Shang Sheng is realized as the mid-high level tone [44]

when it is followed by Yang Ping, Shang Sheng and Yin Ru. In order to verify the descriptions of tone sandhi in the past studies and resolve the discrepancies between the past studies, it is worthwhile to investigate bisyllabic tone patterns in various tonal combinations in this study.

This section concerns tone patterns in bisyllabic lexical tone combinations. The F_0 contours of the Meixian Hakka tones in different tonal combinations are analyzed and discussed. As the purposes of this study are (1) to examine the citation forms and sandhi forms of the tones in Meixian Hakka, (2) to investigate the sandhi environment for each of the citation tones in which tone sandhi occurs, (3) to formulate tone sandhi rules in Meixian Hakka based on the F_0 contours of the citation tones in various tonal combinations, and (4) to verify the descriptions of Meixian Hakka tone sandhi rules in the past studies based on the acoustic results of this study. It is believed that through the manipulation of speech analysis software, objective and scientific results can be obtained for the formulation of tone sandhi rules in Meixian Hakka. The results of this study may provide an objective basis for comparing the tone sandhi rules in Meixian Hakka with the rules in other Chinese dialects.

2.2.2 Methodology

The acoustic data of tones in various bisyllabic lexical tone combinations were provided by a male speaker. He is a native speaker of Meixian Hakka who was born and grew up in the district of Meijiang. The speaker was 69 years old at the time of recording with no history of speech and hearing disorders. The word list consists of Chinese bisyllabic words that exemplify all possible combinations of the six citation tones in Meixian Hakka. Table 2.6 exhibits various tonal combinations of the six citation tones in Meixian Hakka. The columns of the table indicate the citation tones of the first syllable of the bisyllabic word, whereas the rows display the citation tones of the second syllable.

1st syllable / 2nd syllable	Tone 1 [33]	Tone 2 [11]	Tone 3 [41]	Tone 4 [51]	Tone 5 [5]	Tone 6 [<u>41</u>]
Tone 1 [33]	si33 kua33 西瓜	ts^{h}a11 pi33 茶杯	tu41 ki33 肚饥	kui51 fa33 桂花	lat5 tsiau33 辣椒	fat<u>41</u> sau33 发烧
Tone 2 [11]	tsu33 p^{h}i11 猪皮	p^{h}i11 k^{h}iu11 皮球	pu41 t^{h}eu11 斧头	tsa51 t^{h}eu11 蔗头	sak5 t^{h}eu11 石头	fat<u>41</u> ts^{h}oi11 发财
Tone 3 [41]	tsu33 tu41 猪肚	k^{h}i11 tsɿ41 棋子	pu41 k^{h}au41 补考	ki51 tsa41 记者	sət5 tsiu41 食酒	t^{h}iet<u>41</u> kui41 铁轨
Tone 4 [51]	ti33 poi51 里背	p^{h}i11 tai51 皮带	pu41 k^{h}o51 补课	fu51 kui51 富贵	sət5 tsu51 食昼	pit<u>41</u> ki51 笔记
Tone 5 [<u>55</u>]	pi33 sak5 碑石	p^{h}o11 p^{h}ak5 婆白	pu41 sip5 补习	su51 sok5 树勺	p^{h}ak5 sət5 白食	pit<u>41</u> ts^{h}ət5 笔直
Tone 6 [<u>41</u>]	su33 tsok<u>41</u> 书桌	ts^{h}a11 set<u>41</u> 茶色	tu41 pok<u>41</u> 赌博	fu51 tsuk<u>41</u> 腐竹	fat5 p^{h}at<u>41</u> 活泼	pit<u>41</u> fap<u>41</u> 笔法

Table 2.6 The possible combinations of the so-called six Meixian Hakka citation tones in bisyllabic words

The IPA transcription of each bisyllabic word is listed above the Chinese characters. There are 36 test bisyllabic words (6 citation tones x 6 citation tones) in the list. For the ease of data analysis, every syllable in the word list starts with a syllable-initial consonant. In this case, the F_0 contours of the two syllables of a bisyllabic word can be separated by the appearance of consonant. The syllable-initial consonants included in this study are obstruents, namely stops, fricatives and affricates. Each of the test words was pronounced in isolation. The order of the test bisyllabic words was randomized and four readings of the word list were recorded. The speakers were instructed to read the word list in a normal manner and the recordings were carried out in a quiet room. The experimental procedures of F_0 extraction for tones in bisyllabic words are similar to the procedures for tones in monosyllabic words as described in section 2.1.2 regarding the methodology. F_0 trajectories of each syllable of the bisyllabic words were analyzed separately. Wide band spectrograms were manipulated to determine the onset and offset of the F_0 contour. The citation form and sandhi form of each citation tone are compared and discussed in the following section.

2.2.3 Results

F_0 trajectories of the citation tones in various bisyllabic tonal combinations are shown in Figures 2.19-2.24. The x-axis of the figure displays the fundamental frequency values in Hertz while the y-axis exhibits the duration in millisecond. For the ease of comparison, F_0 contour of the tone on the first syllable of a bisyllabic word starts at 0 ms and that on the second syllable begins at 500 ms. Results show that tone sandhi in Meixian Hakka occurs in Yin Ping [33], Shang Sheng [41] and Qu Sheng [51], and two sandhi forms that deviate from the citation forms are generated. The tone patterns of Yang Ping [11], Yin Ru [$\underline{41}$] and Yang Ru [$\underline{55}$ or 5] in monosyllabic words remain unchanged in bisyllabic words. The results of tones that undergo tone sandhi, which are Yin Ping, Shang Sheng and Qu Sheng, are presented in sections 2.2.3.1-2.2.3.3. The sandhi forms and sandhi environments of the three tones are described and examined. In section 2.2.3.4, the results of other three tones, which are Yang Ping, Yin Ru and Yang Ru, are exhibited and the unchanged tone patterns of the citation tones are displayed.

2.2.3.1 Yin Ping [33]

As discussed in section 2.1.3, acoustic data indicated that Yin Ping [33] is the mid level tone in Meixian Hakka. Figure 2.19 displays the mean F_0 trajectories of Yin Ping in combination with other citation tones. The upper diagram shows bisyllabic tonal

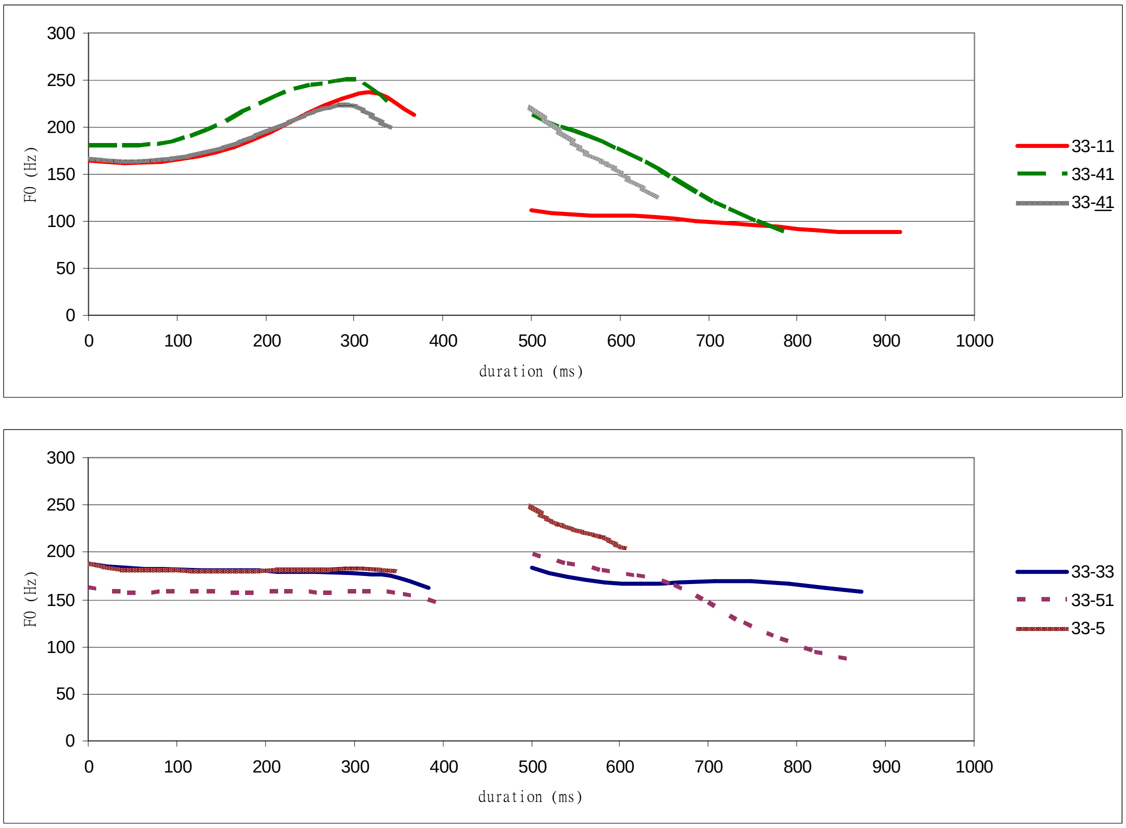
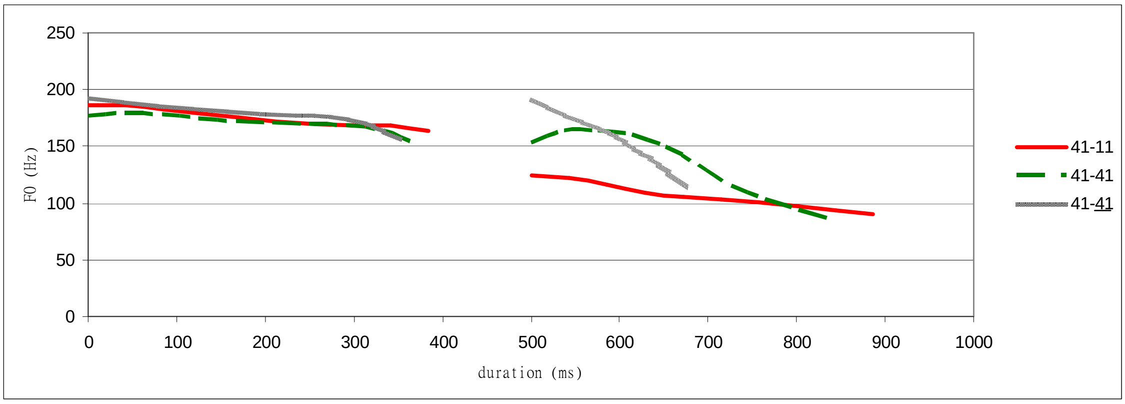


Figure 2.19 Mean F_0 trajectories of the Meixian Hakka tones in bisyllabic words in which Yin Ping [33] occurs as the first syllable

combinations in which tone sandhi occurs, whereas the lower diagram shows the tonal combinations in which the F_0 pattern of Yin Ping remains unchanged. Results indicate that the F_0 contour of Yin Ping deviates from its citation form whenever Yin Ping appears as the first syllable of a bisyllabic word and is followed by Yang Ping [11], Shang Sheng [41] and Yin Ru [41]. It can be noticed that the citation form of Yin Ping, the mid level tone, in two-syllable words becomes its sandhi form, the mid to high rising tone, when it is followed by the low level tone, the mid-high to low falling tone and the short mid-high to low falling tone. As shown in the upper diagram of Figure 2.19, the sandhi form of Yin Ping occurs at similar frequency range where the citation form of Yin Ping begins. A t-test shows that the difference in F_0 between the onset of Yin Ping which undergoes tone sandhi and the onset of Yin Ping which remains unchanged in bisyllabic tonal combinations is not significant [paired-samples t-test, p < .169]. The F_0 contour of the sandhi form rises rapidly to the highest point which is approximate to the starting point of the high to low falling tone [51]. Based on the acoustic data, the sandhi form of Yin Ping is transcribed as the mid to high rising tone and the pitch value of [35] is given. Meixian Hakka contrasts six tones in monosyllabic words, three are level and three are falling, as described in the previous section. It can be seen that the sandhi form of Yin Ping, the mid to high rising tone [35], does not resemble the F_0 contour of any of the citation tones in terms of direction of F_0 movement. In contrast, the F_0 contour of Yin Ping, the mid level tone, maintains its citation form in bisyllabic words, when Yin Ping is followed by Yin Ping, the mid level tone [33], Qu Sheng, the high to low falling tone [51] and Yang Ru, the high level short tone [<u>55</u> or 5]. The results of this study support the descriptions of tone sandhi in Yuan et al.'s study (2001) and Lü's study (2006), but are not consistent with the descriptions in Hashimoto's study (1973) and Huang's study (1995). Hashimoto claimed that Yin Ping undergoes tone sandhi when it is followed by Shang Sheng and Yin Ru, whereas Huang declared that Yin Ping becomes the mid to high rising tone [35] when it is followed by Yang Ping, Shang Sheng, Qu Sheng and Yin Ru.

2.2.3.2 Shang Sheng [41]

The citation form of Shang Sheng in monosyllabic words is the mid-high to low falling tone [41]. Figure 2.20 shows the mean F_0 trajectories of tonal combinations in bisyllabic words having Shang Sheng on the first syllable. The upper diagram displays the tonal combinations in which tone sandhi occurs, whereas the lower diagram displays the bisyllabic tonal combinations in which Shang Sheng maintains its citation form. Results exhibit that Shang Sheng undergoes tone sandhi when it is followed by Yang Ping [11], Shang Sheng [41] and Yin Ru [<u>41</u>]. It can be observed that the sandhi environment of Shang Sheng is identical to that of Yin Ping. As shown in Figure 2.20, the slope of the



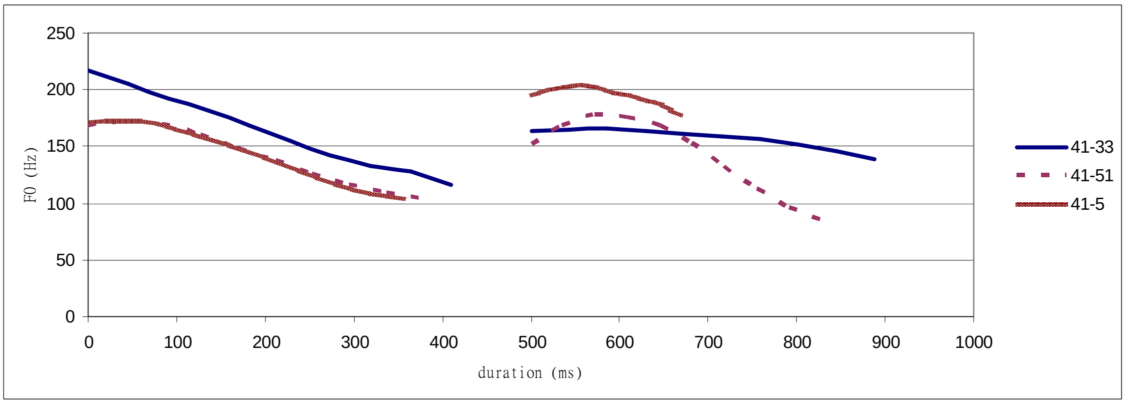


Figure 2.20 Mean F_0 trajectories of the Meixian Hakka tones in bisyllabic words in which Shang Sheng [41] occurs as the first syllable

falling tone evidently decreases as compared to the F_0 contour of Shang Sheng having no tone sandhi change, and the F_0 contour of Shang Sheng becomes quasi-level. The sandhi form of Shang Sheng resembles a level contour with a slight fall and the F_0 curve falls between 150 Hz and 200 Hz. A comparison of the F_0 contour between the sandhi form of Shang Sheng as shown in the upper diagram of Figure 2.20 and the citation form of Yin Ping as shown in the lower diagram of Figure 2.19 exhibits no statistical significance at most of the time points [paired-samples t-test, $p > .05$]. It is suggested in this study that the pitch value of the sandhi form of Shang Sheng is assigned as [33], and it is the same as the pitch value of the citation form of Yin Ping. The tone sandhi rule of Shang Sheng in Meixian Hakka is formulated according to the results. Shang Sheng, the mid-high to low falling tone [41], changes to the mid level tone [33] when it is followed by Yang Ping, the low level tone [11], Shang Sheng, the mid-high to low-falling tone [41] and Yin Ru, the mid-high to low falling short tone [41]. In contrast, the F_0 contour of Shang Sheng remains unchanged in bisyllabic words when the tone on second syllable is Yin Ping [33], Qu Sheng [51] and Yang Ru [5]. As shown in the lower diagram of Figure 2.20, Shang Sheng has a falling contour of which the onset is at around 200 Hz and the offset is at around 100 Hz. In this case, Shang Sheng as the tone on the first syllable of a bisyllabic word has a similar contour shape as the tone in a monosyllabic word. The tone sandhi of Shang Sheng has not been mentioned in the studies of Huang (1995) and Yuan et al. (2001). On the other hand, the tone sandhi of Shang Sheng has been observed in the studies of Hashimoto (1973) and Lü (2006). The sandhi form of Shang Sheng, the mid level tone, observed in this study agrees with the findings of Hashimoto (1973) and Lü (2006). Besides, the sandhi environment of Shang Sheng observed in this study is consistent with the observation of Lü (2006), but not with the observation of Hashimoto (1973). In the study of Hashimoto (1973), the falling contour of Shang Sheng is flattened when it is followed by all citation tones but Yang Ping. However, it is observed in this study that tone sandhi of Shang Sheng occurs when the second syllable is superimposed by Yang Ping [11], Shang Sheng [41] and Yin Ru [41].

2.2.3.3 Qu Sheng [51]

Qu Sheng in Meixian Hakka has a high to low falling contour [51] in monosyllabic words. Figure 2.21 shows the mean F_0 trajectories of possible bisyllabic tonal combinations in which Qu Sheng appears as the first syllable. The upper diagram exhibits the tonal combinations in which tone sandhi occurs, whereas the lower diagram exhibits the tonal combinations in which F_0 contour of the tone remains unchanged. Results show that tone sandhi occurs in Qu Sheng and the sharp falling contour of Qu Sheng is flattened in the sandhi environment. As can be seen in the upper diagram, Qu

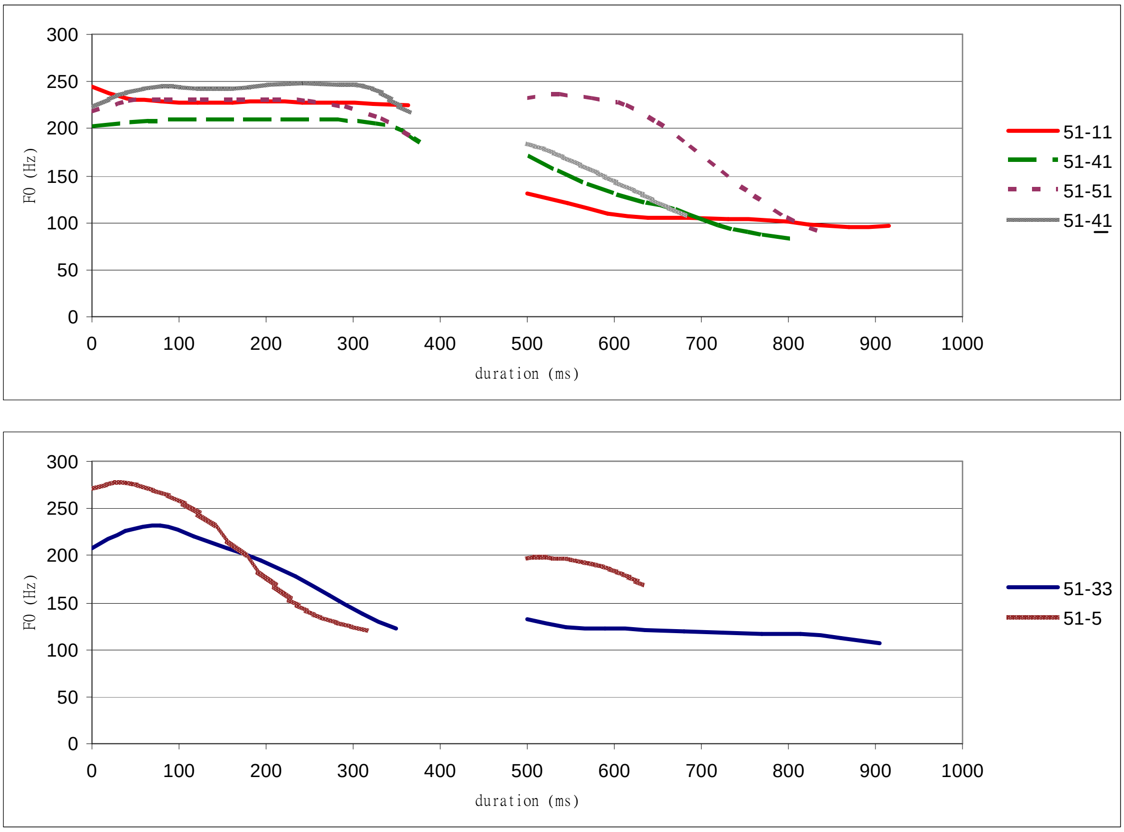


Figure 2.21 Mean F_0 trajectories of the Meixian Hakka tones in bisyllabic words in which Qu Sheng [51] occurs as the first syllable

Sheng as the tone on the first syllable of a bisyllabic word is realized as the high level tone [55] when it is followed by Yang Ping, the low level tone [11], Shang Sheng, the mid-high to low falling tone [41], Qu Sheng, the high to low falling tone [51], and Yin Ru, the mid-high to low falling short tone [41]. The level contour of Qu Sheng falls between 200 Hz and 250 Hz with a slight fall at the end of the F_0 contour. Similar to the case of Ying Ping, the sandhi form of Qu Sheng, the high level tone [55], does not resemble any of the citation forms of the Meixian Hakka tones. On the other hand, tone sandhi of Qu Sheng is blocked when the tone is followed by Yin Ping, the mid level tone [33], and Yang Ru, the high level short tone [55 or 5]. As shown in the lower diagram of Figure 2.21, the F_0 contour of Qu Sheng drops from the highest point at around 250 Hz to the lowest point at around 130 Hz. The sharp falling contour of Qu Sheng in bisyllabic tonal combinations resembles the falling contour of the tone in monosyllabic words. In comparison to the past studies, the tone sandhi rule of Qu Sheng observed in this study is consistent with the observations of Huang (1995) and Lü (2006). It is suggested that Qu Sheng changes to its sandhi form, the high level tone [55], when it is followed by tones ended in the low frequency range, namely Yang Ping [11], Shang Sheng [41], Qu Sheng [51] and Yin Ru [41]. The tone sandhi of Qu Sheng was not found in the study of Hashimoto (1973). The tone sandhi rule of Qu Sheng was also formulated in the study of Yuan et al. (2001). The sandhi form of Qu Sheng in their study, the slight falling tone [54 or 53], is similar to the sandhi form observed in this study, the high level tone [55]. However, the sandhi environment described in their study is different from that in this study. They claimed that Qu Sheng changes to its sandhi form when followed by Yang Ping, Shang Sheng, Qu Sheng and Yin Ping. As shown in the acoustic data of this study, Yin Ping as the tone on the second syllable of a bisyllabic word does not trigger the tone sandhi, but Yin Ru.

2.2.3.4 Yang Ping [11], Yin Ru [41], Yang Ru [55 or 5]

Yang Ping, Yin Ru and Yang Ru in monosyllabic words are realized as the low level tone [11], the mid-high to low falling short tone [41] and high level short tone [55 or 5], respectively. Figures 2.22-2.24 display the mean F_0 trajectories of possible bisyllabic tonal combinations in which Yang Ping, Yin Ru and Yang Ru occur as the tone on the first syllable, respectively. Figure 2.22 displays that the F_0 contours of Yang Ping in different bisyllabic tonal combinations are quasi-level and there is a slight fall in F_0 before the 20% time point. As can be seen, the F_0 contour of Yang Ping remains relatively constant throughout the overall duration and it falls between 100 Hz and 150 Hz. In comparison to Yin Ping which falls between 150 Hz and 200 Hz as shown in Figure 2.19, Yang Ping is at a relatively low frequency range in Figure 2.22. As the citation form of Yang Ping [11] in isolation, which has a low level contour with a slight

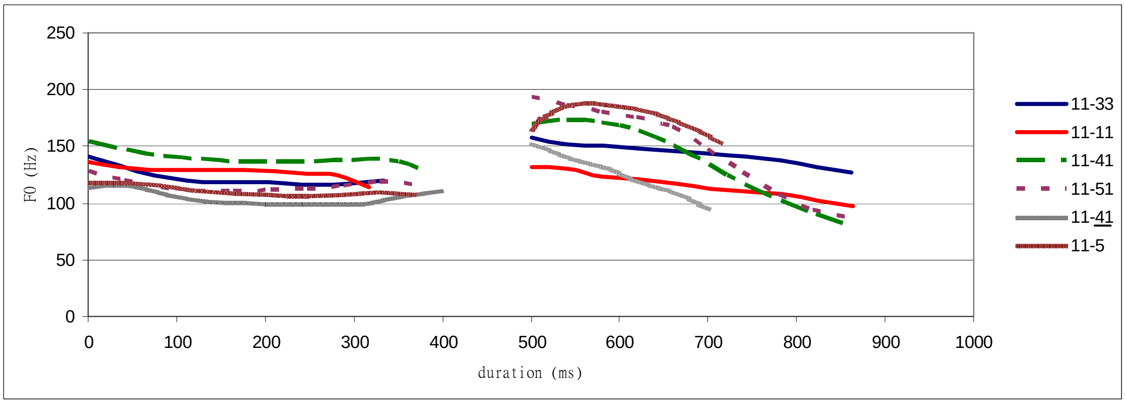


Figure 2.22 Mean F_0 trajectories of the Meixian Hakka tones in bisyllabic words in which Yang Ping [11] occurs as the first syllable

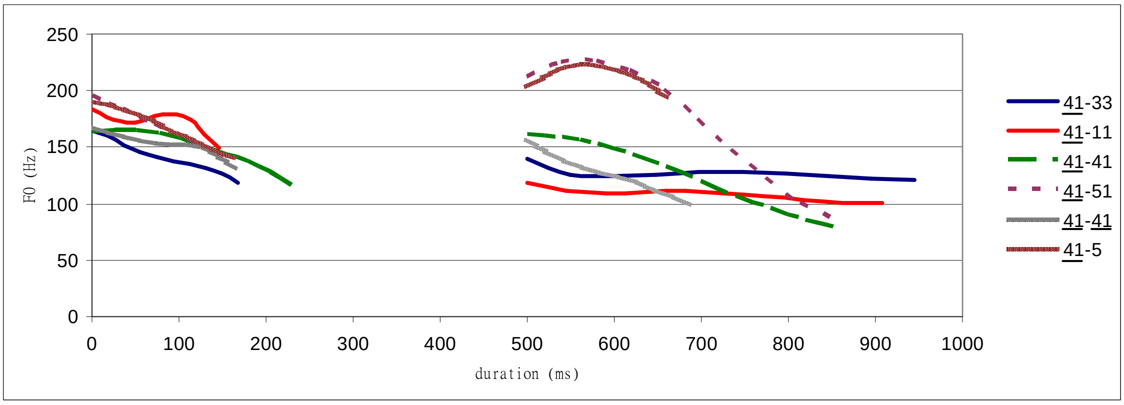


Figure 2.23 Mean F_0 trajectories of the Meixian Hakka tones in bisyllabic words in which Yin Ru [41] occurs as the first syllable

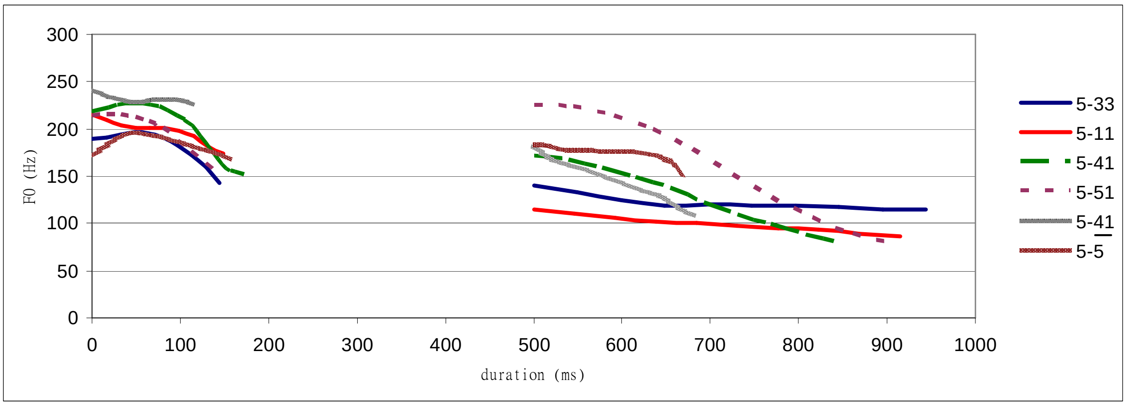


Figure 2.24 Mean F_0 trajectories of the Meixian Hakka tones in bisyllabic words in which Yang Ru [55 or 5] occurs as the first syllable

fall, is similar to the sandhi form of Yang Ping in bisyllabic tonal combinations, it seems that no tone sandhi is found in Yang Ping. Similarly, Figure 2.23 displays that the F_0 contours of Yin Ru [41] in different tonal combinations are falling in general. The F_0

contours start at around 170 Hz and finish at around 120 Hz. That is, the F_0 contour of Yin Ru falls from the mid-high frequency range to the low frequency range. This indicates that Yin Ru maintains its citation form in bisyllabic words and there is no tone sandhi observed in various tonal combinations. The F_0 contours of Yang Ru in combination with different citation tones are exhibited in Figure 2.24. It can be observed that the F_0 contour of Yang Ru becomes either a delayed falling or simple falling contour in various bisyllabic tonal combinations. The onsets of the F_0 contours of Yang Ru in bisyllabic tonal combinations are in between 170 Hz and 240 Hz, whereas the offsets are in between 135 Hz and 210 Hz. It has been described in section 2.1.3.1 that the citation form of Yang Ru in monosyllabic words is the high level short tone [5]. Nonetheless, it is observed that Yang Ru as the tone on the second syllable of bisyllabic tonal combinations in which tone sandhi does not occur has a delayed falling contour as shown in Figures 2.20, 2.22 and 2.23, a simple falling contour as shown in Figures 2.19 and 2.21, as well as a level contour with a slight fall at the end as shown in Figure 2.24. This suggests that the difference in contour shape between Yang Ru in monosyllabic word and Yang Ru as the tone on the first syllable of a bisyllabic word does not imply the occurrence of tone sandhi, but may be due to the influence of phonetic context. Therefore, it is believed that Yang Ru does not undergo tone sandhi in bisyllabic tonal combinations.

In summary, the acoustic data of Meixian Hakka show that when two syllables occurs in combination, the tone on the first syllable changes and tone sandhi occurs. Yin Ping, Shang Sheng and Qu Sheng undergo tone sandhi when they are on the first syllable of a bisyllabic word and the sandhi environment is fulfilled. On the other hand, tone sandhi is not found in bisyllabic tonal combination in which Yang Ping, Yin Ru and Yang Ru appear as the first syllable. Three tone sandhi rules in Meixian Hakka have been formulated in this study: Yin Ping, the mid level tone [33], changes to the mid to high rising tone [35] when it is followed by Yang Ping [11], Shang Sheng [41] and Yin Ru [41]; Shang Sheng, the mid-high to low falling tone [41], is realized as the mid level tone [33] when it precedes Yang Ping [11], Shang Sheng [41] and Yin Ru [41]; and Qu Sheng, the high to low falling tone [51], changes to the high level tone [55] when it is followed by Yang Ping [11], Shang Sheng [41], Qu Sheng [51] and Yin Ru [41]. It can be noticed that the sandhi environments of Yin Ping and Shang Sheng are the same and they are different from the sandhi environment of Qu Sheng. The F_0 contour of Qu Sheng undergoes tone sandhi when it is followed by the citation tones with a low F_0 offset. In addition, two sandhi forms that are different from the citation forms of tones in Meixian Hakka are generated. The sandhi form of Yin Ping is realized as the mid to high rising tone [35], whereas that of Qu Sheng is realized as the high level tone [55]. In comparison with the past studies, results of this study indicate that only the observation of Lü (2006) is considered to be an appropriate description of tone sandhi in Meixian Hakka, but not the observations of Hashimoto (1973), Huang (1995) and Yuan et al. (2001) in which either the sandhi form or the sandhi environment is inconsistent with the results of this study based on the acoustic data.

CHAPTER THREE
PERCEPTION OF LEXICAL TONES IN MEIXIAN HAKKA

3.1 Introduction

As discussed in chapter two, there are six citation tones in contemporary Meixian Hakka. The historically tonal category of Yin Ping corresponds to the mid level tone [33]; the tonal category of Yang Ping to the low level tone [11]; the tonal category of Shang Sheng to the mid-high to low falling tone [41]; the tonal category of Qu Sheng to the high to low falling tone [51]; the tonal category of Yin Ru to the mid-high to low falling short tone [41]; and the tonal category of Yang Ru to the high level short tone [55 or 5]. Though the tonal system of Meixian Hakka has been described phonologically and historically in the past studies (Hanyu Fangyan Zihui, 1989; Huang, 1992; Xie, 1994; Yuan et al., 2001), the acoustic data provided are limited and no perceptual study on Meixian Hakka tones has been done. Therefore, it would be worthwhile to examine the possible correlation between physical dimension and perceptual dimension of the tones in Meixian Hakka. The acoustic data show that three types of F_0 contours are found in the citation forms of words in Meixian Hakka, which are level, simple falling and delayed falling. The rising contour does not occur in the citation forms of the words, but in tonal combination patterns of bisyllabic words. The target tones of this perceptual study are the four lexical long tones in Meixian Hakka, which are the mid level tone [33], the low level tone [11], the high to low falling tone [51] and the mid-high to low falling tone [41]. The high level short tone [55 or 5] and the mid-high to low short tone [41] were not included in this study as they are in complementary distribution with the four long tones. As in other Chinese dialects, the two short tones in Meixian Hakka occur only in CVC syllables, whereas the long tones occur in other syllable types but not CVC. The tone values indicate that the four long tones differ in tonal height and slope of contour from the physical dimension. If the perception of tones is mainly based on the acoustic characteristics of tones, it may be intuitively assumed that (1) the difference between [33] and [11] is tonal height; (2) the difference between [41] and [11] is slope of contour; and (3) the difference between [51] and [41] is slope of contour as well. The tones [11], [41] and [51] are different in slope of contour as their offsets are at the low frequency range and the three tones differ in their F_0 onset. On the other hand, though [11] and [33] are considered to be phonologically level, it should be noted that [11] and [33] differ acoustically not only in their tonal height but also in the contour shape. As shown in the acoustic data, the mid level tone [33] is generally level, whereas the low level tone [11] in Meixian Hakka has a slight drop in F_0 throughout the overall duration. In consideration of the contour shape of the long tones, Meixian Hakka has three falling tones and one level tone. The three falling tones include the high to low falling tone [51], the mid-high to low falling tone [41] and the low level tone [11] with a slight fall; the level tone includes the mid level tone [33].

Gandour and Harshman (1978) conducted a cross-language study of tone perception on English, Yoruba and Thai. The data obtained from multidimensional scaling indicated that five perceptual dimensions were reported, which were ‘average pitch’, ‘direction’, ‘length’, ‘extreme endpoint’ and ‘slope’. The importance of each of the five perceptual dimensions varied depending on the language background of the listeners. The perceptual dimension ‘length’ was reported to be not applicable to tone languages. Tone languages including Cantonese, Mandarin, Taiwanese and Thai as well as a non-tone language, namely English, were investigated in the cross-language study of Gandour (1983). Results of multidimensional scaling showed that two-dimensional solution was optimal to the listeners' perception of tones. In other words, listeners had the best performances in dissimilating pairs of tones with the use of two perceptual dimensions and it was the appropriate number of dimensions underlying the perception of tones. The 'height' dimension and the 'direction' dimension were found to be more salient in tone perception. The former dimension was interpreted as the level of fundamental frequency and the latter dimension was interpreted as the direction of F_0 movement. For listeners from the non-tone language, the 'height' dimension was more important and the 'direction' dimension was least important than did for listeners from the tone languages. In addition, listeners from the three Chinese languages gave more weight to the 'height' dimension and less weight to the 'direction' dimension in comparison to Thai listeners. Gandour (1983) further proposed that despite the linguistic experience of listeners, the difference in relative importance of the two dimensions between different language group memberships was correlated with the types of lexical tones found in the tonal space of a language. For example, Cantonese is a tone language which has more tones in homogeneous direction of the F_0 movement. Therefore, Cantonese listeners tend to give less weight to the 'direction' dimension in contrast with Mandarin and Taiwanese listeners. In terms of the directional characteristic, the case of Meixian Hakka is similar to that of Cantonese. Three of the four long tones in Meixian Hakka, namely [51], [31] and [11], have falling contours and they are in relatively homogeneous direction. Taking the five perceptual dimensions of tone into account, Meixian Hakka listeners may attach less importance to the dimensions of 'direction', 'length' and 'extreme endpoint' in the perception of tones. The 'extreme endpoint' dimension is considered to be less important as the offsets of the three falling tones are at similar frequencies in the low frequency range. Hence, it is supposed that Meixian

Hakka listeners may place more emphasis on the dimensions of ‘average pitch’ and ‘slope’. In this study, ‘average pitch’ is investigated with reference to starting point and endpoint of F_0.

A perceptual experiment was designed based on the acoustic data obtained in this study to investigate the perceptual characteristics of the four long tones in Meixian Hakka. The potency of various perceptual cues is examined to evaluate their influence on the identification of Meixian Hakka tones. A matrix of F_0 onset values and F_0 offset values was proposed and it reflected all possible falling and level F_0 contours that commonly found in the real speech of Meixian Hakka. The purposes of this study are (1) to examine the potency of tonal height, slope, starting point and endpoint of F_0 contour in the identification of the target tones in Meixian Hakka, (2) to discuss whether the same perceptual cue or different perceptual cues are used for the identification of the four long tones in Meixian Hakka, (3) to investigate the perceptual space that occupied by the four long tones, and (4) to establish the correlation between the acoustic characteristics of the Meixian Hakka tones and listeners’ internal representation of the speech signal.

3.2 Methodology

3.2.1 Stimuli

A matrix was designed to examine the perceptual potency of (1) height, (2) slope, (3) starting point, and (4) endpoint of F_0 contour as shown in Figure 1. The matrix was composed of the F_0 onset variants and F_0 offset variants. Every stimulus had a frequency interval of 6 Hz. The F_0 onset continuum was divided in 17 steps, whereas the offset continuum was divided in 8 steps. The F_0 onset continuum had a varying F_0 with values ranging from 90 Hz to 186 Hz, and the F_0 offset values, combined with the F_0 onset variants, range from 84 Hz to 126 Hz. The stimulus set consisted of a total of

115 possible F_0 contours in Meixian Hakka. The 115 F_0 contours are all unidirectional with 7 level contours and 108 falling contours. Each of the synthesized F_0 contours in the stimulus set was superimposed on a base syllable [ts^hu]. That is, a set of 115 speech stimuli was constructed with varying pitch contours on a CV syllable which was phonetically transcribed as [ts^hu]. The original syllable of the base syllable in this study was [ts^hu11] which was produced by a male native speaker who was at the age of 53 at the time of recording. Syllable [ts^hu] associated with the low level tone [11] was chosen in order to minimize or even eliminate the influence of duration on the perception of tone. Acoustic data in chapter 2 showed that the duration of the low level tone [11] for most speakers are similar to the duration of the high to low falling tone [51] and the mid-high to low falling tone [41]. In contrast, the mid level tone [33] has the greatest duration among all. Therefore, it is assumed that the duration of [11] would not provide favorable condition to the identification of any of the four tones in Meixian Hakka.

The frequency range of the F_0 onsets and that of the F_0 offsets were designed according to the F_0 measurements of the four target tones in Meixian Hakka for the male speaker. All stimuli in this study were in equal duration of 421 ms, of which the vocalic portion of the syllable had the duration of 298 ms. With manipulation of the amplitude normalization function in CoolEdit 2000 software, the loudness of the stimuli was adjusted to approximately equal level. Praat 5.0.19 was used to synthesize the test stimuli. The original F_0 contour of the base syllable was removed and then replaced with each of the combinations of the F_0 onset and offset values in the matrix. All segmental information of the base syllable was retained. In this case, it is supposed that the difference in identification responses of tones would be resulted from the variation in F_0 onsets and offsets of the F_0 contour. As shown in Figure 3.1, the combinations of F_0 onsets and offsets that resemble the average F_0 patterns of the four tones [33, 11, 51,

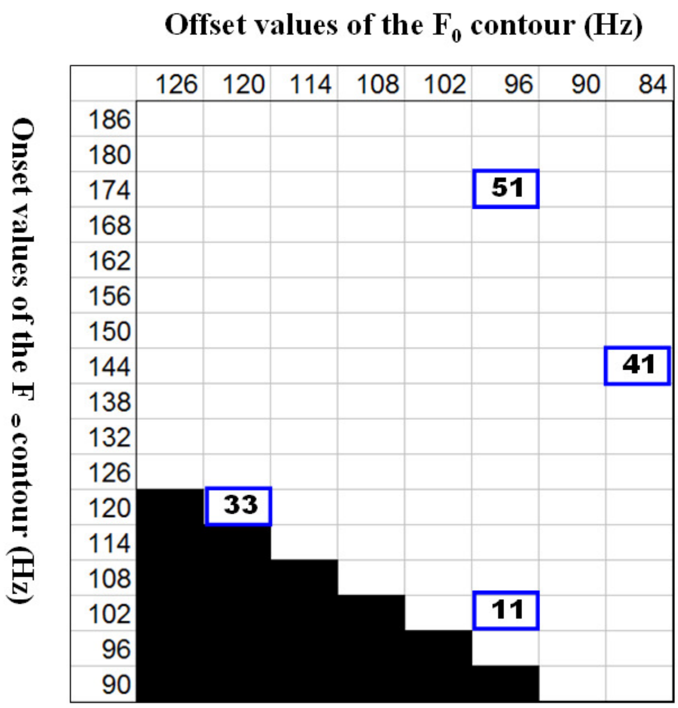


Figure 3.1 The 2-D perceptual matrix manipulated in the experiment showing the possible perceptual space for the four long tones in Meixian Hakka. The average acoustical spaces of the four tones based on the acoustic data are highlighted with the assigned pitch values.

41] produced by the male speaker are highlighted. A synthesized tone with a steady F_0 contour at 120 Hz resembles the average F_0 contour of the mid level tone [33] in a male voice; the synthesized tone with an F_0 onset of 102 Hz and an F_0 offset of 96 Hz resembles the average F_0 contour of the low level tone [11]; the synthesized tone with an F_0 onset of 174 Hz and an F_0 offset of 96 Hz resembles the average F_0 contour of the high to low falling tone [51]; the synthesized tone with an F_0 onset of 144 Hz and an F_0 offset of 84 Hz resembles the average F_0 contour of the mid-high to low falling tone [41].

3.2.2 Subjects

Thirty seven native speakers of Meixian Hakka were served as listeners in the perceptual experiments. They were twenty females and seventeen males who were born and grew up in the district of Meijiang. In order to make the results of the perceptual study comparable to the acoustic data obtained in chapter two, the listeners in this study were elderly. All native speakers were at the ages between 45 and 70 with no history of hearing loss. They were paid for their participation.

3.2.3 Procedure

Praat 5.0.19 was manipulated to play the stimuli. All stimuli were arranged in random order using the scripting language function in Praat software. In this case, all subjects had different randomization results. The stimulus set of the identification test included three repetitions of each stimulus. The stimuli were delivered to a Sennheiser HD340 headphone. Field trips in the district of Meijiang were carried out and the perceptual tests were conducted in a silent room. The test words used in the forced-choice identification test were meaningful CV syllables [ts^hu] that associated with the four long tones in Meixian Hakka. In other words, the test monosyllabic words for identification were [ts^hu33] 'to take out' 抽 , [ts^hu11] 'to remove' 除, [ts^hu 41] 'ugly' 醜 and [ts^hu 51] 'beast' 獸 in Meixian Hakka. Subjects were asked to identify the monosyllabic words in Chinese character that they heard. After a sound was played, the subjects had to click on one of a set of labeled keyboard keys showing the four test words. The interstimulus interval was 0.5 second. There was no limitation on the response time and the subjects would have had sufficient time to respond. The identification responses of subjects were recorded automatically by Praat software for further analysis. A trial test was conducted to acquaint subjects with the nature of the base syllable and the pitch range variation among different stimuli.

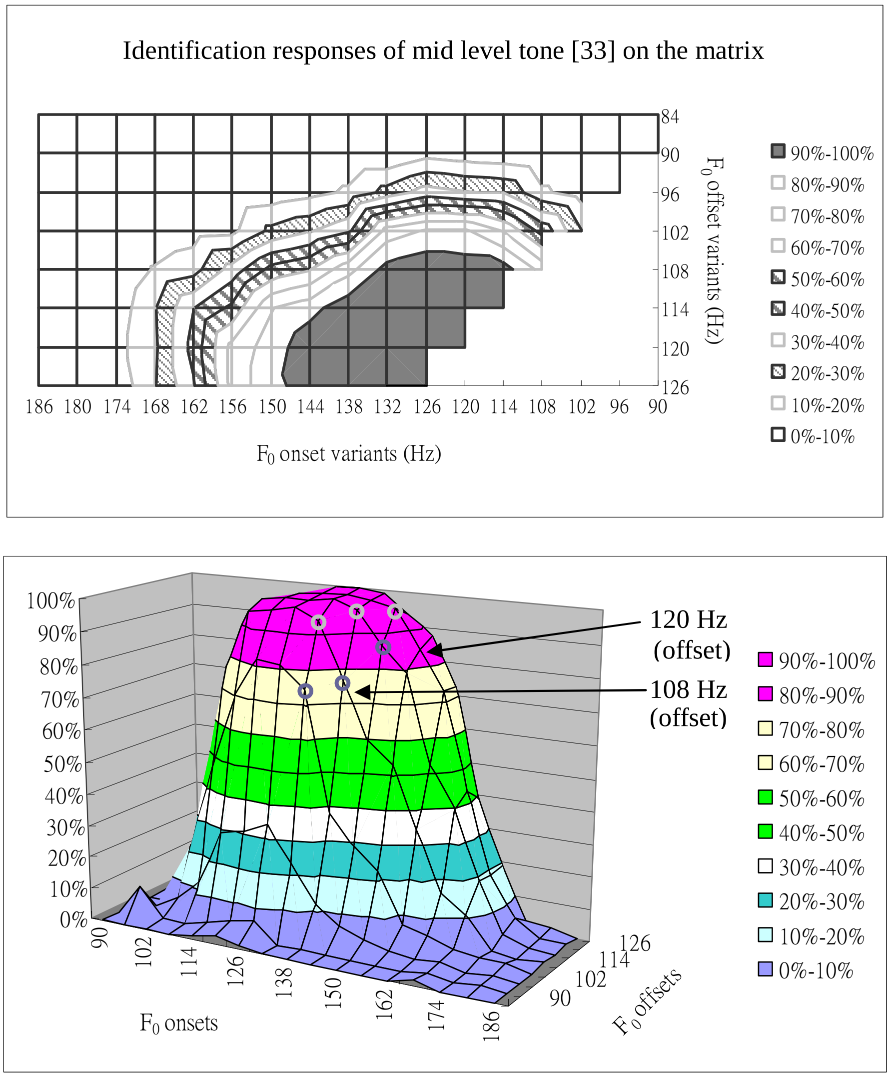
3.3 Results of the identification test

An analysis of the identification responses was performed with varying F_0 onset and offset values as factors. The identification responses of stimuli were obtained from 37 subjects, having a total of 111 data points (37 listeners x 3 repetitions) for each stimulus. Figures showing the identification responses in this chapter are in the same patterns. For the identification responses of each tone, a wireframe contour chart and a 3-D surface chart are presented. A wireframe contour chart provides a 2-D view of the surface chart from an upper view, which is similar to a 2-D topographic map. The abscissa of the contour chart indicates the varying F_0 onset values of the F_0 contour and the ordinate of the chart indicates the varying F_0 offset values. In 3-D surface chart, the F_0 onset values ranging from 90 to 186 Hz appear along the x-axis; the F_0 offset values ranging from 84 to 126 Hz appear along the y-axis; the percentages of tone identification are plotted along the z-axis. The blank space on the wireframe contour chart shows the perceptual space of tones that does not exist in Meixian Hakka. The stimuli in the matrix within the blank space have a rising F_0 contour, which are not found in the citation forms of Meixian Hakka tones. For the ease of data presentation, four percentage ranges of tone identification are highlighted in the 2-D wireframe contour chart. The four percentage ranges are: (1) 90% - 100% representing nearly perfect identification, (2) 40% - 60% indicating the perceptual boundary at 50% of tone identification and confusion occurs between the target tone and at least one of the other tones, (3) 20% - 30% displaying the perceptual boundary at chance level (25%) when there are four choices in the identification test, and (4) 0% - 20% showing seldom identification of the test word. Except these four ranges, other percentage ranges are also displayed for every 10% of tone identification with light gray lines. In the 3-D surface chart, different percentage ranges of tone identification are presented and seven percentage ranges, which are (1) 80% - 100% for nearly perfect tone identification, (2) 60% - 80%, (3) 40% - 60% for confusion level between two tones (50%), (4) 30% - 40%, (5) 20% - 30% for perceptual boundary at chance level (25%), (6) 10% - 20% and (7) 0% - 10% for seldom tone identification, are separated in different colors.

In the followings, the perceptual characteristics of each of the four long tones in Meixian Hakka will be analyzed and described. The perceptual saliency of tonal height, slope, starting point and endpoint of the F_0 contour in tone perception, as well as the perceptual space of the four tones correlated with the F_0 onsets and offsets will be discussed. The identification responses of subjects on the four target tones will be compared to examine whether confusion occurs in identification among the four tones and to investigate the patterns of confusion between the four tones. Lastly, the correlation between the acoustic characteristics and the perceptual results of the four long tones will be discussed.

3.3.1 The identification of the mid level tone [33]

The mid level tone [33] is the only one of four long tones in Meixian Hakka with a steady F_0 contour. It is expected that slope of the F_0 contour may play a more important role than tonal height in the perception of [33]. The identification results of the mid level tone [33] are displayed in Figure 3.2. Different percentage ranges of tone [33] identification are indicated with different gradients of gray as shown in Figure 3.2a. The shaded area in gray color on the matrix shows that for stimuli with F_0 onset values ranging from around 114 Hz to around 144 Hz and the F_0 offset values ranging from around 108 Hz to around 126 Hz, the percentage of mid level tone identification is over 90%. As shown in the figure, the mid level tone [33] occupies the bottom right corner of the perceptual matrix. That is, almost all Hakka listeners in this study identified a sound as syllable with the mid level tone [33] when the F_0 contour of the stimulus has an onset value at the middle region of the matrix and a corresponding offset value at the high frequency range of the matrix. This can also be found in Figure 3.2b. The 3-D surface chart shows that nearly perfect identification (80% - 100%) of [33] is distributed over



(a)

(b)

Figure 3.2 Identification responses of the mid level tone [33] on the matrix. Responses in different percentage category were presented in (a) a wireframe contour chart and (b) a 3-D surface chart. The data were pooled responses of 37 subjects in three repetitions of the identification test.

the central part of the perceptual space with high offset values. As can be seen, the onset and offset values of the stimuli for the identification of [33] are positively correlated. Within the aforementioned frequency range, a high F_0 onset has to correspond to a high F_0 offset, and vice versa, for a high identification of [33]. A comparison between the identification results of the four tones (Figure 3.2) and the designed matrix of the experiment that based on the acoustic data (Figure 3.1) shows that the synthesized steady tone at 120 Hz which resembles the average F_0 contour of [33] is located inside the shaded area of nearly perfect identification of [33]. In addition to the steady frequency trajectories at 114, 120 and 126 Hz, it should be noted that falling frequency trajectories were also perceived as the mid level tone [33] by the native speakers of Meixian Hakka. Figure 3.2a shows that listeners identified the stimulus as [33] for most of the time when the difference in frequency value between F_0 onset and F_0 offset was less than 24 Hz. For example, the identification of [33] is over 90% when stimuli have falling frequency trajectories with onset values at 144, 138 and 132 Hz and offset values at 120, 114 and 108 Hz, respectively. When the difference in frequency value between F_0 onset and F_0 offset becomes greater, i.e. over 24 Hz, the percentage of identification of [33] decreases. As shown in Figure 3.2b, falling frequency trajectories (in light gray circles) from 144 to 120 Hz, 138 to 114 Hz and 132 to 108 Hz, of which the difference between F_0 onset and F_0 offset was 24 Hz, were perceived as the mid level tone [33] over 90% of the time. In contrast, falling frequency trajectories (in dark gray circles) from 144 to 114 Hz, 138 to 108 Hz, and 132 to 102 Hz, of which the difference between F_0 onset and F_0 offset was over 24 Hz, were perceived as [33] around 80% of the time. In other word, listeners of Meixian Hakka would tolerate 24 Hz difference between F_0 onset and offset of a falling F_0 trajectory for the identification of the mid level tone [33]. If the difference between onset and offset values exceeds the acceptable F_0 range, the identification responses of [33] abruptly decrease.

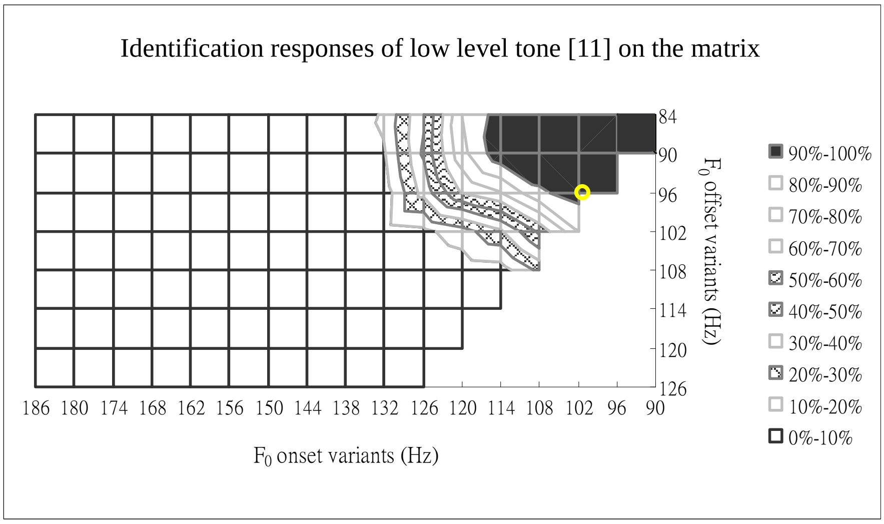
The shaded area in thick diagonal stripes in Figure 3.2a and the middle level of the dome-shaped figure in Figure 3.2b display the 40% - 60% of identification responses for the mid level tone [33]. This area indicates the confusion level at which the stimuli were identified as [33] by listeners only 50% of the time. Figure 3.2a displays that the identification function of [33] lowers to the confusion level (50%) when the frequency trajectories of stimuli have an onset value between 156 and 162 Hz at one end and at 108 Hz at the other end, and a corresponding offset value between 120 and 126 Hz and at 102 Hz, respectively. It can be noticed that in the relatively high frequency range, stimuli were perceived as [33] at the confusion level (50%) even when the difference between onset and offset of the F_0 trajectory was 42 Hz. For example, the falling F_0 trajectory with an onset value of 156 Hz and an offset value of 114 Hz was perceived as [33] 50% of the time. That is, listeners tend to overlook a great drop in frequency value at high frequency for the identification of the mid level tone [33]. In contrast, in the relatively low frequency range, the mid level tone [33] was perceived 50% of the time only when the F_0 trajectory was relatively steady. Listeners would tolerate just a small difference in frequency value between the onset and offset of the F_0 contour. For example, stimuli were identified as [33] at the confusion level when the frequency trajectories dropped from 108 to 102 Hz and from 114 to around 102 Hz. The shaded area in thin diagonal stripes in Figure 3.2a and the lower level of the dome-shaped figure in Figure 3.2b display the 20% - 30% of tone [33] identification in Meixian Hakka. The perceptual boundary, which is at the chance level (25%) of the identification, is located within this area. The chance level is 25% as there were four choices of identification in the perceptual test. The chance level indicates the perceptual boundary at which listeners identify the target tone without any cue but just by chance. The perceptual boundary at the chance level of tone [33] in Figure 3.2a shows that the mid level tone [33] was perceived when the frequency trajectories of the stimuli had an onset value between 102 and 168 Hz and an offset value between 96 and 126 Hz. The perceptual results of tone [33] at the chance level (25%) are similar to those of the tone at the confusion level (50%). Stimuli in the high frequency range (e.g., stimulus with a falling F_0 trajectory from 156 to 108 Hz) were identified as tone [33] by chance even though the difference between onset and offset of a falling F_0 trajectory was around 48 Hz. However, in low frequency range, listeners hardly identified the stimuli as tone [33] when the difference between onset and offset of a falling F_0 trajectory was 30 Hz or more.

In view of the identification responses of tone [33] at the confusion level and those at the chance level, the perceptual salience of tonal height and slope of F_0 contour across different frequency range seems to be different. A comparison of the relative importance of various perceptual cues on the identification of four Meixian Hakka citation tones will be discussed later. Perceptual results of this study indicate that tonal height plays a very important role in the identification of the mid level tone [33] in Meixian Hakka. As shown in Figure 3.2a, listeners may identify stimuli as the mid level tone [33] only when the onset variants of the stimuli are in the middle frequency range of the designed matrix and the offset variants are in the high frequency range of the matrix. At the confusion level (50%) of identification, the mid level tone [33] confuses with other tones in Meixian Hakka when the stimuli have an onset value lower than 108 Hz or higher than 162 Hz as well as an offset value lower than 102 Hz. It is surprised that slope of F_0 contour as a perceptual cue might not be as important as what is expected. Acoustic data in this study showed that native speakers in Meixian Hakka produce the mid level tone [33] with a relatively steady F_0 trajectory for the overall duration of the F_0 contour. However, it seems that the tone [33] is not perceived as what it reflects in the acoustic data and the perceptual results indicate that listeners can identify the tone [33] perfectly for stimuli with either a steady F_0 trajectory or a falling F_0 trajectory of less than 24 Hz.

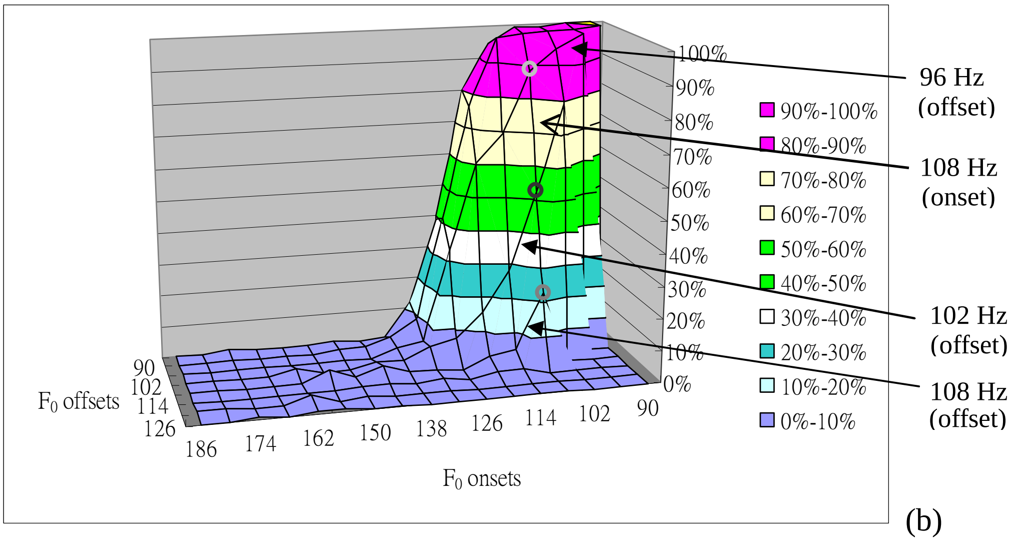
3.3.2 The identification of low level tone [11]

The identification results of the low level tone [11] are showed in Figure 3.3. As shown in Figure 3.3a, the shaded area in black color on the matrix indicates the nearly perfect identification of the low level tone [11]. Results show that listeners identified the stimuli as tone [11] for over 90% of the time when the frequency trajectories of the stimuli had an onset value between 90 and 114 Hz and a corresponding offset value between 84 and 96 Hz,. The shaded area of tone [11] occupies the top right corner of the perceptual matrix. This indicates that stimuli were easily perceived as tone [11] when they had an F_0 trajectory with low onset and offset values within an individual's pitch range. The perceptual responses of tone [11] on the matrix are also displayed in the 3-D surface chart. As shown Figure 3.3b, the responses of nearly perfect identification (80% - 100%) of tone [11] gather at the top right corner of the perceptual space with low onset and offset values. In comparison to the perceptual results of the mid level tone [33], the distribution of the nearly perfect identification for the low level tone [11] in the perceptual matrix is smaller. The identification responses of tone [11] in the Figure 3.3a show that more than four steps of the varying onset from 90 to 114 Hz and two steps of the varying offset from 84 to 96 Hz were perceived as the tone [11] for most of the time. On the other hand, stimuli within the shaded area of the perceptual matrix which covers near six steps of the varying onset from 114 to 144 Hz and more than three steps of the varying offset from 108 to 126 Hz were perceived as the tone [33] as shown in Figure 3.2a.

As mentioned in chapter two, acoustically the low level tone [11] has a slight falling F_0 trajectory. A comparison of the perceptual results in Figure 3.3a with the designed matrix as shown in Figure 3.1 indicates that the synthesized tone which resembles the average falling F_0 contour of the low level tone [11] from 102 to 96 Hz is found inside the shaded area of nearly perfect identification of tone [11] and it is near the margin of the area. The empty circle in Figure 3.3a displays the location of the average F_0 contour of the low level tone [11] on the confusion matrix. Similar to the perceptual results found in the identification of tone [33], falling frequency trajectories



(a)



(b)

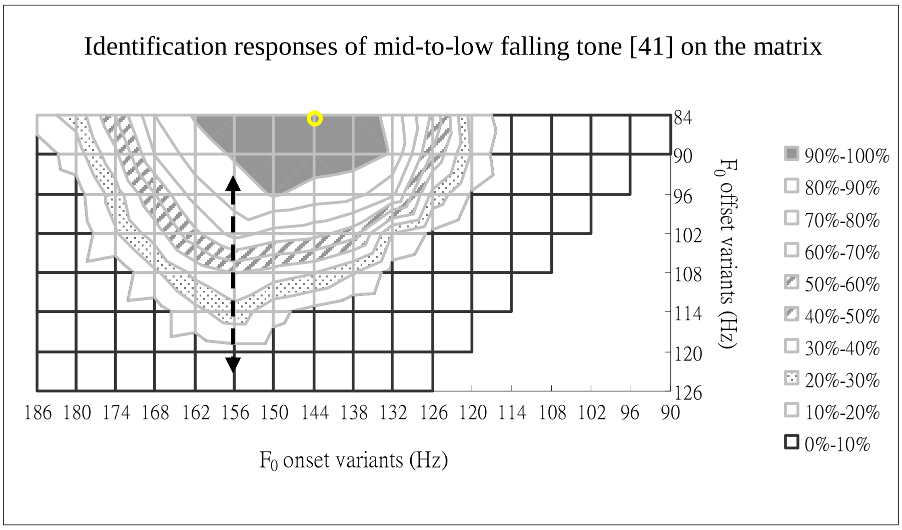
Figure 3.3 Identification responses of the mid level tone [11] on the matrix. Responses in different percentage category were presented in (a) a wireframe contour chart and (b) a 3-D surface chart. The data were pooled responses of 37 subjects in three repetitions of the identification test.

as well as steady frequency trajectories at 90 and 96 Hz were easily perceived as tone [11] by the Meixian Hakka listeners. Figure 3.3a shows that listeners identified the synthesized F_0 contour undoubtedly as tone [11] when the difference in frequency value between F_0 onset and offset was less than 30 Hz. For example, it can be seen that the stimulus which has a falling F_0 trajectory with an onset value at 114 Hz and an offset value at 90 Hz was perceived as tone [11] most of the time. Results show that for the identification of tone [11], the values of offsets within the shaded area can be predicted depending on the corresponding onset values as the onset and offset values are negatively correlated. As shown in Figure 3.3a, when the onset value of a stimulus is at a relatively high frequency of 114 Hz, its corresponding offset value for nearly perfect identification of tone [11] have to be low, i.e., either 84 or 90 Hz. In contrast, the frequency range of the offsets for the nearly perfect identification of tone [11] becomes wider (i.e., from 84 to 96 Hz) and the offset values can be higher when the onset values is at a relatively low frequency such as 102 Hz. It implies that for the identification of tone [11], listeners would tolerate a relative high offset value only when the corresponding onset value is low enough. For stimuli with a relatively high onset value, tone [11] would be easily identified only when the corresponding offset of the stimuli is at the lowest frequency range.

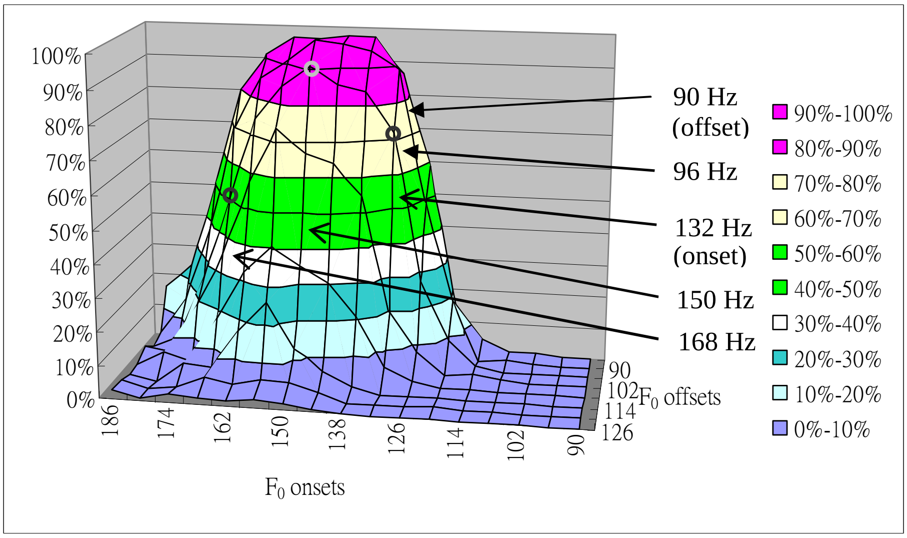
The shaded area in arrow stripes in Figure 3.3a and the middle level of the dome-shaped figure in Figure 3.3b display the 40% - 60% of identification for the low level tone [11]. The perceptual boundary indicating the confusion level (50%) of the identification is located within this area. Figure 3.3a shows that the identification function of tone [11] lowers to the confusion level when the frequency trajectories of the stimuli have an onset value ranging from 108 and 126 Hz, and a corresponding offset value ranging from 84 and 102 Hz. The results display that tone [11] was perceived when the F_0 trajectory was relatively steady. For example, listeners would perceive low level tone [11] at the confusion level when a stimulus has an onset value of 108 Hz and a corresponding offset value of 102 Hz. In this case, even though the offset value of a stimulus is relatively high, tone [11] may be perceived at 50% of the time as the F_0 trajectory of the stimulus is rather steady and the corresponding onset value is low. In contrast, when the onset value increases to 126 Hz, listeners may identify the stimuli as tone [11] 50% of the time only for stimuli with a corresponding low offset value at either 84 or 90 Hz. Based on the identification results, low offset is proved to be necessary for the perception of tone [11] and listeners would overlook the difference in frequency value between onset and offset of 42 Hz only when the offset value is at the lowest frequency range of an individual. The perceptual responses of tone [11] at chance level (25%) are displayed in cross stripes in Figure 3.3a. Stimuli with onset values between 108 to 126 Hz and offset values between 84 and 108 Hz were identified as tone [11] by chance. Similar to the identification results of tone [11] at the confusion level, stimuli with a relatively high onset value were perceived as tone [11] by chance only when the corresponding offset value was lower than 96 Hz. Figure 3.3a displays that the identification of tone [11] is better than by chance when the stimuli have an offset value of lower than 108 Hz. Perceptual results indicate that tonal height is the most potent cue in the identification of the low level tone [11] in Meixian Hakka. The low level tone [11] is perceived only in low frequency range of an individual. Listeners may confuse the low level tone [11] with other tones in Meixian Hakka when stimuli have an onset value of over 114 Hz and an offset value of over 96 Hz. A decrease in identification of tone [11] results from an increase in offset values is illustrated in Figure 3.3b. Here we use stimuli with an onset value of 108 Hz as an example. As indicated by a light gray circle at the top of Figure 3.3b, stimuli with an offset value of 96 Hz is perceived as tone [11] nearly 90% of the time. The identification of tone [11] lowers dramatically to nearly 50% when the offset value increases to 102 Hz as indicated by a dark circle. For stimuli with an offset value of 108 Hz as indicated by a dark grey circle at the lower part of the diagram, the identification of tone [11] further decreases and listeners identify the stimuli as tone [11] by chance. It may suggest that in addition to the tonal height which provides sufficient information for tone [11] identification, the endpoint of F_0 contour may facilitate the distinction between tone [11] and other tones, especially when the F_0 onset value is relatively high. In conclusion, the perceptual results of the low level tone [11] match the acoustic characteristics of the tone observed in this study. The low level tone [11] is perceived when stimuli (1) are in low frequency range of an individual's pitch range and (2) have either a relatively steady F_0 contour or a falling F_0 contour. When the offset value of the stimuli is significantly low, listeners can tolerate a 36 Hz difference between onset and offset values of the F_0 contour and the stimuli tend to be identified as tone [11] for more than 80% of the time.

3.3.3 The identification of mid-high to low falling tone [41]

The identification responses of the mid-high to low falling tone [41] are showed in Figure 3.4. As shown in Figure 3.4a, the shaded area in gray color on the matrix indicates the nearly perfect identification (90% - 100%) of tone [41]. Results show that stimuli with an F_0 onset value between 138 and 162 Hz and an F_0 offset value between 84 to 96 Hz were identified as tone [41] for more than 90% of the time. Figure 3.4a displays that the shaded area of tone [41] occupies the top central portion of the perceptual matrix. That is, Meixian Hakka listeners tend to identify stimuli as tone [41] when the F_0 onset values of the stimuli are in the middle range of the matrix and the corresponding F_0 offset values are in low frequency range. It can be noticed that the identification responses of the mid level tone [33] (Figure 3.2a) and the low level tone [11] (Figure 3.3a) resemble a pool-shaped area in the perceptual matrix. In contrast, a quadrilateral can be found for the identification results of tone [41]. The distribution of shaded area for tone [41] identification in the perceptual matrix is relatively small and it is similar to the distribution for tone [11] identification. As can be seen in the figure, almost five steps of the varying onset from 132 to 162 Hz and two steps of the varying offset from 84 to 96 Hz were perceived as tone [41] most of the time. This result is also



(a)



(b)

Figure 3.4 Identification responses of the mid level tone [41] on the matrix. Responses in different percentage category were presented in (a) a wireframe contour chart and (b) a 3-D surface chart. The data were pooled responses of 37 subjects in three repetitions of the identification test.

exhibited in Figure 3.4b. The 3-D surface chart shows that the nearly perfect identification (80% - 100%) of tone [41] is located at the top center of the perceptual space with low offset values. The empty circle in Figure 3.4a indicates the location of the stimulus that resembles the average F_0 contour of the mid-high to low falling tone [41] on the designed matrix. The F_0 contour of the stimulus falls from 144 to 84 Hz and it is inside the shaded area of nearly perfect identification of tone [41] as shown in the figure. Perceptual results show that only stimuli with a falling F_0 trajectory were perceived as the falling tone [41] by the native speakers of Meixian Hakka but not the stimuli with a steady F_0 trajectory. Listeners tend to identify the stimuli as tone [41] for more than 90% of the time when the difference in frequency value between the onset and offset is greater than 48 Hz and smaller than 78 Hz. For example, stimuli with a falling F_0 trajectory of which the onset value is 138 Hz and the offset value is 90 Hz were perfectly perceived as the mid-high to low falling tone [41]. It can be noticed that, on one hand, when the onset value of the stimuli decreases from 162 to 150 Hz, the frequency range of the corresponding offset values widen from 84 to 96 Hz for the nearly perfect identification of tone [41]. This indicates that when stimuli have a relatively high onset value (which is over 150 Hz as shown on the matrix), a low corresponding offset value is crucial for the identification of tone [41] as no or little perceptual cue of tone [41] can be retrieved from the high F_0 onset. On the other hand, when the onset value of the stimuli decreases from 150 to 132 Hz, the frequency range of the corresponding offset values narrow from 96 to 90 Hz for the nearly perfect identification of tone [41]. Based on the obtained acoustic data, the falling nature of the F_0 contour is supposed to be a potent characteristic of the mid-high to low falling tone [41] in Meixian Hakka. This may explain the decrease in offset value from 96 to 90 Hz when the corresponding onset value decreases from 150 to 132 Hz. As the slope of the falling contour is a potent cue for the identification of tone [41], stimuli with a low onset value tend to have a corresponding low offset value in order to maintain the slope of the falling.

The shaded area in diagonal stripes in Figure 3.4a and the middle level of the dome-shaped figure in Figure 3.4b show the 40% - 60% of identification for the mid-to-low falling tone [41] in Meixian Hakka. As can be seen in Figure 3.4a, the identification function of the falling tone [41] lowers to the confusion level (50%) when stimuli have an F_0 trajectory with an onset value ranged between 126 and 174 Hz and a corresponding offset value ranged between 84 and 108 Hz. The identification responses of tone [41] decreases to the confusion level (50%) when the onset and offset of stimuli have a distance of 6 to 12 Hz from the perceptual boundary of the nearly perfect identification of tone [41]. It can be noticed that the frequency range of the corresponding offset values decreases gradually from 108 to 84 Hz (1) when the onset values increase gradually from 156 to 174 Hz for stimuli with high onset values of more than 156 Hz as shown on the left of the dashed arrow line, and (2) when the onset values decrease gradually from 156 to around 132 Hz for stimuli with onset values of less than 156 Hz as shown on the right of the dashed arrow line. As discussed in the perceptual responses of nearly perfect identification of tone [41], there is a decrease in frequency range of the offset values for stimuli with high onset values as low offset values tend to become a potent perceptual cue for the identification of tone [41] when the high onset values show little information of tone [41]. This can explain the decrease in offset value in the first case. In contrast, the frequency range of the offset values deceases in the second case when the corresponding onset values are relatively low as a low offset value is necessary to maintain the slope of falling in order to distinguish the falling tone [41] from the low level tone [11]. It is obvious that the identification percentage decreases from over 90% to the confusion level (50%) when the offset values of stimuli increase and is over 96 Hz. The shaded area in dots in Figure 3.4a displays the identification responses of tone [41] at the percentage range of 20% - 30%. Results show that the identification responses decrease from the confusion level (50%) to the chance level (25%) when there is an increase or a decrease of around 6 Hz in either the onset value or the offset value. In addition, perceptual asymmetry can be found in the variation of identification responses between different frequency ranges.

Figure 3.4a indicates that the frequency range of onset tend to be separated at 150 Hz. Figure 3.4b displays that stimuli with an onset value of 150 Hz and an offset value of 96 Hz as indicated by a light gray circle was identified as tone [41] 90% of the time by the listeners. Results show that the identification responses of tone [41] decrease sharply when the onset values of the stimuli increase from 150 Hz, whereas the decrease in identification responses of tone [41] is relatively gentle when the onset values of the stimuli decrease from 150 Hz. The dark circles in Figure 3.4b indicate the identification responses of tone [41] when the offset of the stimuli is 96 Hz and the corresponding onsets are three steps apart from 150 Hz. Results indicates that there is a 40% decrease in identification of tone [41] when the high onset of the stimuli increases from 150 to 168 Hz. In contrast, the identification responses of tone [41] only decrease 20% (from 90% to 70%) as the onset of the stimuli decrease from 150 Hz to 132 Hz. This indicates that the decrease in percentage of tone [41] identification for stimuli with an onset value in the high frequency range is twice as much as for stimuli with an onset value in the relatively low frequency range. In conclusion, perceptual results of the mid-high to low falling tone [41] indicate that both slope and endpoint of the F_0 contour may play an important role in the identification of tone [41] in Meixian Hakka. Endpoint of the F_0 contour tends to be the potent cue in the identification when the F_0 contours of stimuli start with a high onset value. In contrast, slope of falling tends to be important for the identification when the F_0 contours of stimuli start with a relatively low onset value. As shown in Figure 3.4a, stimuli were identified as tone [41] when the onset values of stimuli are in the middle frequency range of an individual and the offset values are in the low frequency range.

3.3.4 The identification of high to low falling tone [51]

The identification responses of high to low falling tone [51] are showed in Figure 3.5. As shown in Figure 3.5a, the shaded area in light gray color on the matrix indicates the nearly perfect identification of tone [51]. Identification responses display that stimuli with onset values ranging from 174 to 186 Hz, and offset values ranging from 90 to 126 Hz, were identified as tone [51] over 90% of the time. As shown in Figure 3.5a, the shaded area for the identification of tone [51] occupies the left of the perceptual matrix. Despite the offset value of stimuli, stimuli with an onset value at the high frequency range of an individual tend to be identified as tone [51] by the Meixian Hakka listeners. It can be noticed that the pattern of tone [51] identification on the matrix as shown in Figure 3.5 is different from the patterns of identification of tones [33, 11, 41]. The distribution of shaded area for tone [51] identification in the perceptual matrix is relatively broad in comparison to the perceptual results of other tones in Meixian Hakka. As can be seen in the Figure 3.5a, almost all steps of the varying offsets from 84 to 126 Hz and two steps of the varying onsets from 174 to 186 Hz were included in the shaded area of the nearly perfect identification of tone [51]. This result is also displayed in Figure 3.5b. The 3-D surface chart shows that the nearly perfect identification (80% - 100%) of falling tone [51] is found on the left of the perceptual space that associated with high onset values and it comes across nearly all offset values on the matrix. When comparing the perceptual results in Figure 3.5 to the location of the stimulus that resembles the average F_0 contour of the high to low falling tone [51] in the designed matrix in Figure 3.1, it is surprised that the stimulus with an onset value of 174 Hz and an offset value of 96 Hz, which is the resemblance of the F_0 contour of tone [51], is not included in the shaded area of the nearly perfect identification of tone [51]. The location of the stimulus that resembles the F_0 contour of tone [51] is indicated by the empty

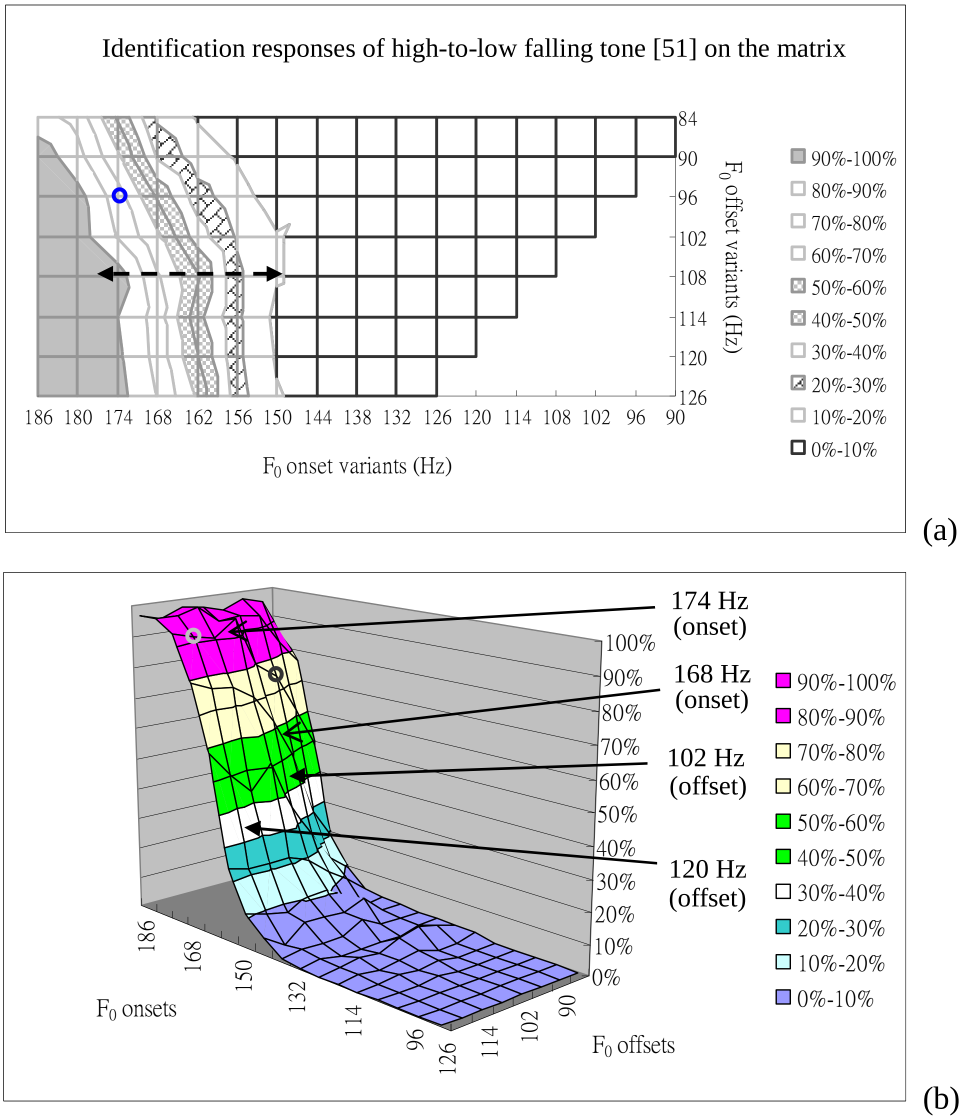


Figure 3.5 Identification responses of the mid level tone [51] on the matrix. Responses in different percentage category were presented in (a) a wireframe contour chart and (b) a 3-D surface chart. The data were pooled responses of 37 subjects in three repetitions of the identification test.

circle in Figure 3.5a. It shows that the stimulus was identified as tone [51] only around 70% of the time and there is confusion between tone [51] and tone [41]. This finding cannot be explained and further investigation may be needed. Acoustically, the high to low falling tone [51] has a delayed falling F_0 trajectory and the onset value of tone [51] is the highest as compared to other tones in Meixian Hakka. As shown in Figure 3.5a, only stimuli with a sharp falling F_0 trajectory were perceived as tone [51]. Meixian Hakka listeners tend to identify stimuli as tone [51] for more than 90% of the time when the difference between the onset and offset values is more than 48 Hz. For example, stimuli with a falling F_0 trajectory that drops from 174 to 126 Hz were perfectly identified as tone [51]. The differences between the onset and offset values are even greater for other stimuli in the shaded area. The perceptual results indicate that the F_0 onset value plays an important role in the identification of tone [51] in Meixian Hakka. That is, the high onset value of stimuli may contain sufficient information that elicits the identification of the high to low falling tone [51]. In contrast, the offset value appears to be less crucial in the identification. This may be explained in consideration of the distribution of F_0 contours of the four long tones in Meixian Hakka. As the F_0 onsets of tones [33, 11, 41] and the F_0 offsets of tones [33, 11, 41] in Meixian Hakka are in the middle and low frequency ranges, a high F_0 onset may be sufficient to distinguish the high to low falling tone [51] from other tones. That may be why stimuli with an onset value at the high frequency range were perceived as tone [51] regardless of the corresponding offset value.

The shaded area in pattern of check in Figure 3.5a indicates the 40% - 60% of identification for tone [51] and the shaded area in brick pattern in Figure 3.5a displays the identification of 20% -30% which indicates the perceptual boundary at chance level (25%). As can be seen, the identification function of tone [51] is at the confusion level (50%) when stimuli have an onset value between 162 and 174 Hz and a corresponding offset value between 126 and 84 Hz. Results show that the identification responses of tone [51] decrease from over 90% to the confusion level (50%) when there is a two-step increase of 12 Hz in the onset value. The perceptual boundary at chance level is located on the matrix where the onset values are between 156 and 168 Hz and the corresponding offset values are between 126 and 84 Hz. As shown in Figure 3.5, the dashed arrow line divides the matrix into two parts according to the F_0 offset value. For stimuli with an offset value of over 108 Hz as indicated by the area below the dashed arrow line, the identification responses of tone [51] remain relatively constant and the percentage of identification is not affected by the variation of offset value from 108 to 126 Hz. In contrast, for stimuli with an offset value of below 108 Hz as indicated by the area above the dashed arrow line, there is a shift in identification responses of tone [51] towards left of the matrix. It shows that the identification percentage of tone [51] for stimuli with a high onset value decreases when the offset values of stimuli decrease gradually from 108 to 84 Hz. In addition, stimuli with a low offset value of 90 Hz or 96 Hz would be identified perfectly as tone [51] only if the corresponding onset value is over 180 Hz.

In summary, perceptual results of tone [51] indicate that the starting point of F_0 contour is a potent cue in tone [51] identification in Meixian Hakka. Regardless of F_0 offset value, stimuli tend to be identified as tone [51] when the onset values of stimuli are at the high frequency range of an individual. There is a confusion of identification between the high to low falling tone [51] and the tones [33, 41] in Meixian Hakka when the high onset value of stimuli decreases. For stimuli with an offset value of above 108 Hz, starting point of F_0 contour at the highest frequency may be sufficient for the identification of tone [51] and the variation in F_0 offset value would not result in a change in the percentage of tone [51] identification. In contrast, for stimuli with an offset value of below 108 Hz, a relatively high F_0 offset value may provide a more favourable condition for the identification of tone [51] than do the low F_0 offset value such as 84 Hz and 90 Hz. It can be noticed that the higher the F_0 onset value, the wider the range of the corresponding F_0 offset value for tone [51] identification. As there are acoustically three falling tones, namely [51, 41, 11], in Meixian Hakka, it is hypothesized that slope of F_0 contour would be an important cue to distinguish between the three tones and higher identification responses of tone [51] should be achieved for stimuli with a steeper slope of the F_0 contour. However, perceptual results of tone [51] do not match the hypothesis. As can be seen in Figure 3.5b, stimuli with an onset value of 174 Hz and a corresponding high offset value of 120 Hz were perceived as falling tone [51] over 90% of the time as indicated by the light gray circle. When the offset value lowers to 96 Hz and the slope of the F_0 contour becomes steeper, the identification responses of tone [51] surprisingly decrease to around 70% as indicated by the dark gray circle. This indicates that high starting point of F_0 contour instead of steep slope is proved to be a potent cue for the identification of the high to low falling tone [51] in Meixian Hakka.

3.4 Discussion

The perceptual results of each of the four tones [33, 11, 41, 51] in Meixian Hakka have been described. Results show that [33] is perfectly identified by listeners when stimuli has either a steady F_0 trajectory or a slight falling F_0 trajectory and with an F_0 onset value in the middle frequency range and a corresponding F_0 offset value in the relatively high frequency range. The low level tone [11] is perfectly identified when the F_0 contour of stimuli is either steady or falling and both the onset and offset of stimuli are in the lowest part of an individual's frequency range. Tone [41] is perfectly identified when F_0 contour of stimuli is falling and stimuli have an onset value in the middle frequency range and an offset value in the low frequency range, whereas nearly perfect identification of tone [51] is achieved when F_0 contour of stimuli is falling and stimuli has an onset value in the highest part of an individual's frequency range. To investigate the importance of different perceptual cues to identification of tones in Meixian Hakka, a comparison of identification results of the four tones is presented. A comparison of identification responses of the four tones [33, 11, 41, 51] in Meixian

Hakka is displayed in Figure 3.6. Two different percentage ranges of tone identification are highlighted, which are (1) percentage range of 80% - 100% representing nearly perfect identification and (2) percentage range of 40% - 60% indicating the location of perceptual boundary at confusion level. Within the area of nearly perfect identification, there is a line for each tone separating 90% - 100% of identification in the core part of the shaded area from 80% - 90% of identification. For the ease of data presentation, the identification responses of 20% - 30% at chance level are not displayed. Figure 3.6 shows that the identification responses of the four tones at confusion level (40% - 60%) almost completely overlap with each other and it leaves only a small space in between uncovered. This indicates that listeners tend to confuse one tone with another one, but not with different tones when stimuli are at the perceptual boundaries on the matrix. In this case, perceptual boundaries at chance level (25%) seem to be less significant than

Identification responses of four Meixian Hakka tones [33, 11, 41, 51] on the matrix

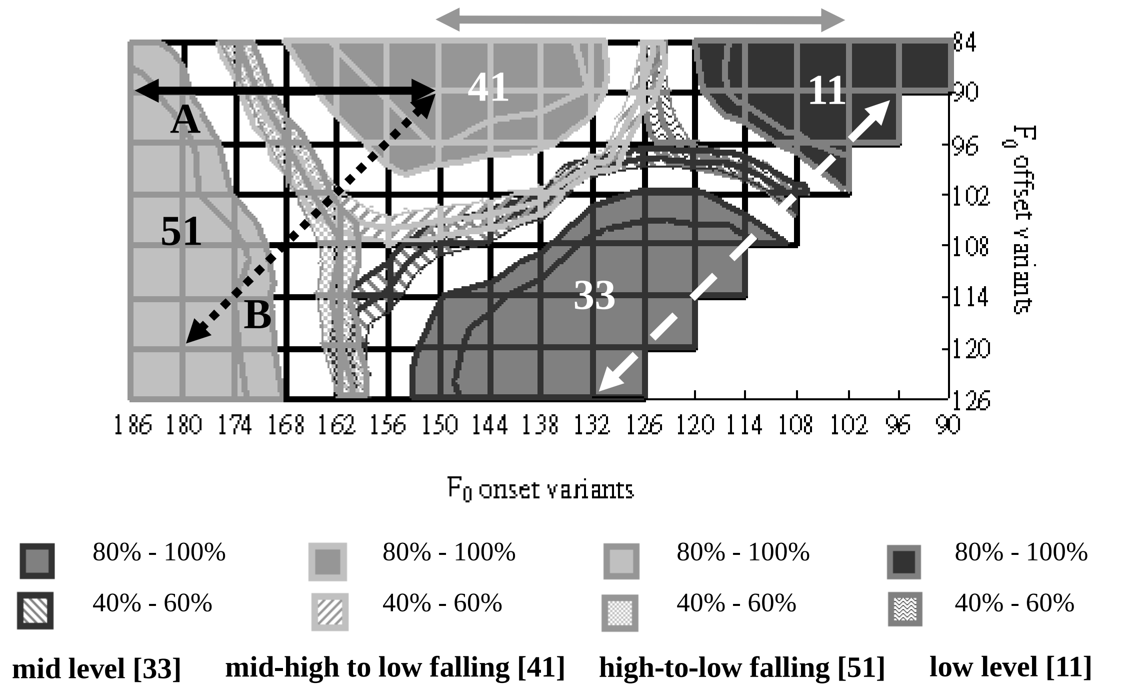


Figure 3.6 Identification responses of the four long tone [33, 11, 41, 51] on the matrix. Responses in different percentage category were presented in wireframe contour chart. The data were pooled responses of 37 subjects in three repetitions of the identification test.

the perceptual boundaries at confusion level (50%). As shown in Figure 3.6, nearly perfect identifications (80% - 100%) of the four tones are displayed in shaded area with different gradient of gray. There is no overlap between the four tones and each tone is located at a specific area within the perceptual space. It seems that listeners may manipulate different perceptual cues to distinguish the four tones in Meixian Hakka and the synthesized F_0 contours on the matrix are perceived as different tones by matching acoustic features of stimuli to pattern of tone perception. By comparing the identification responses of the four tones [33, 11, 41, 51] as shown on the stimulus matrix in Figure 3.6, the perception pattern of tones in Meixian Hakka is investigated in terms of F_0 onset value and F_0 offset value.

Taking F_0 onset value into account, the frequency range of the matrix can be separated into three parts, which are high, mid and low, according to the perceptual responses of tones in Meixian Hakka. As shown in Figure 3.6, stimuli are perceived as either the mid level tone [33] or the mid-high to low falling tone [41] when the onset of stimuli is in the middle frequency range. The high to low falling tone [51] is perceived when stimuli have an onset value at the highest part of frequency range, whereas the low level tone [11] is perceived when stimuli have an onset value at the lowest part of the frequency range. Regardless of F_0 offset value, tone [51] is perfectly identified when stimuli have an onset value of over 168 Hz. High F_0 onset is shown to be a sufficient cue to distinguish tone [51] from the other three tones [33, 11, 41] as shown in Figure 3.6. This finding matches the acoustic characteristic of tone [51] that the highest point of F_0 contour of [51] is always at the highest part of an individual's frequency range. It can be observed that stimuli with an onset value at the middle frequency range tend to be perceived as either [33] or [41] and the frequency range of onset value for the nearly perfect identification of the two tones are generally overlapped. This indicates that it may be hard for the two tones to distinguish from one another in terms of F_0 onset value. This finding may be induced from the acoustic characteristic of [33] and [41] that onsets of F_0 contours of the two tones are both in the middle frequency range of an individual. Perceptual results of [11] displays that stimuli with an onset value at the low frequency range are perfectly perceived as the low level tone [11]. This finding matches the acoustic results of [11] that the overall F_0 contour of the tone is in the low frequency range. However, regardless of F_0 offset, there is an overlap in nearly perfect identification between [11] and [33] when stimuli have an onset value between 108 and 120 Hz as shown in Figure 3.5. Acoustically, the onset of F_0 contour of [33] is significantly higher than that of [11] [t (9) = 6.015, $p < .000$] for all speakers. This indicates that the perceptual result does not match the acoustic characteristic of tone [11] and low onset value alone may not provide a sufficient cue for the tone [11] perception.

Regarding the offset value, the frequency range of the matrix can be separated into two parts, which are high and low, according to the identification responses of the tones in Meixian Hakka as shown in Figure 3.6. It can be seen that the nearly perfect identification of tone [51] is found across almost all offset values. In this case, the potency of offset value in the identification of tone [51] tends to be low even though acoustic data in this study showed that the F_0 contour of [51] for all speakers has a low F_0 offset. Such difference between the perceptual result and the acoustic feature of tone [51] seems to be reasonable as the acoustic feature of low F_0 offset is found in many tones, namely [51, 41, 11], in Meixian Hakka and the acoustic feature of tone [51] having a high F_0 onset tends to be an important perceptual cue to distinguish tone [51] from other tones in Meixian Hakka. Figure 3.6 displays that F_0 offset value tend to provide a sufficient cue to distinguish tone [33] from the other two tones [41, 11] in Meixian Hakka. As can be seen, regardless of onset value, stimuli with an offset value of over 102 Hz tend to be perfectly identified as tone [33], whereas stimuli with an offset value of below 102 Hz tend to be identified as either [41] or [11]. In consideration of F_0 offset value, the overlap of nearly perfect identification of tones [41, 11] in the frequency range between 84 and 102 Hz indicates that F_0 offset is not a potent cue for the differentiation between [41] and [11] in Meixian Hakka. The finding matches the acoustic characteristics of tones [33, 11, 41] that the offset of tone [33] is at the middle frequency range of an individual, whereas the offsets of tones [41, 11] are at the low frequency range.

To investigate the potency of different perceptual cues for perception of tones [33, 11, 41, 51] in Meixian Hakka, different dimensions showing the shift in identification from one tone to another tone are analyzed. Regarding the perception of tones in Meixian Hakka, three assumptions are made based on the acoustic data presented in chapter two. The first assumption is the difference between [33] and [11] would be the tonal height as the F_0 contour of [33] is in the higher frequency range than that of [11]. The second assumption is the difference between [41] and [11] would be slope of F_0 contour as [41] has a relatively steep falling contour from the mid frequency range to the low frequency range while [11] has a slight falling contour in the low frequency range. The third assumption is the difference between [51] and [41] would be slope of F_0 contour as the offsets of both [51] and [41] are in the low frequency and the F_0 contour of [51] with a higher onset seems to have a steeper slope than that of [41] with an onset in the mid frequency range. For the ease of presentation, sketches displaying the shift in identification from one tone to another are shown in Figure 3.7. The four sketches in Figure 3.7 correspond to different dimensions of the perceptual matrix in Figure 3.6. Figure 3.7a showing the perceptual boundary between [33] and [11] corresponds to the dashed arrow line in Figure 3.6 which displays the dimension taken

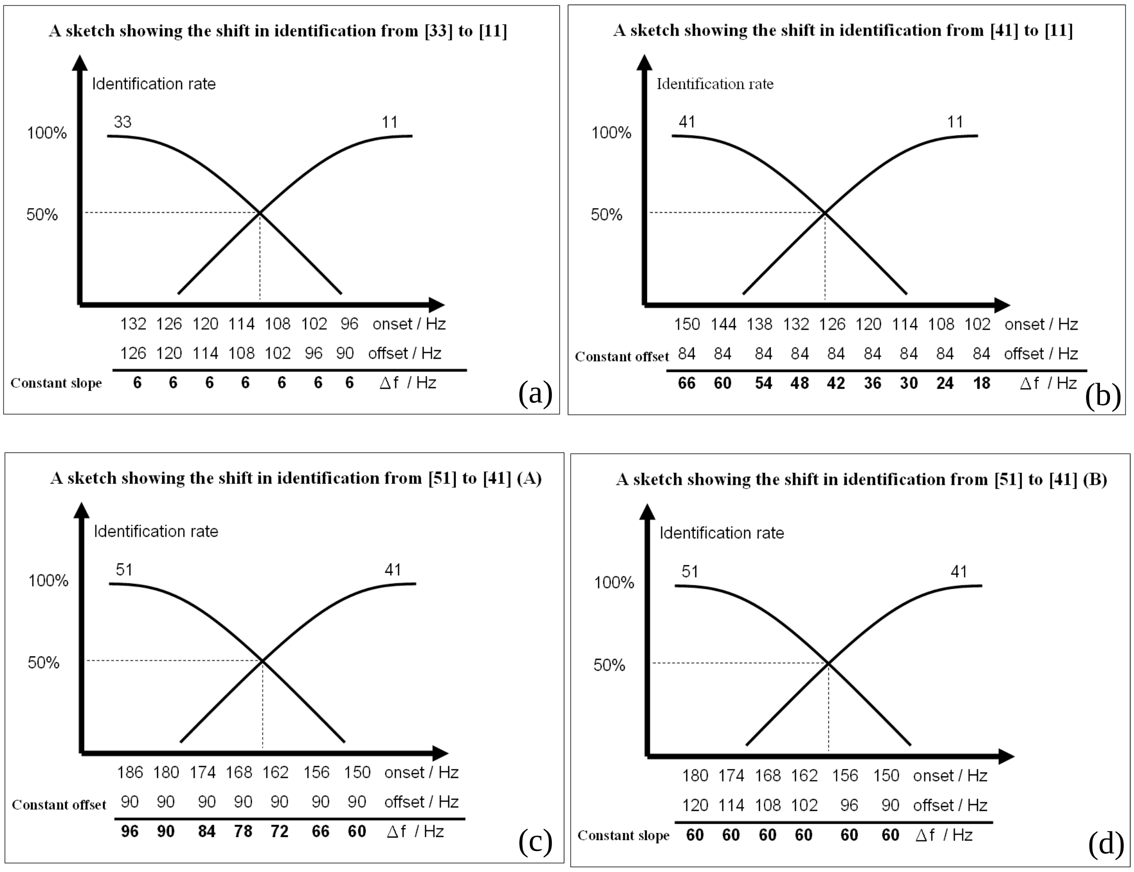


Figure 3.7 Sketches showing the shift in identification (a) from [33] to [11]; (b) from [41] to [11]; (c) from [51] to [41] with constant offset value; and (d) from [51] to [41] with constant slope of F_0 contour

along the matrix showing identification percentage of stimuli with varying onset and offset values. Figure 3.7b showing the perceptual boundary between [41] and [11] corresponds to the light gray arrow line in Figure 3.6 which exhibits the dimension taken along the matrix showing identification percentage of stimuli with a constant offset value of 84 Hz. In this study, two dimensions (A and B) along the matrix regarding the identification of [51] and [41] are investigated. Figure 3.7c showing the perceptual boundary between the two tones corresponds to the dark arrow line in Figure 3.6 which displays the dimension A showing the identification percentage of stimuli with a constant offset value of 90 Hz. In contrast, Figure 3.7d showing the perceptual boundary between the two tones corresponds to the dotted arrow line in Figure 3.6 which exhibits the dimension B showing the identification percentage of stimuli with varying onset and offset values. As displayed in Figure 3.7, onset values, offset values and difference in frequency value between onset and offset which indicates slope of F_0 contour are shown in x-axis, while the identification percentage of stimuli is sketched in y-axis.

The shift in identification from [33] to [11] as shown in Figure 3.7a displays that tonal height is an important cue for the distinction between [33] and [11] in Meixian Hakka. As shown in Figure 3.7a, the slope of F_0 contour is constant for stimuli along this dimension. F_0 onset and offset values are the varying factors which indicate change in tonal height. Results show that the identification shifts from [33] to [11] when the absolute tonal height decreases. In other words, keeping the slope of F_0 contour constant, stimuli with a higher tonal height are perceived as the mid level tone [33], whereas stimuli with a lower tonal height are perceived as the low level tone [11]. Tonal height as the potent perceptual cue to distinguish between the two level tones [33, 11] matches the first assumption based on the acoustic data. As shown in Figure 3.7b, the F_0 values keeps constant at 84 Hz for stimuli along the dimension and the varying factor is slope of F_0 contour. Perceptual results show that there is a shift in identification from [41] to [11] when the slope of F_0 contour decreases. That is, keeping the offset value constant, stimuli are perceived as [41] when the slope of F_0 contour is relatively steep. In contrast, stimuli with a gentle slope are identified as [11] instead of [41]. This indicates that slope is an important cue for the distinction between the mid-high to low falling tone [41] and the low level tone [11] in Meixian Hakka. Similar to the case of the two level tones [33, 11], slope as the potent cue to distinguish between [41] and [11] matches the second assumption based on the acoustic data. As shown in Figure 3.6, two dimensions (A and B) are investigated for the distinction between the high to low falling tone [51] and the mid-high to low falling tone [41] as different perceptual cues seem to be manipulated for tone perception in these two dimensions. Figure 3.7c shows the dimension A along the perceptual matrix. It can be seen that the F_0 offset value is constant along the continuum and the decrease in F_0 onset value results in the decrease in slope of F_0 contour. Results show that the identification shifts from [51] to [41] when there is a decrease in slope. In other words, stimuli with a steep falling contour are easily identified as [51], whereas stimuli with a less steep contour are identified as [41]. This indicates that slope is the potent cue for the distinction between [51] and [41] and this finding matches the third assumption that based on the acoustic data. However, the importance of tonal height for the perception of [51] and [41] is found in another dimension as shown in Figure 3.7d. In dimension B, the slope of F_0 contour along the continuum keeps constant and tonal height changes along the continuum. Results show that the identification shifts from [51] to [41] when the tonal height decreases. In other words, keeping the slope of falling F_0 contour constant in dimension B, stimuli in the relatively high frequency range are identified as [51], whereas stimuli in the lower frequency range are identified as [41]. Tonal height as a potent cue for the distinction between [51] and [41] does not match the third assumption that slope should be the only potent cue for the perception of the two falling tones. The perceptual results of the two dimensions indicate that both slope of F_0 contour and tonal height are important for the distinction between [51] and [41] in Meixian Hakka. Listeners tend to manipulate different perceptual cues in the identification of [51] and [41] when either one of the perceptual cues does not provide sufficient information for tone perception.

In conclusion, this perceptual study investigated the perceptual characteristics of the four long tones [33, 11, 41, 51] in Meixian Hakka. Results show that tonal height is the potent cue for the identification of the two level tones [33, 11]. The mid level tone [33] perceptually has a higher tonal height than the low level tone [11]. Stimuli are identified as level tone when they have either a steady frequency trajectories or a slight falling F_0 trajectory with a less than 30 Hz drop. The starting point of the F_0 trajectory plays an important role in the identification of the high to low falling tone [51] in Meixian Hakka. Results show that regardless of F_0 offset value, stimuli with an onset values at the highest part of an individual's pitch range are perceived as [51]. A comparison of identification responses between [51] and [41] shows that both slope and tonal height are the potent cues for the distinction between the two tones. On one hand, tone [51] has a steeper slope than tone [41] in tone perception. On the other hand, stimuli in the higher frequency range are identified as [51], but not as [41]. It is assumed that the importance of tonal height for the distinction of [51] and [41] may be mainly related to the importance of high F_0 onset for tone [51] perception as F_0 offset of tone [51] seems to be perceptually insignificant. The perceptual salience of slope is shown to be potent to distinguish the mid-high to low falling tone [41] from the high to low falling tone [51] and the low level tone [11]. Results show that the F_0 contour of tone [41] perceptually has a steeper slope than that of tone [11], but has a less steep slope than that of tone [51]. A comparison between the potency of different perceptual cues in tone perception and the three assumptions based on the acoustic data shows that the perceptual patterns of tones in Meixian Hakka generally match the acoustic characteristics of the tones. It indicates that there is a strong correlation between acoustics and perception of tones in Meixian Hakka.

CHAPTER FOUR
FORMANT FREQUENCY AND TEMPORAL ORGANIZATION OF VOWELS, DIPHTHONGS AND TRIPHTHONGS

4.1 Introduction

This chapter describes the spectral properties of monophthongs, diphthongs and triphthongs in Meixian Hakka. The vowel system of Meixian Hakka has been described in the past studies (Hashimoto, 1973; Huang, 1992; Wen, 2006; Xie, 1994; Yuan et al., 2001). Though there were published descriptions of the sound system of Meixian Hakka, most of the authors of the past studies were dialectologists and the phonemic transcriptions were mainly based on an individual's impressionistic description. So far, no experimental data on formant frequencies and duration of the Meixian Hakka vowels has been presented. As discrepancies are found among the past studies, it is worthwhile to investigate the spectral characteristics of vowels, diphthongs and triphthongs using speech analysis software. In this study, the recordings of monophthongs, diphthongs and triphthongs were analyzed from a phonetic perspective. The frequency values of the first three formants (F_1, F_2 and F_3) and the temporal organization of vowels are to be presented. The phonetic characteristics of vowels, diphthongs and triphthongs in Meixian Hakka are to be uncovered and the transcription of the Meixian Hakka vowel system is to be established based on the acoustic data. Furthermore, the results of this study are to be compared with the findings of the past studies on other Chinese dialects as well as on other languages.

Table 4.1 displays the transcriptions of vowels, diphthongs and triphthongs in open syllables, i.e., (C)V, (C)D and (C)T syllables, in Meixian Hakka described by different linguists. As can be seen, discrepancies in the transcription of Meixian Hakka vowels occur among the past studies. It was generally agreed that there are 5 vowels [i, e, a, o, u], 1 syllabic approximant [ɹ̩/ ɿ], 11 diphthongs [ia, ie, io, iu, ua, uo, ui, ai, oi, au, eu] and 4 triphthongs [iai, iui, iau, uai] in open syllables in Meixian Hakka.

	Monophthongs	Diphthongs	Triphthongs
Hashimoto (1973) phonetic transcription	[i] [ɛ] [ɑ] [ɔ] [u] [z̩]	[ĭɑ] [ĭɛ] [iu] [ĭɔ] [ui] [ʊ̆ɔ] [ʊ̆ɑ] [ʊ̆ɛ] [aĭ] [ɑʊ̆] [ɛʊ̆] [ɔĭ]	[ĭaĭ] [ĭɑʊ̆] [ʊ̆aĭ] [ĭɛʊ̆] [ĭɔĭ] [ʊ̆ɔĭ]
phonemic transcription	/i/ /e/ /a/ /o/ /u/	/ia/ /ie/ /iu/ /io/ /ui/ /uo/ /ua/ /ue/ /ai/ /au/ /eu/ /oi/	/iai/ /iau/ /uai/ /ieu/ /ioi/ /uoi/
Huang (1992; 1995)	/i/ /e/ /a/ /o/ /u/ /ɿ/	/ia/ /ie/ /iu/ /io/ /ui/ /uo/ /ua/ /ai/ /au/ /eu/ /oi/	/iui/ /iai/ /iau/ /uai/ /ieu/ (Huang, 1992)
Xie (1994)	/i/ /e/ /a/ /o/ /u/ /ɿ/	/ia/ /ie/ /iu/ /io/ /ui/ /uo/ /ua/ /ai/ /au/ /eu/ /oi/	/iui/ /iai/ /iau/ /uai/ /ieu/
Yuan et al. (2001)	/i/ /e/ /a/ /o/ /u/ /ɿ/	/ia/ /ie/ /iu/ /io/ /ui/ /uo/ /ua/ /ue/ /ai/ /au/ /eu/ /oi/	/iui/ /iai/ /iau/ /uai/
Wen (2006)	/i/ /e/ /a/ /o/ /u/ /ɿ/	/ia/ /ie/ /iu/ /io/ /ui/ /uo/ /ua/ /ai/ /au/ /eu/ /oi/	/iui/ /iai/ /iau/ /uai/

Table 4.1 The transcriptions of vowels, diphthongs and triphthongs in Meixian Hakka that occur in the open syllables in the past studies

The syllabic approximant [ɹ̩/ ʅ] occurs in many Chinese dialects. It was labelled as an apical alveolar vowel, which was represented by the non-IPA symbol [ɿ], in Beijing Mandarin by the Chinese linguists (Karlgren, 1915-1926; Chao, 1948; Howie, 1976). The palatograms and linguagrams of the so-called apical alveolar vowel displayed that [ɿ] in Beijing Mandarin is articulatorily a syllabic apico-laminal alveolar approximant (Zee, 2002). Lee and Zee (2003) further suggested that the sound should be transcribed as the IPA symbol [ɹ̩]. Similar to the case in Beijing Mandarin and other Chinese dialects, [ɿ] in Meixian Hakka occurs obligatorily after homorganic alveolar obstruent [s, ts, ts^h] in open syllables. In this study, the non-IPA symbol for the apical vowel [ɿ] is used for ease of presentation. In the study of Hashimoto (1973) on Meixian Hakka, the frictionless syllabic continuant [z̩] was used to represent either an apical vowel [ɿ] or a schwa [ə] described in other studies of Meixian Hakka. This indicates that Hashimoto perceived no audible difference between the apical vowel [ɿ] and schwa [ə] in Meixian Hakka. The schwa [ə] in Meixian Hakka, which occurs only in closed syllables and before syllable-final consonant [m, n , p, t], is not the focus of this study and it will be left for further investigation.

The patterns and frequency values of the first three formants as well as the temporal characteristics of the five vowels [i, e, a, o, u] and the apical vowel [ɿ] in Meixian Hakka have been investigated. The acoustic characteristics of vowels are presented in the first part of this chapter. Besides, the spectral characteristics and temporal organization of diphthongs are presented in the second part. It can be observed from the comparison of transcriptions in the past studies in Table 4.1 that there are four diphthongs starting with [i-], three or four diphthongs starting with [u-], two starting with [a-], one starting with [e-] and one starting with [o-]. Two types of transcription of the Meixian Hakka diphthongs were found in the study of Hashimoto (1973), namely the phonetic transcription and the phonemic transcription.

Allophones of the Meixian Hakka vowels and diphthongs were also described in the studies of Huang (1992; 1995) and Yuan et al. (2001). As stated in their studies, the phonemes of vowels may be realized as different allophones when they occur in different phonological environments. In consideration of the number of diphthongs, Hashimoto (1973) and Yuan et al. (2001) declared that there were 12 diphthongs in Meixian Hakka, whereas only 11 diphthongs were described in the studies of Huang (1992; 1995), Xie (1994) and Wen (2006). The 11 Meixian Hakka diphthongs that were mentioned in all of the past studies are [ia, ie, io, iu, ua, uo, ui, ai, oi, au, eu]. It can be noticed that diphthong [ue] appeared only in the studies of Hashimoto (1973) and Yuan et al. (2001). Though [ue] was listed in the study of Yuan et al. (2001), no further description and no test word were provided for the diphthong [ue]. In contrast, Hashimoto (1973) gave one example in his study for the diphthong [ue] and the monosyllabic word in which [ue] occurred was transcribed as [kue11]. According to the description of Hashimoto, diphthong [ue] was found in two monosyllabic words in Meixian Hakka. In addition to the word in which [ue] occurred after the voiceless unaspirated velar stop [k], it was claimed that [ue] also occurred in isolation. However, there was no Chinese character for the two words and the transcription indicated that the two words were imitative words.

Discrepancies are also found in the number and vowel quality of the Meixian Hakka triphthongs among the past studies. As exhibited in Table 4.1, it was generally agreed among the past studies that triphthongs /iui/, /iai/, /iau/ and /uai/ exist in the sound system of Meixian Hakka. Huang (1995), Wen (2006) and Yuan et al. (2001) declared that there were only four triphthongs in Meixian Hakka. However, Hashimoto

(1973), Huang (1992) and Xie (1994) indicated that there were five triphthongs instead. In addition to /iai/, /iui/, /iau/ and /uai/, they showed that /ieu/ was considered as one of the Meixian Hakka triphthongs (Huang, 1992; Xie, 1994). In contrast, Hashimoto (1973) claimed that there were phonologically 6 triphthongs in Meixian Hakka and they were /iai/, /uai/, /iau/, /ieu/, /uoi/ and /ioi/. Among the six triphthongs, /uoi/ and /ioi/ have not been described in other studies. In comparison, only two Meixian Hakka triphthongs have been mentioned in Chen's study (1993) and they were /iau/ and /uai/. With the assistance of phonetic data obtained in this study, such discrepancy in transcription among the past studies would be clarified and we would have a better understanding of the phonetic characteristics and temporal organization of vowels, diphthongs and triphthongs in Meixian Hakka.

4.2 Methodology

4.2.1 Speakers

The speech data were provided by 10 male and 10 female speakers. They were all native speakers of Meixian Hakka who were born and grew up in the district of Meijiang (梅江区). The district of Meijiang is located in the centre of the county of Meixian (梅县) and the county of Meixian is under the jurisdiction of Meizhou City (梅州市). All speakers were in the old age group and they were the elders aged between fifty and seventy years with no history of speech and hearing disorders.

4.2.2 Test Materials

In this study, 5 vowels [i, e, a, o, u], 11 diphthongs [ia, ai, io, oi, iu, ui, au, ua, ie, eu, uo], 4 triphthongs [iui, iai, iau, uai] and the apical vowel [ɿ] in Meixian Hakka described in the past studies were recorded for analysis. The test monosyllabic words in this study are listed in Table 4.2. All test vowels, diphthongs and triphthongs were in (C)V

Test vowel	Test word in (C)V syllable	Test diphthong	Test word in (C)V syllable	Test triphthong	Test word in (C)V syllable
[i]	[i44] 医 ‘cure’	[ia]	[ia44] 野 ‘wild’	[iui]	[iui53] 锐 ‘sharp’
[e]	[ke44] 鸡 ‘chicken’	[ai]	[ai44] 挨 ‘to be close to’	[iai]	[kiai53] 芥 ‘mustard’
[a]	[a44] 阿 a prefix for addressing people	[io]	[iɔ53] ‘potbellied’	[iau]	[iau44] 腰 ‘waist’
[o]	[o44] 屙 ‘to discharge excrement or urine’	[oi]	[ɔi44] 爱 ‘love’	[uai]	[kuai44] 乖 ‘well-behaved’
[u]	[p^hu44] 铺 ‘to spread’	[iu]	[iu44] 有 ‘have’		
[ɿ]	[sɿ44] 师 ‘teacher’	[ui]	[tui53] 对 ‘contrast’		
		[au]	[au53] 拗 ‘to bend or twist so as to break’		
		[ua]	[kua44] 瓜 ‘a creeping plant bearing large edible fruit’		
		[ie]	[ie44] ‘a kind of taros’		
		[eu]	[ɛu53] 沤 ‘to decay’		
		[uo]	[kuo44] ‘to curse’		

Table 4.2 The test monosyllabic words and their corresponding vowels, diphthongs as well as triphthongs in Meixian Hakka included in this study.

syllables which were associated with either a mid level tone or a high falling tone. It should be noted that the transcription of the test words in Table 4.2 was generally proposed by the dialectologists of the past studies. The transcription of vowels, diphthongs and triphthongs in Meixian Hakka that based on the acoustic data obtained in this study is to be discussed later.

Though there were 12 diphthongs described in the past studies (Hashimoto, 1973; Yuan et al., 2001), only 11 diphthongs have been investigated in this study and the diphthong [ue] was not included. The diphthong [ue] in (C)V syllables was not examined in this study due the absence of test word. As aforementioned, there was no listed test word for [ue] in the study of Yuan et al. (2001) and it made it difficult to find a test word that consists of [ue] in this research. Hashimoto (1973) showed that [ue] in (C)V syllables was found in two monosyllabic words in Meixian Hakka. Nonetheless, the two words were onomatopoeic words and there was no corresponding Chinese character for the words. In addition, the conventionality of the onomatopoeic words in Meixian Hakka that consists of [ue] may be questionable as it was not listed in other past studies. Onomatopoeia is a word that imitates the sound it is describing. The phonemic system of a particular language constrains the phonetic and phonological representations of the onomatopoeic words (Dofs, 2008; Gao, 2007). In other words, onomatopoeia should be composed of vowels and/or consonants that exist in the sound inventory of the language. Saussure declared that the number of onomatopoeia in English and French was limited and he concluded that “onomatopoeic words are never organic elements of a linguistic system” (1983: 69). It can be noticed that no word in (C)V syllable consists of [ue] in Meixian Hakka except the onomatopoeic words. Besides, diphthong [ue] in (C)VC syllables can be observed in the transcription of Meixian Hakka across studies (Hashimoto, 1973; Xie, 1994; Huang, 1995; Yuan et al., 2001). It showed that [ue] occurs in rimes, namely, /uet/ and /uen/ in Meixian Hakka. As it is believed that onomatopoeic words may not be appropriate to be used as test word, the diphthong [ue] in (C)V syllables described in the studies of Hashimoto (1973) and Yuan et al. (2001) were not included in this study and it will be left for further study that examines syllables with a syllable final consonant.

Table 4.1 in page 118 shows that there were in total 7 different Meixian Hakka triphthongs described by different authors of the past studies. The triphthongs were phonologically described as /iai/, /iui/, /iau/, /uai/, /ieu/, /uoi/ and /ioi/. In Meixian Hakka, /iai/ occurs either in isolation or in syllables with a syllable-initial consonant [k], whereas /uai/ occurs only in syllables with syllable-initial consonants [k] and [k^h]. The triphthong /iui/ can occur only in isolation. Huang (1995), Wen (2006) and Xie (1994) stated that there was only one possible morpheme in Meixian Hakka which is [iui53] ‘sharp’. In contrast, triphthong /iau/ in Meixian Hakka can occur in various phonological environments. On one hand, /iau/ occurs in isolation. On the other hand, it can be preceded by the majority of syllable-initial consonants, i.e., [p, p^h, m, t, t^h, n, l, ts, ts^h, s, k, k^h, ŋ, h], in Meixian Hakka.

Triphthong /ieu/ was described in the studies of Hashimoto (1973), Huang (1992) and Xie (1994), but not in the studies of Huang (1995), Wen (2006) and Yuan et al. (2001). Hashimoto (1973) has distinguished the triphthong /ieu/ from the diphthong /eu/, whereas in the studies of Huang (1992) and Xie (1994), /ieu/ was preceded by the syllable-initial consonants [k] and [k^h] only. Huang (1992) further indicated that triphthong /ieu/ and diphthong /eu/ were in complementary distribution. In contrast, there was no triphthong /ieu/ in the sound system of Meixian Hakka as shown in the studies of Huang (1995), Wen (2006) and Yuan et al. (2001). The rime of the test words which was transcribed as /ieu/ by the linguists of the earlier studies was considered as a diphthong /eu/ in the studies of Huang (1995), Wen (2006) and Yuan et al. (2001). It should also be noted that in the latter study of Huang (1995), the phonological distinction between triphthong /ieu/ and diphthong /eu/ has completely disappeared. The triphthong /ieu/ in the former study of Huang has been treated as the diphthong /eu/ in the latter study. In this study, the triphthong /ieu/ was not included in the word list. The acoustic data of a test word which was transcribed as [k^{h}ieu] in the earlier studies, but as [k^{h}eu] in the later studies are to be presented, showing that the existence of triphthong /ieu/ in Meixian Hakka is not supported by the phonetic evidence.

Triphthongs /uoi/ and /ioi/ were not included in the word list for recording. For the triphthong /uoi/, Hashimoto (1973) considered the first segment in /uoi/ as a vowel. However, the segment was treated as an approximant [ʋ] instead of vowel [u] in many of the past studies such as Hanyu Fangyin Zihui (1989), Huang (1992; 1995) and Yuan et al. (2001). On the contrary, Xie (1994) and Wen (2006) stated that the first segment of the triphthong should be transcribed as a voiced labio-dental fricative [v] instead. To differentiate between labiodentals, a comparison of the acoustic parameters such as median of the harmonics-to-noise ratio, centre of gravity and duration among different labiodentals is necessary (Gordon et al., 2002; Hamann & Sennema, 2005; Jassem, 1979). The controversy over the first segment of the rime might be resolved through further investigation and it would not be discussed in this study. Similar to the case of diphthong [ue], there was only one listed test word displayed in Hashimoto's study in which triphthong /ioi/ occurred. The transcription of the test word in Hashimoto's study was [kj'iɔɪ] which means 'to be tired' (1973: 68). In contrast, the transcription of the same word in the studies of Huang (1995) and Xie (1994) indicated that the rime of the syllable should be a diphthong [oi] instead. To examine the so-called triphthongs /uoi/ and /ioi/ in Meixian Hakka, a supplementary study has been conducted. Video and audio recordings for the triphthong /uoi/ as well as audio recordings for the triphthong /ioi/ were obtained and analyzed. The data are to be presented and discussed in next section of the chapter.

Digital audio recordings were made of the native speakers of Meixian Hakka reading a randomized list of 21 meaningful monosyllabic words as listed in Table 4.2. Each test word was embedded in a carrier sentence, [ŋai11 t^huk5 ____ pun35 m̩11 t^haŋ33] "I read ____ for you (to) listen.". The order of the test monosyllabic words was randomized and four readings of the word list were recorded. The speakers were instructed to read the word list at a normal rate of speech and audio recordings were conducted in a quiet room.

4.2.3 Data Analysis

Praat 4.3.19 speech analysis software was used for both spectral analysis and temporal measurement of vowels, diphthongs and triphthongs in Meixian Hakka. Speech data were captured at a sampling rate of 10,000 samples per second which produce an upper frequency cut-off of 5,000 Hz. Formant trajectories were obtained for determining the steady states of vowels, the two elements of diphthongs and the three elements of triphthongs. The frequencies of the first three formants, i.e., F_1, F_2 and F_3, were measured for each vowel at its steady state, close to the midpoint of the vowel if possible; each diphthong at the points of nearest approach to the targets of the onset steady state and offset steady state; and each triphthong at the points of nearest approach to the targets of the first, second and third elements of the sequence. The onset and offset of the transition for either a diphthong or a triphthong were mainly determined by the F_2 trajectory. It is because with respect to the formant movement of the F_2 trajectory, a diphthong can be differentiated from other diphthongs in a specific language. However, it is not the case for the F_1 and F_3 trajectories of which the formant movement may look alike across different diphthongs. In this study, the formant frequencies of each element of diphthongs and triphthongs are compared with the formant frequencies of the ideal vowel targets. The vowel target is defined by the formant frequency values obtained from the midpoint of a monophthong at which the formant trajectories are generally steady and the formant pattern is acoustically prominent.

In addition, the duration of different portions of each vowel sound have been measured with the aid of the formant trajectories, visual inspection of the waveforms, and auditory judgment of the investigator. The duration data of the vowel sounds that have been obtained in this study are as follows: (1) total duration of each monophthong, diphthong and triphthong, (2) duration of quasi-steady states for each element of the each diphthong and triphthong, and (3) transition duration between every two elements of the sound sequence. The temporal organization of the monophthongs, diphthongs and triphthongs in Meixian Hakka will be presented and discussed at the end of each subsection of the next section.

4.3 Results

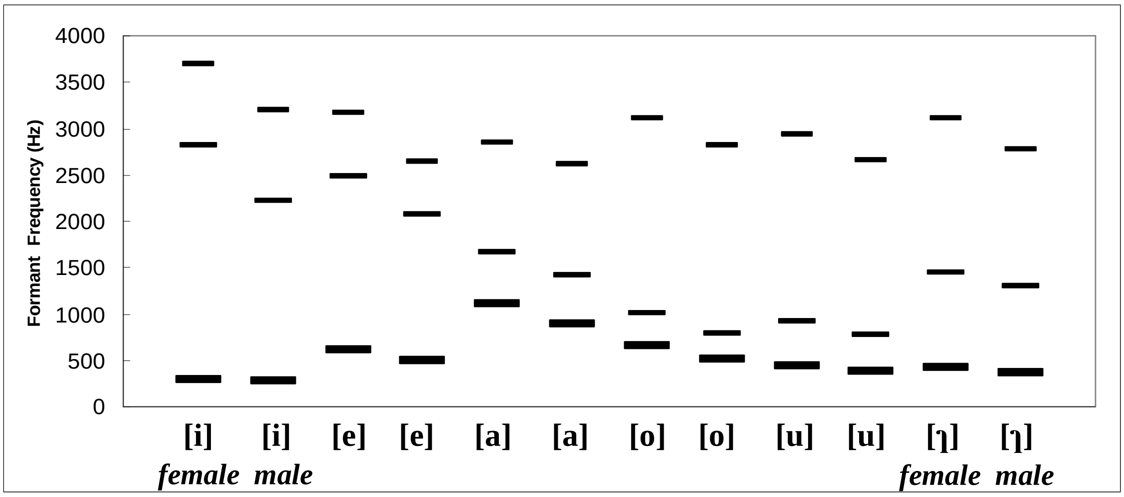
This section presents the results for each of the monophthongs, diphthongs and triphthongs in Meixian Hakka that listed in Table 4.2. The results are divided into three subsections, which are monophthongs, diphthongs and triphthongs. According to the transcriptions in the past studies (Hashimoto, 1973; Huang, 1992; 1995; Wen, 2006; Xie, 1994; Yuan et al., 2001), the rimes of the test words have been grouped into different subsections. The 5 monophthongs [i, e, a, o, u] and the apical vowel [ɿ] in Meixian Hakka are presented in the subsection of vowels; The 11 diphthongs [ia, ai, io, oi, iu, ui, au, ua, ie, eu, uo] in the subsection of diphthongs; The 4 triphthongs [iui, iai, iau, uai] as well as the 2 triphthongs [ioi, uoi] proposed by Hashimoto (1973) in the subsection of triphthongs. The formant pattern and temporal organization of each of the monophthongs, diphthongs and triphthongs will be described and discussed in each subsection. In addition, the vowels, diphthongs and triphthongs in Meixian Hakka are transcribed based on the spectral and temporal data obtained in this study and the transcription of vowels in this study will be compared with the traditional transcriptions proposed by other linguists. At the end of this section, the transcription of the vowel system in Meixian Hakka which is evidenced by the acoustic data in this study is presented.

In consideration of diphthong, there have had long debates about its definition in the past and so far no general consensus on this matter can be reached. Generally speaking, there are two views of a diphthong that have been proposed by the linguists. On one hand, some linguists defined a diphthong as either a sequence of two vowels or one vowel and one glide (Arlund, 2006; Bond, 1978; Catford, 1977; Gottfried, 1993, Jha, 1985; Lindau et al., 1990). A diphthong is classified as a sound segment that composed of quasi-steady-state vowel onset, quasi-steady-state vowel offset and a transition with an apparent rate of change in between. Lehiste and Peterson (1961) investigated different syllable nuclei in American English and /ɑɪ/, /ɑʊ/ and /ɔɪ/ were phonetically classified as diphthongs based on the slope of transition and the duration of the target position. It has been proposed that a diphthong was characterized by two target positions. In other words, a diphthong was defined as a syllable nucleus consisting of two vowel targets with definite steady state and a prominent transition between the two targets (Gay, 1968; Lehiste & Peterson, 1961). In the study of Lehiste and Peterson, /ɑɪ/, /ɑʊ/ and /ɔɪ/ were further distinguished from /e$^{\text{ɪ}}$/, /o$^{\text{ʊ}}$/ and /ɝ/ which have only one vowel target for each syllable nucleus. On the contrary, Trager and Smith (1957) declared that /ɑɪ, ɑʊ, ɔɪ, eɪ, oʊ/ in American English were sequences of a vowel plus a semivowel. Though it seems that the definition of a diphthong in the past studies was mainly based on the data of American English, such definition has been widely adopted by many scholars in the researches of the world languages. The definition of a diphthong which defined a diphthong as a sequence of two perceptually different vowels within one syllable nucleus is generally recognized as the duality (sequence) view of diphthongs. On the other hand, there are scholars who have considered a diphthong as a phonemic unit with continuously changing formant trajectories within a syllable (Kent & Read, 2002; Ladefoged, 2001). It was proposed that a diphthong involves a change in vowel quality from one vowel to another. This view is named the unity (single vowel) view of diphthongs.

The analysis of this paper has been undertaken on the basis of the duality view and a diphthong is divided into three portions, namely, the vowel onset, the vowel offset and the transition in between. Unlike the diphthongs in American English, the data of this study displays that many diphthongs in Meixian Hakka have a quasi-steady-state onset element, a quasi-steady-state offset element and a transition between the two elements. In that case, it may be more rational to consider the diphthongs in Meixian Hakka as syllable nuclei containing two target positions and the formants of each diphthong shift from the onset target position to the offset target position. The data points of the formants were taken from each of the two quasi-steady-state elements in a diphthong. In addition to the first three formant frequencies of the onset and offset elements of a diphthong, the duration of the two target elements, the duration of the transition between the two elements, as well as the rate of change of the F_2 transition have been measured and calculated. In comparison to diphthongs, the number of triphthongs that occur in the world languages is much fewer and the definition of a triphthong has not been discussed. In this study, a triphthong is defined as a sequence of three vocalic elements which are perceptually different. It is supposed that a triphthong should involve a relatively quick change in vowel quality from one quasi-steady-state vowel onset to one quasi-steady-state vowel offset by passing over a target vowel in the middle that may or may not have the steady-state portion.

4.3.1 *Vowels*

Figure 4.1 displays the formant patterns of F_1, F_2 and F_3 for the 5 monophthongs [i, e, a, o, u] and the apical vowel [ɿ] in the (C)V syllables in Meixian Hakka. Each IPA symbol on the left indicates the first three formants of each vowel produced by 10 female speakers, whereas that on the right the first three formants of each vowel produced by 10 male speakers. Each mean value was obtained from 40 data points for a particular vowel (4 repetitions x 10 speakers). Table 4.3 shows the mean frequency values of F_1, F_2 and F_3 and their standard deviation for the monophthongs and apical vowel that



Monophthongal Vowels in Meixian Hakka

Figure 4.1 Mean values of the first three formants (F_1, F_2, F_3) for the monophthongal vowels [i, e, a, o, u] in Meixian Hakka (left for female speakers and right for male speakers).

Test vowels	Female speakers			Male speakers		
	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)
[i]	298 (44)	2812 (133)	3686 (222)	270 (26)	2219 (234)	3194 (199)
[e]	617 (63)	2481 (125)	3161 (231)	501 (41)	2076 (174)	2644 (157)
[a]	1104 (115)	1662 (76)	2839 (306)	888 (64)	1414 (77)	2616 (283)
[o]	654 (51)	1002 (88)	3110 (222)	508 (46)	789 (61)	2817 (287)
[u]	442 (48)	916 (78)	2935 (156)	374 (28)	776 (66)	2653 (228)
[ɿ]	430 (55)	1438 (129)	3116 (194)	369 (37)	1293 (122)	2767 (226)

Table 4.3 Mean frequency values of F_1, F_2 and F_3 (in Hz) and their standard deviations (s.d.) for the monophthongs [i, e, a, o, u] and the apical vowel [ɿ] in Meixian Hakka that produced by 10 female and 10 male speakers.

produced by 10 male and 10 female speakers. The spectral data of the vowels for each individual speaker are shown in Appendix C.

As can be seen, for both male and female speakers, the high front vowel [i] has a lower F_1 and a higher F_2 in comparison to the close-mid front vowel [e] (Figure 4.1). The frequency values of F_1, F_2 and F_3 for [i] are 270 Hz, 2219 Hz and 3194 Hz, respectively, for the male speakers, whereas they are 298 Hz, 2812 Hz and 3686 Hz, respectively, for the female speakers. In contrast, the frequency values of F_1, F_2 and F_3 for [e] are 501 Hz, 2076 Hz and 2644 Hz, respectively, for the male speakers, whereas they are 617 Hz, 2481 Hz and 3161 Hz, respectively, for the female speakers. It can be noticed that for the male speakers, the mean frequency values of F_2 for [i] and [e] are similar and the difference in F_2 value between the two vowels is relatively small, which is less than 150 Hz. However, it should be noted that the result of the formant pattern that [i] has a lower F_1 and a higher F_2 than [e] is consistent for every single speaker in this study.

Likewise, the mean frequency values of F_2 for [o] (789 Hz) and [u] (776 Hz) are similar for the male speakers. Compare the formant pattern of F_1, F_2 and F_3 for [o] with that for [u], results show that [o] has higher mean values of the first three formants than [u]. As shown in Table 4.3, the mean frequency values of F_1, F_2 and F_3 for [o] are 508 Hz, 789 Hz and 2817 Hz, respectively, for the male speakers, whereas the values are 654 Hz, 1002 Hz and 3110 Hz, respectively for the female speakers. On the other hand, the mean values of F_1, F_2 and F_3 for [u] are 374 Hz, 776 Hz and 2653 Hz, respectively, for the male speakers, whereas they are 442 Hz, 916 Hz and 2935 Hz, respectively, for the female speakers. Differ from the case of [i] and [e], the spectral difference between [o] and [u] varies across speakers. Though the mean frequency values of the first three formants for [o] are higher than those for [u], results show that the difference in F_2 and

F_3 values between the two vowels is not consistent for all speakers. For some speakers, the F_2 value for [o] can be similar to or lower than that for [u] and such result can be found in the data of four male and two female speakers in this study. Similarly, the data of two male and two female speakers show that the F_3 for [o] is not necessary to be higher than the F_3 for [u]. In contrast, the F_1 value for [o] is consistently higher than that for [u] for every speaker and the difference in F_1 value between [o] and [u] is higher for the female speakers than that for the male speakers.

The F_1 of [a] (male: F_1 = 888 Hz, female: F_1 = 1104 Hz) has the highest frequency value among all vowels in Meixian Hakka and the F_2 value of [a] (male: F_2 = 1414 Hz, female: F_2 = 1662 Hz) is at the middle range of the frequency value. It displays that the formant pattern of F_1, F_2 and F_3 for the vowel [a] can be easily distinguished from other monophthongs in Meixian Hakka. In addition to the phonological constraints of the apical vowel that [ɿ] occurs only after alveolar affricates [ts, tsʰ] and alveolar fricative [s] in Meixian Hakka, the F_1F_2 pattern of [ɿ] as shown in Figure 4.1 differentiates [ɿ] from other vowels as well. The acoustic data indicates that low F_1 (male: F_1 = 369 Hz, female: F_1 = 430 Hz) and F_2 in middle frequency range (male: F_2 = 1293 Hz, female: F_2 = 1438 Hz) characterize the apical vowel [ɿ]. The mean frequencies of the first and second formants for [ɿ] in Meixian Hakka corresponds to a high central vowel and results show that the apical vowel has an F_1 value which is similar to that of [u].

As can be seen in Figure 4.1 and Table 4.3, the formant patterns of [i], [e], [a], [o] and [u] as well as the apical vowel [ɿ] for the male speakers are similar to the formant patterns of the monophthongs for the female speakers. The overall formant frequency values, particularly the F_1 and F_2 values, of the vowels for the male speakers are lower than those for the female speakers as expected. It is due to the differences in vocal tract

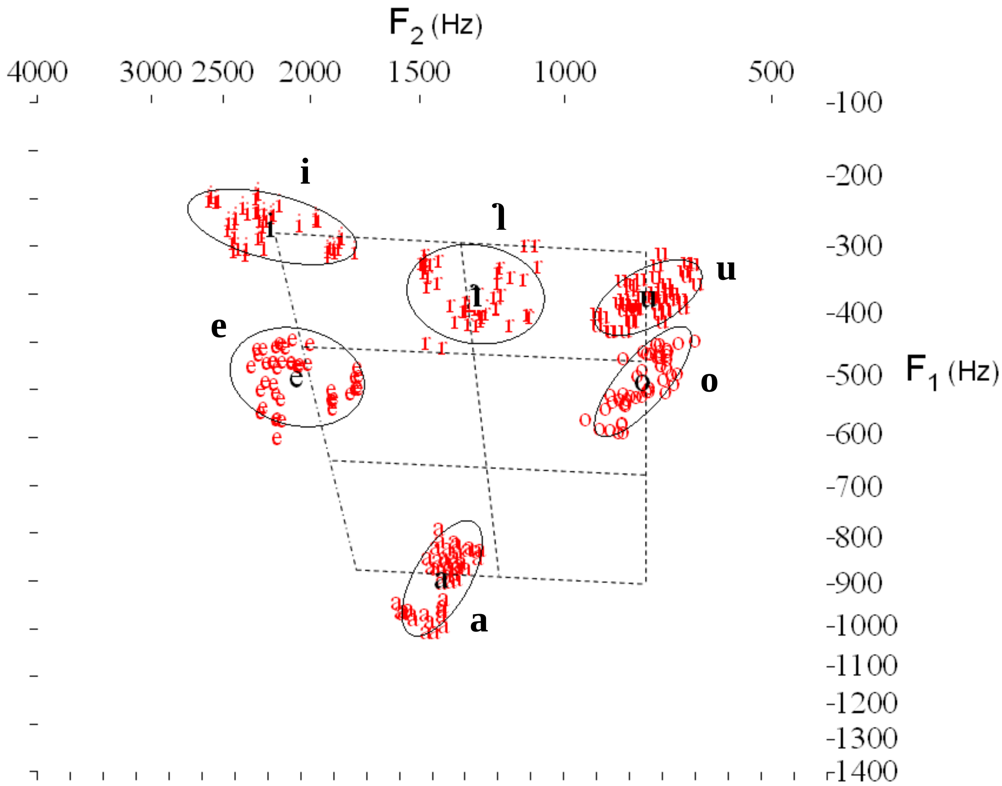


Figure 4.2 Vowel ellipses for [i, e, a, o, u, ɿ] in Meixian Hakka on the F_1/F_2 plane for 10 male speakers. Vowel quadrilateral is in dotted line for reference.

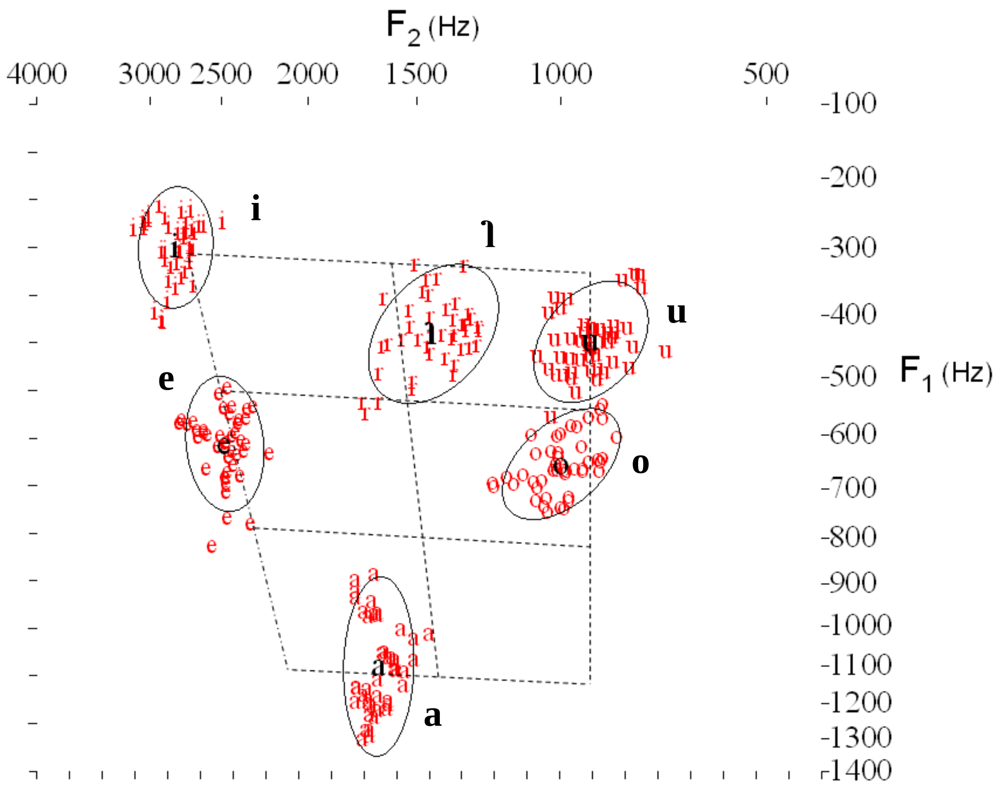


Figure 4.3 Vowel ellipses for [i, e, a, o, u, ɿ] in Meixian Hakka on the F_1/F_2 plane for 10 female speakers. Vowel quadrilateral is in dotted line for reference.

dimensions between male and female speakers (Peterson & Barney, 1952). That is, the vocal tract dimension is smaller for a female speaker than that for a male speaker in general and it in turn results in higher formant frequencies values for the female speakers as shorter vocal tract resonates at higher frequencies.

Figure 4.2 shows the positions of the vowel ellipses for the monophthongal vowels [i, e, a, o, u] and the apical vowels [ʅ] in Meixian Hakka in the acoustical vowel chart (F1/F2 plane) on Bark scale, with F_1 on the y-axis and F_2 on the x-axis. Vowel ellipses are drawn on the basis of the distribution of 40 data points produced by 10 male speakers and the data points are scattered on the graph. Figure 4.3 shows the vowels ellipses for the monophthongs and apical vowels that produced by 10 female speakers and Figure 4.4 shows a comparison of the vowel ellipses between male and female speakers. Each vowel ellipse shows an area containing the data points of a vowel category with radii of two standard deviations along the mean frequency value. According to the normal statistical assumptions, it is predicted that approximately 95% of the data points from the sample of the speakers are included within the ellipse. The acoustical vowel chart has been converted to the Bark scale for better approximation of the perceived distance within the phonetic space.

As shown in Figures 4.2 and 4.3, the ellipse for each vowel in Meixian Hakka occupies distinct location in the F1/F2 plane and there is no overlap in between. It shows that Meixian Hakka vowels differentiate from one another clearly in the acoustical vowel space. A vowel quadrilateral has been drawn in Figures 4.2 and 4.3, respectively. Results show that there are three levels of vowel height, namely high, mid and low, in the vowel system of Meixian Hakka. The three-level vowel height distinctions may be due to the fact that the vowel system of Meixian Hakka is relatively simple. It is supposed that three levels of vowel height rather than four levels as traditionally shown in the IPA vowel chart are enough for the listeners to distinguish between vowels. As can be seen, [i] and [u] are at the highest level of the vowel height and it should be reasonable to classify them as close or high vowels. It also shows that, for both male and female speakers, [u] in Meixian Hakka is not as high as [i] on the F1/F2 plane. On the contrary, [a] is at the lowest level of the vowel height and it is the only open or low vowel in Meixian Hakka.

It can be noticed that there is no phonological distinction between high-mid vowel and low-mid vowel in Meixian Hakka. The vowel ellipses for the Meixian Hakka mid vowels show that the data points distribute mainly over an area in the acoustic space between the high-mid level and low-mid level of the vowel height. That is, for the mid front vowel, the vowel is in between the high-mid front vowel (Cardinal Vowel 2), IPA [e], and the low-mid front vowel (Cardinal Vowel 3), IPA [ɛ], while for the mid back vowel, the vowel is in between the high-mid back vowel (Cardinal Vowel 7), IPA [o], and the low-mid back vowel (Cardinal Vowel 6), IPA [ɔ]. The positions of the mid vowels indicate that the two vowels can be interpreted as either [e, o] or [ɛ, ɔ]. As the acoustic distant between mid vowel and high vowel is relatively smaller than that between mid vowel and low vowel, it is more rational to interpret the mid front vowel as [e] and the mid back vowel as [o]. The acoustic results of this study also support the transcription of the past studies which treated the Meixian Hakka mid vowels as [e] and [o] (Huang, 1995; Wen, 2006; Yuan et al., 2001). Huang (1995) stated that [e] and [o] in Meixian Hakka were realized as high-mid vowel when it occurred in the open syllables (p. 5). That is, /e/ was realized as [e] when it occurred in rimes such as [e, ie, eu], whereas /o/ was realized as [o] when it occurred in rimes such as [o, io, uo, oi].

A comparison of the relative positions of the vowel ellipses in Figure 4.4 indicates

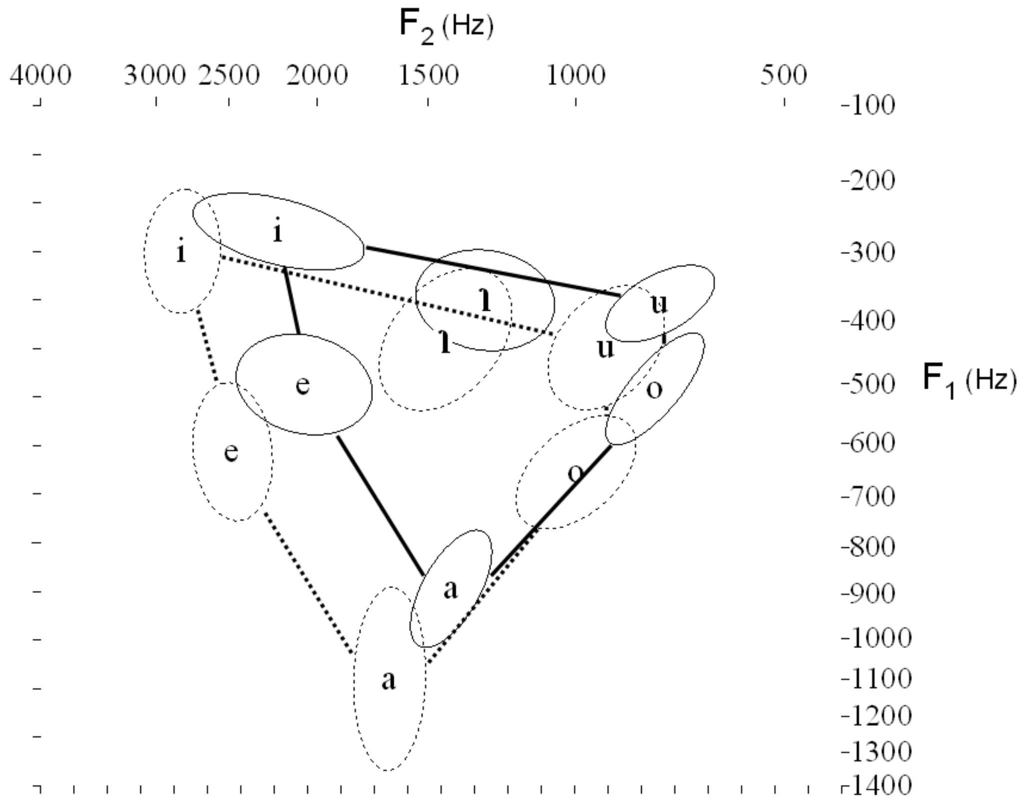


Figure 4.4 Vowel ellipses for [i, e, a, o, u] in Meixian Hakka on the F_1/F_2 plane. Solid ellipses indicate the acoustic data for the male speakers and dotted ellipses indicate those for the female speakers.

that the vowel ellipses for the female speakers are located more to the left and lower area of the acoustical vowel space than those for the male speakers. It is resulted from higher formant frequency values of the resonant sounds produced by the female speakers as mentioned. In addition, a comparison of Figures 4.2 and 4.3 shows that the vowels produced by the female speakers occupy a larger acoustical vowel space than those produced by the male speakers do. This finding can also be noticed in Figure 4.4. As can be seen, in terms of vowel backness in the acoustic space, the distance of vowel ellipses between [i] and [u] for the female speakers is relatively larger than that for the male speakers. Similar case is found for the distance between [e] and [o]. In terms of vowel height, it shows that the distance of vowel ellipses between [ɿ] and [a] for the female speakers is larger. Results also show that the relative distance between the mid vowels [e, o] and the high vowels [i, u] is greater for female speakers than that for the

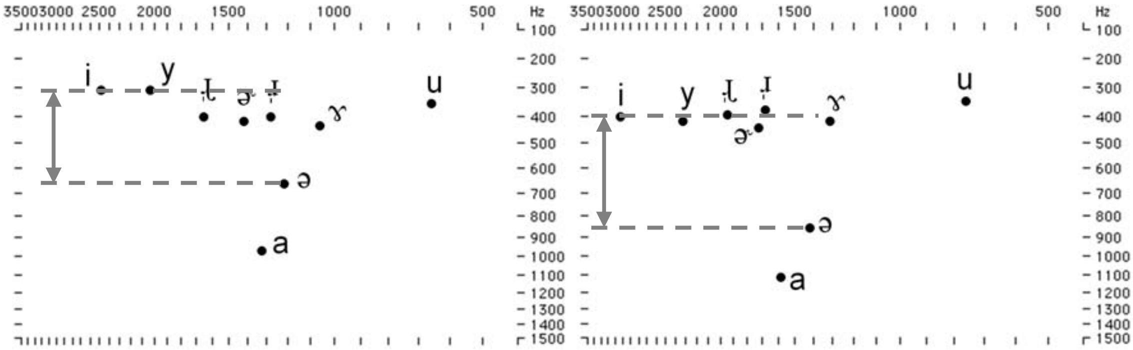


Figure 4.5a The vowel ellipses for the Beijing Mandarin vowels which produced by 10 male (the left diagram) and 10 female (the right diagram) speakers. The vowel chart is reproduced based on the data from Zee and Lee, 2001.

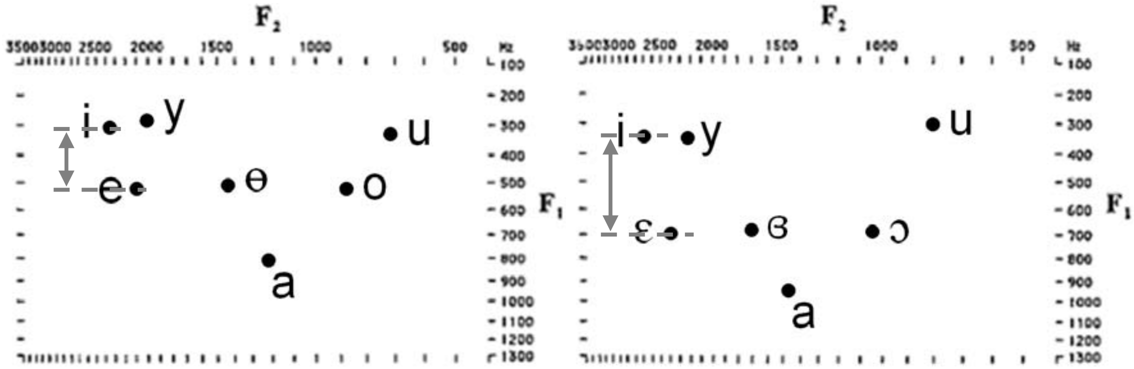


Figure 4.5b The vowel ellipses for the Hong Kong Cantonese monophthongs (indicated by empty circles) which produced by 10 male (the left diagram) and 10 female (the right diagram) speakers. The vowel charts are reproduced based on the data from Zee, 2000.

male speakers. The acoustic difference between male and female speakers was also found in the vowel systems of Beijing Mandarin and Hong Kong Cantonese as presented in the past studies (Zee, 2000; 2001; Zee & Lee, 2001) that based on the data of 20 native speakers. The vowel charts of the monophthongs in Beijing Mandarin and Hong Kong Cantonese from Zee (2000; 2001) and Zee and Lee (2001) has been reproduced and are shown in Figures 4.5a and 4.5b, respectively. As Hakka, Mandarin and Cantonese are major dialects of the Chinese language, it suggests that this phenomenon may occur in all Chinese dialects. However, it is no doubt that further investigation on other dialect groups of the Chinese language is necessary to verify the hypothesis.

In addition, it can be observed that in Meixian Hakka, the distance between high and mid vowel is larger for the front vowels than that for the back vowels. This finding holds true for both male and female speakers. In contrast, the difference in relative distance between high and mid vowel with regard to vowel backness does not exist in the acoustic data of other dialects such as Beijing Mandarin and Hong Kong Cantonese. Moreover, for both [i] and [u] in Beijing Mandarin and Hong Kong Cantonese, their values of F_1 are similar, at about 300 Hz, whereas the acoustic data of this study shows that [u] in Meixian Hakka has a higher F_1 value than [i]. As shown in Table 4.3 in page 130, the value of F_1 for [u] is about 400 Hz (male: F_1 = 374 Hz; female: F_1 = 442 Hz), whereas that for [i] is only about 300 Hz (male: F_1 = 270 Hz; female: F_1 = 298 Hz). As shown in Figure 4.4, the position of the vowel ellipse for [u] in terms of vowel height, i.e., F_1 value, is lower than that for vowel [i].

In addition to the formant patterns, the temporal organization of the monophthongs has been investigated. The mean durations of 5 monophthongs [i, e, a, o, u] and the apical vowels [ɿ] in Meixian Hakka for the male and female speakers are displayed in Figure 4.6. As can be seen, [i] (male: 479.6 ms, female: 499.5 ms) has the longest duration among the vowels, while the apical vowel [ɿ] (male: 365.5 ms, female: 414.1 ms) has the shortest duration for both male and female speakers. For the other four vowels [e, a, o, u], there is some discrepancy in the temporal organization between male and female speakers. For the male speakers, the durations of [o] (422.1 ms) and [a] (420.7 ms) are similar and they are the second longest. Following [o] and [a] are [e] (408.1 ms) and [u] (373.4 ms), ranging from longer to shorter vowel duration. The duration of [u] for the male speakers is similar to that of the apical vowel [ɿ] and the two vowels are shortest in comparison to other vowels. In contrast, for the female speakers, [o] (485.4 ms) is the second longest and its duration is similar to [i]. It is followed by [e]

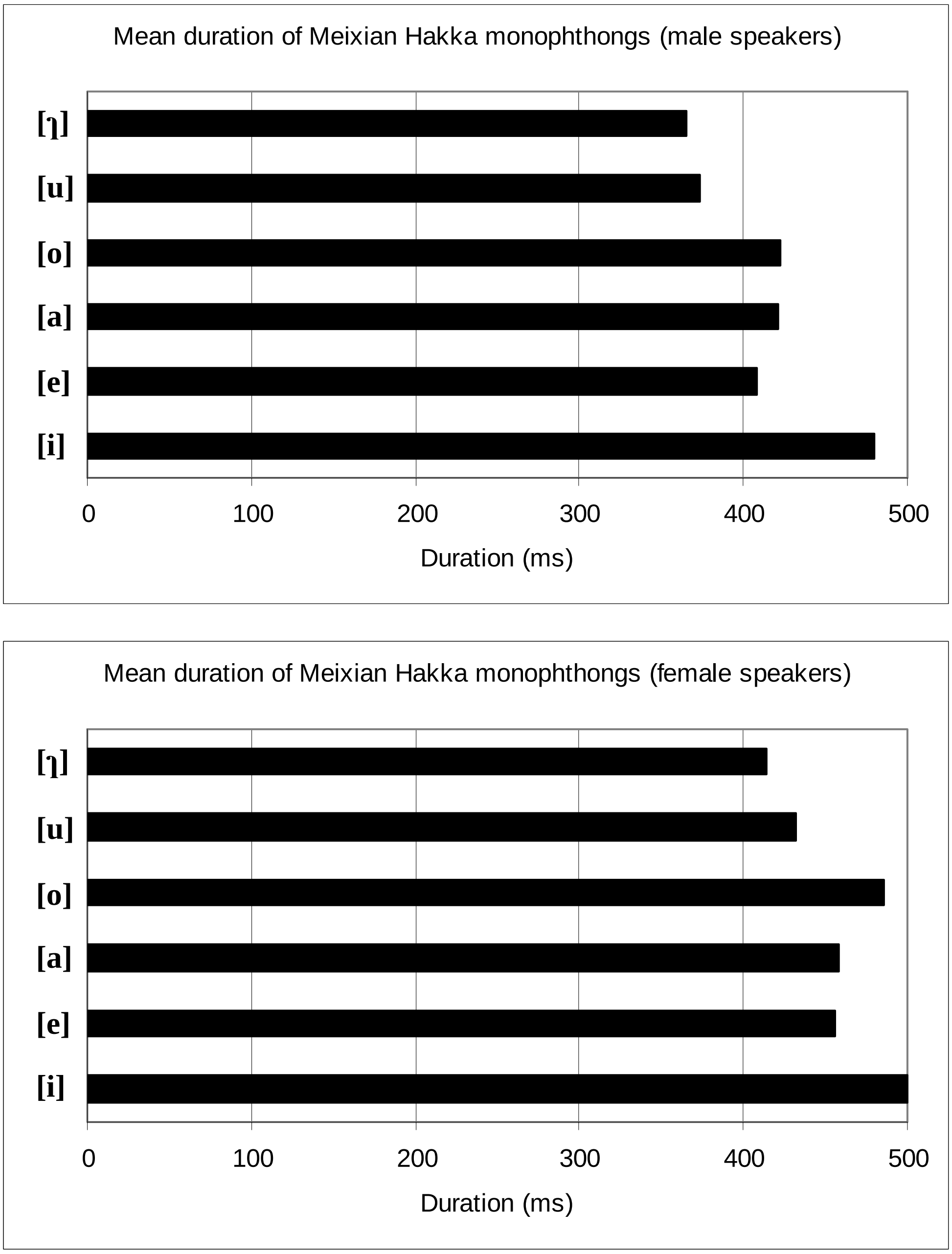


Figure 4.6 Mean duration of 5 monophthongs [i, e, a, o, u] and the apical vowel [ʅ] in (C)V syllables in Meixian Hakka for the male speakers (upper chart) and female speakers (lower chart)

(455.6 ms) and [a] (458.6 ms) which have similar duration. Similar to the data of male speakers, [u] (431.4 ms) is the second shortest among the vowels for the female speakers.

However, if we take a closer look at the mean duration of the Meixian Hakka monophthongs and the apical vowel for each individual speaker, it shows that duration

Figure 4.6a Mean duration of Meixian Hakka monophthongs [i, e, a, o, u] and apical vowel [ɿ] that produced by 10 male speakers

Figure 4.6b Mean duration of Meixian Hakka monophthongs [i, e, a, o, u] and apical vowel [ɿ] that produced by 10 female speakers

may not be an effective cue for differentiating the vowels in Meixian Hakka. The mean duration of the Meixian Hakka monophthongs for every individual speaker are displayed in Figures 4.6a and 4.6b. It shows that even though vowel [i] has the greatest mean vowel duration as displayed in Figure 4.6, it does not hold true for every speaker in this study. The data of male speakers 2 and 6 denote that vowel [i] does not have the greatest vowel duration. Furthermore, the duration of [i] for four female speakers (female speakers 4, 5, 6 and 7) also indicates that [i] may not be the longest vowel in Meixian Hakka. Although the duration of [i] might not be the greatest for all speakers, [i] is longer than [u] and [ɿ] in most cases. As mentioned, the mean duration of [ɿ] is the shortest for male and female speakers as shown in Figure 4.6. However, the data of two male (male speakers 3 and 8) and more than half of the female speakers (female speakers 1, 2, 4, 6, 8 and 9) displays that the mean duration of [ɿ] may be greater than [u] or [a]. Generally speaking, the duration of [ɿ] is relatively shorter than other vowels for most Meixian Hakka speakers. The mean duration of [e, a, o] varies across speakers and they are generally neither the longest vowel nor the shortest vowel in Meixian Hakka. It may be concluded that there is no consistent pattern of the intrinsic duration for different vowels in Meixian Hakka. This may be explained by the fact that formant frequency pattern should be sufficient for the identification of the Meixian Hakka vowels, and vowel duration in contrast may vary due to different speaking contexts and different speaking styles. However, despite of the apical vowel [ɿ], there is a tendency for [i] having the greatest duration and [u] having the shortest duration in Meixian Hakka.

It has long been claimed that the duration of vowels can be influenced by various factors such as tense-lax distinction of vowel, vowel height, voicing and place of articulation of the consonantal context of vowels, syllable stress and speaking rate

(Erickson, 2000; House & Fairbanks, 1953; Klatt, 1976; Lehiste & Peterson, 1961; Peterson & Lehiste, 1960). Among these factors, tense-lax distinction of vowel and vowel height contribute the inherent attributes of vowel duration. Lehiste and Peterson (1961) showed the dynamic characteristics of various syllable nuclei in American English. They found that tense vowel in CVC syllables had a relatively greater intrinsic duration and a relatively short offglide, whereas lax vowel had a relatively shorter intrinsic duration and a relatively long offglide. Early studies on the perceptual influence of duration reported that intrinsic duration was an important cue for the identification of vowels (Ainsworth, 1972; Bennett, 1968). Idson and Massaro (1980) provided evidence that perceived duration is crucial to the tense-lax distinction of vowels. The perceptual results indicated that the judgment of vowels changed gradually from lax to tense with gradual increases in the target vowel duration, intervowel interval and masking vowel duration. In terms of the influence of vowel height on vowel intrinsic duration, Peterson & Lehiste (1960) displayed that open vowels such as /ɑ/ and /æ/ had greater intrinsic durations than close vowels such as /i/. The correlation between vowel intrinsic duration and vowel height was also found in study of Yadav (1979) which has investigated the vowel duration in Maithili. Results showed that the average duration of Maithili vowels with a more open jaw position such as /a/ are longer than vowels with a closer jaw position such as /i/ and /u/. However, it seems that the difference in intrinsic vowel duration that found in American English and Maithili was not supported by the acoustic data of this study. In this study, results show that [i] is the longest among all vowels in Meixian Hakka, but there seems to be no correlation at all between intrinsic duration and vowel height.

4.3.2 *Diphthongs*

In this section, the acoustical data of the 11 diphthongs in Meixian Hakka, which have been transcribed as [ai, ia, oi, io, au, ua, iu, ui, ie, uo, eu] are presented. As mentioned in Chapter 1, the duality view of diphthongs (Arlund, 2006; Bond, 1978; Catford, 1977; Gottfried, 1993, Jha, 1985; Lindau et al., 1990) has been taken for the analysis. The duality view has considered a diphthong as a combination of two perceptually distinct vowels and phonetically, a diphthong was interpreted as syllable nucleus having two target positions. According to the duality view, a diphthong should be separated into three parts which are (1) an onset target, (2) an offset target, and (3) a transition between the two targets with an apparent rate of change. Past studies of diphthongs showed that formant frequencies of the onset and/or offset of a diphthong, rate of change of F_2 transition and temporal organization of a diphthong might be optimal to characterize diphthongs in a specific language (Chan & Ren, 1988; Collier & t'Hart, 1983; Gay, 1968; 1970; Holbrook & Fairbanks, 1962; Jha, 1985). Manrique (1979) suggested that Spanish diphthongs might be described and identified in terms of the onset or offset steady states of the open vowels, rate of change of F_2 transition and temporal relationship between the onset and offset steady states. The relevance of onset steady state and second formant rate of change on the perception of English diphthongs has also been verified in the studies of Gay (1968; 1970).

In this study, formant frequencies of the two elements, rate of change of F_2 transition, and temporal organization of diphthongs in Meixian Hakka have been investigated. The formant frequencies of the first two formants, i.e., F_1 and F_2, for the two elements of a diphthong, which are the onset target and offset target, are plotted on the graphs. The formant patterns of Meixian Hakka diphthongs are presented and each of them is compared with other diphthongs that are similar in terms of their target positions. The F_1 and F_2 values of the two elements of a diphthong are also compared with the formant frequencies of their vowel targets in monophthongal vowels. It should be noted that the transcription of diphthongs in this chapter is based on the phonological transcriptions used in the past literatures (Huang, 1995; Yuan et al., 2001). Based on the acoustic data obtained in the study, the phonetic transcription of each Meixian Hakka diphthong will further be discussed. Among various acoustic characteristics of diphthongs, the dynamic properties of a diphthong have been widely recognized as the potent cues for differentiating diphthongs in a language. In this study, the calculation of the F_2 range and F_2 range of change for each Meixian Hakka diphthong has been made. The data have been used to examine (1) the relevance of F_2 rate of change on the distinction among Meixian Hakka diphthongs in terms of vowel height and backness and (2) the correlation between F_2 rate of change and F_2 frequency range of the transition between the two elements. The dynamic properties of Meixian Hakka dipthhongs in this study are also compared with the findings of the past studies which investigated diphthongs in other languages and Chinese dialects. Such comparison may contribute to a better understanding of the acoustic properties of diphthongs that may be language-specific or language universal.

The mean formant frequencies of each element of the 11 diphthongs in Meixian Hakka are shown on the F_1/F_2 plane for 10 male speakers in Figures 4.7 and 4.8, and for 10 female speakers in Figures 4.9 and 4.10. The formants of the first and second elements of Meixian Hakka diphthongs are plotted against the formants of the corresponding monophthongs. Each of the two end points of an arrow shows the first and second elements of a diphthong. The direction of an arrow indicates the movement of the formants on the F1/F2 plane. The dotted circles in the figures indicate the positions of the monophthongs in the acoustical vowel space.

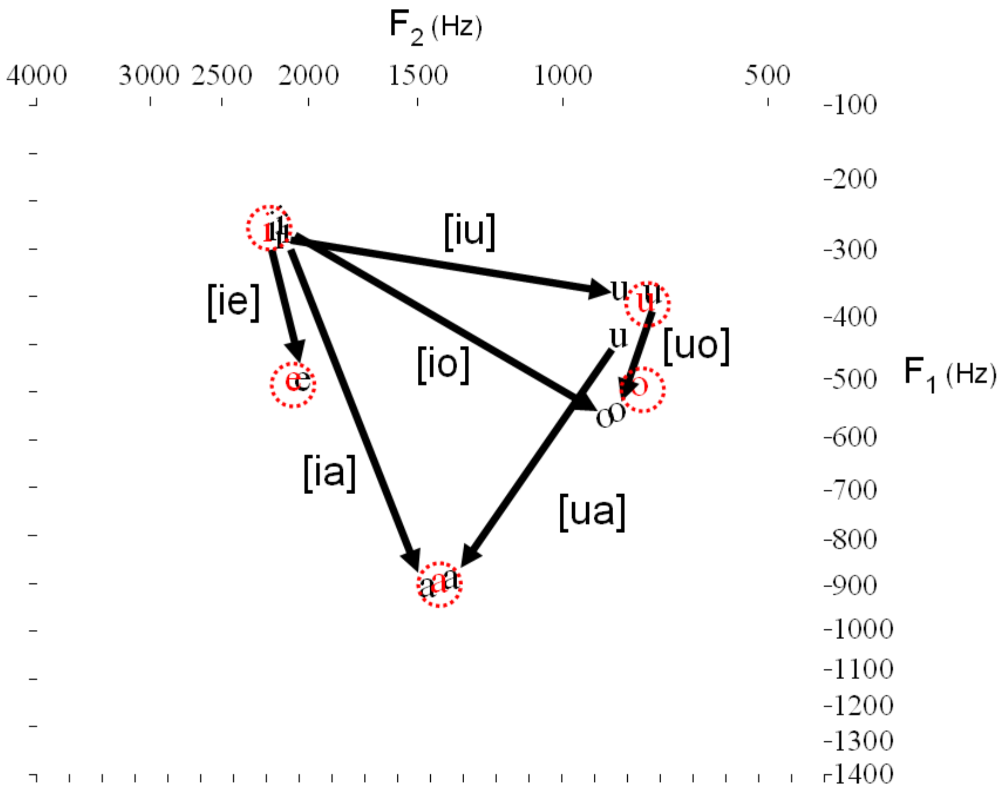


Figure 4.7 Formant movements for the diphthongs [ie, ia, io, iu, ua, uo] and the 5 monophthongal vowels [i, e, a, o, u] in Meixian Hakka for 10 male speakers.

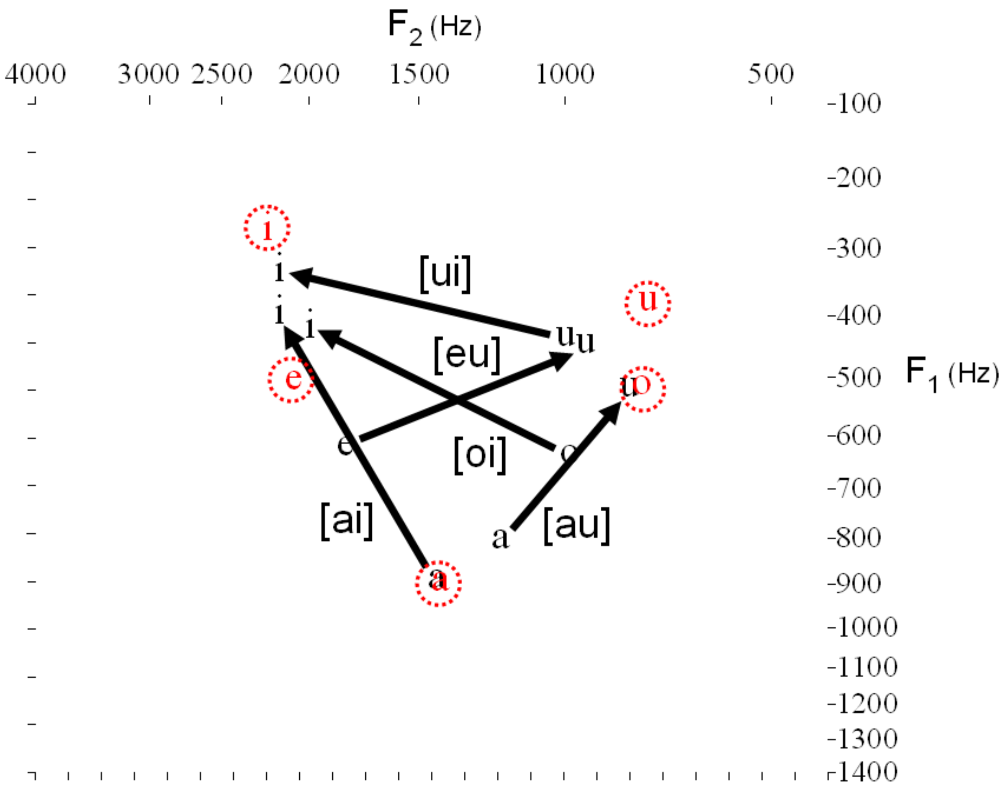


Figure 4.8 Formant movements for the diphthongs [eu, ui, oi, au, ai] and the 5 monophthongal vowels [i, e, a, o, u] in Meixian Hakka for 10 male speakers.

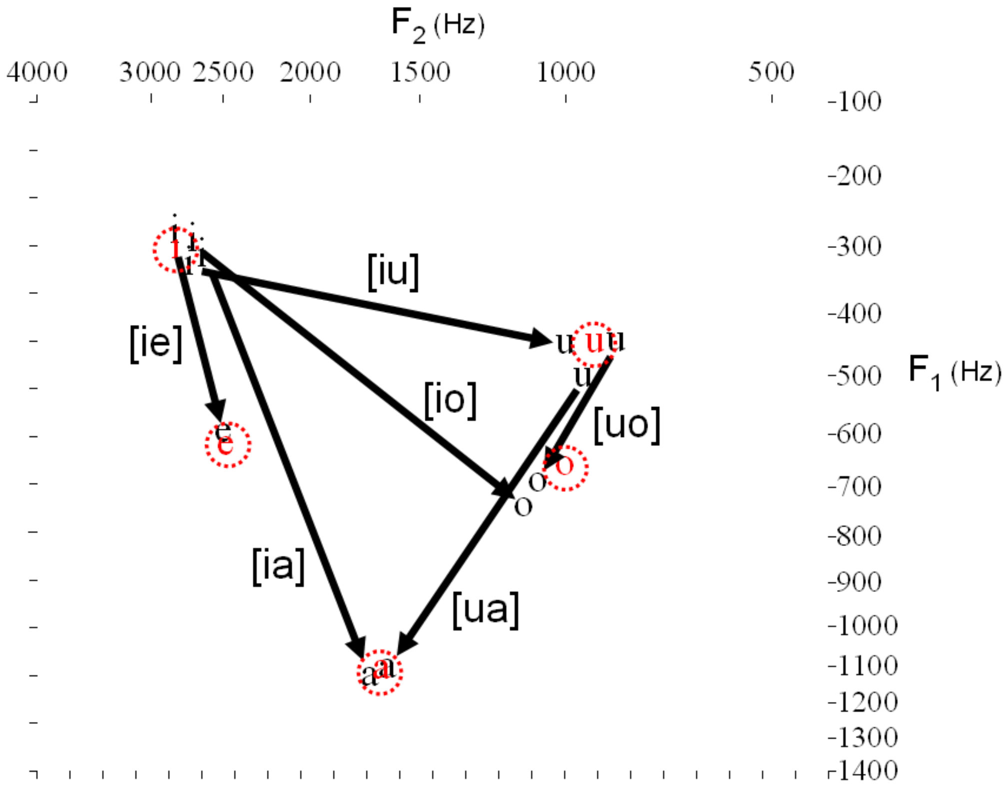


Figure 4.9 Formant movements for the diphthongs [ie, ia, io, iu, ua, uo] and the 5 monophthongal vowels [i, e, a, o, u] in Meixian Hakka for 10 female speakers.

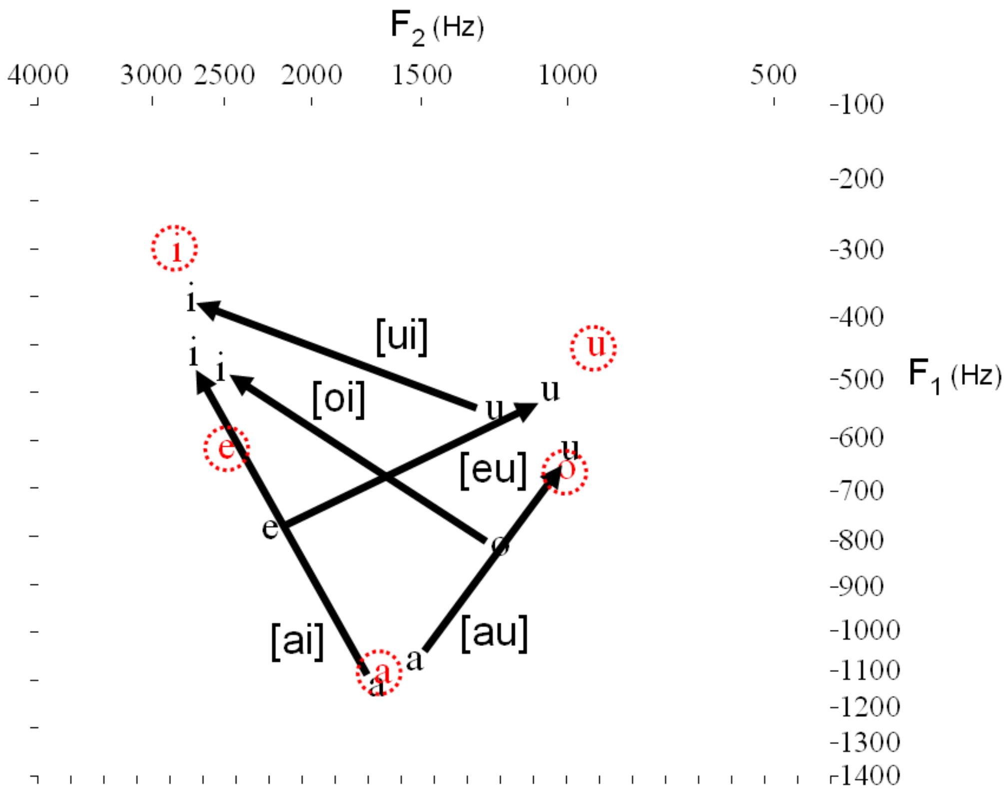


Figure 4.10 Formant movements for the diphthongs [eu, ui, oi, au, ai] and the 5 monophthongal vowels [i, e, a, o, u] in Meixian Hakka for 10 female speakers.

The acoustic data in Figures 4.7 and 4.9 indicates that the first and second elements of [ie, ia, io, iu, ua, uo] have F_1 and F_2 values that are equivalent to those of the monophthongal vowels, which are the corresponding target vowels of the diphthongs. In other words, the positions of both the first and second elements of the diphthongs in the acoustical vowel space are close to the positions of the monophthongs [i, e, a, o, u], respectively. As can be seen, the first and second elements of the diphthongs cluster around the positions of their target vowels in the acoustical vowel space. For examples, the positions of the first element of the diphthongs [ie, ia, io, iu] are located almost at the same position of [i] that produced in isolation. In contrast, Figures 4.8 and 4.10 indicate that the first and second elements of [eu, ui, oi, au] as well as the second element of [ai] do not achieve the formant frequencies of the corresponding target vowels. In comparison to the positions of their target vowels on the F1/F2 plane, the first and second elements of [eu, ui, oi, au] and the second element of [ai] appear to be centralized. For instance, [i] as the second element of [ui, oi, ai] has higher F_1 values than the monophthong [i] and they are located at lower position in the vowel space than the monophthongal vowel [i]. Given the spectral data, it may be rational to separate the Meixian Hakka diphthongs into two categories. The first category consists of [ie, ia, io, iu, ua, uo], in which the F_1 and F_2 of the two elements for a diphthong resemble those of the monophthongal vowels. The second category consists of [eu, ui, oi, au, ai], in which vowel undershoot occurs and the two elements of a diphthong do not reach the corresponding target vowels. This phenomenon has not been found in the past studies of diphthongs in other Chinese dialects (Ren, 1986; Zee, 1999; 2001) and it may be a specific characteristic for the Hakka dialects.

Figures 4.8 and 4.10 show that the positions of the onset element and offset element of diphthongs [eu, ui, oi, au, ai] on the F1/F2 plane are similar for both male and female speakers. Unlike the first element [i] in diphthongs [ie, ia, io, iu] which resembles the target vowel [i] in monophthong (Figures 4.7 and 4.9), the second element [i] for diphthongs [ai, oi, ui] is located in between [i] and [e] on the acoustical vowel space (Figures 4.8 and 4.10). There is different degree of increase in F_1 and decrease in F_2 for [i] in the diphthongs [ai, oi, ui]. A comparison of the positions of [i] on the F1/F2 plane displays that [i] in the diphthong [oi] is at the lowest position as it has the highest F_1 value (male: 416 Hz; female: 482 Hz). The F_2 value of vowel [i] in [oi] (male: 1990 Hz; female: 2515 Hz) is similar to the mean F_2 value of the monophthong [e] (male: 2076 Hz; female: 2481 Hz). Besides, the mean F_2 values of [i] in [ai] and [ui] are similar and they are lower than the F_2 value of the monophthong [i] (male: 2219 Hz; female: 2812 Hz) and higher than the F_2 value of the monophthong [e] (male: 2076 Hz; female: 2481 Hz). In terms of F_1 value, [i] in [ai] has a relatively higher value than [i] in [ui] and the former is located at a lower vertical position in the acoustical vowel space. According to the positions of [i] on the F1/F2 plane, it is rational to transcribe the second element in the diphthong [ui] as a close front unrounded vowel [i] and that in the diphthongs [ai] and [oi] as a near-close near front unrounded vowel [ɪ].

In addition to [i] in diphthongs of which the F_1 value is higher and F_2 value is lower than those of the monophthong [i], vowel undershoot is also found in [u]. As can be seen, the F_1 and F_2 values of [u] in diphthongs [eu, ui, au] are higher than those of the monophthong [u]. The positions of [u] in diphthongs are closer to the centre of the acoustical vowel space, in comparison with [u] in monophthong. The positions of [u] in [ui, eu] on the F1/F2 plane are in between a close back rounded vowel [u] and a close-mid back rounded vowel [o], regardless of whether it is the onset or offset of a diphthong. The positions of [u] in [ui, eu] indicate that the first element of [ui] and the

	Male speakers			Female speakers		
	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)
[ai] (1st)	882 (72)	1431 (79)	2571 (282)	1138 (117)	1686 (52)	2878 (294)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[ai] (2nd)	394 (60)	2155 (234)	2907 (237)	457 (96)	2697 (189)	3314 (180)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[ia] (1st)	273 (30)	2118 (205)	3233 (261)	311 (38)	2634 (170)	3685 (255)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[ia] (2nd)	898 (56)	1462 (86)	2526 (246)	1120 (98)	1712 (89)	2751 (286)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[ui] (1st)	430 (41)	997 (114)	2694 (195)	547 (55)	1228 (160)	3052 (142)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[ui] (2nd)	330 (42)	2151 (209)	2800 (238)	373 (39)	2715 (174)	3273 (155)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iu] (1st)	281(27)	2155 (214)	3048 (190)	323 (61)	2719 (158)	3505 (242)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iu] (2nd)	356 (40)	839 (92)	2562 (210)	444 (54)	1001 (96)	2950 (171)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[oi] (1st)	628 (46)	982 (48)	2702 (321)	806 (49)	1207 (103)	2967 (364)
[o]	508 (46)	789 (61)	2817 (287)	654 (51)	1002 (88)	3110 (222)
[oi] (2nd)	416 (90)	1990 (255)	2635 (252)	482 (74)	2515 (244)	3186 (174)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[io] (1st)	261 (26)	2145 (228)	3094 (219)	288 (43)	2697 (182)	3536 (271)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[io] (2nd)	561 (45)	879 (48)	2792 (323)	733 (89)	1128 (125)	3134 (232)
[o]	508 (46)	789 (61)	2817 (287)	654 (51)	1002 (88)	3110 (222)
[au] (1st)	795 (59)	1200 (70)	2642 (379)	1069 (92)	1525 (76)	2865 (372)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[au] (2nd)	512 (68)	822 (81)	2800 (347)	618 (83)	992 (112)	3019 (296)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[ua] (1st)	428 (58)	844 (111)	2511 (235)	498 (44)	950 (70)	2674 (226)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[ua] (2nd)	880 (75)	1371 (88)	2492 (349)	1101 (93)	1637 (82)	2740 (258)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[eu] (1st)	614 (61)	1818 (140)	2580 (140)	772 (67)	2221 (146)	2986 (201)
[e]	501 (41)	2076 (174)	2644 (157)	617 (63)	2481 (125)	3161 (213)
[eu] (2nd)	441 (62)	938 (119)	2424 (294)	516 (73)	1050 (128)	2911 (243)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[ie] (1st)	268 (31)	2193 (206)	3239 (242)	276 (34)	2820 (188)	3814 (226)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[ie] (2nd)	500 (50)	2027 (148)	2685 (199)	594 (51)	2501 (148)	3185 (204)
[e]	501 (41)	2076 (174)	2644 (157)	617 (63)	2481 (125)	3161 (213)
[uo] (1st)	363 (41)	758 (83)	2416 (352)	441 (51)	860 (73)	2755 (282)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[uo] (2nd)	552 (40)	847 (50)	2771 (289)	684 (61)	1081 (105)	3027 (174)
[o]	508 (46)	789 (61)	2817 (287)	654 (51)	1002 (88)	3110 (222)

Table 4.4 Mean values of F_1 and F_2 (in Hz) and their standard deviations for the first and second elements of diphthongs [ai, ia, oi, io, au, ua, iu, ui, ie, uo, eu] and their target vowels in Meixian Hakka for 10 female and 10 male speakers.

second element of [eu] should be treated as a near-close near back rounded vowel [ʊ]. It should be noted that for both male and female speakers, the formants of the second element in [au] resemble closely the formants of the monophthong [o].

Table 4.4 shows the formant frequency values of F_1, F_2 and F_3 of the diphthongs and the corresponding target vowels in Meixian Hakka produced by 10 male and 10 female speakers. The standard deviation of each element is also shown in bracket. The mean frequency values of F_1, F_2, and F_3 of the second element in [au] are 512 Hz, 822 Hz, and 2800 Hz, respectively, for the male speakers and they are 618 Hz, 992 Hz and 3019 Hz, respectively, for the female speakers. The formant pattern of the second element in [au] corresponds to the monophthong [o] rather than the monophthong [u]. The mean F_1, F_2 and F_3 values of [o] are 509 Hz, 789 Hz and 2817 Hz for the male speakers, and 654 Hz, 1002 Hz, and 3110 Hz for the female speakers, whereas those of [u] are 374 Hz, 776 Hz and 2653 Hz for the male speakers, and 442 Hz, 916 Hz, and 2935 Hz for the female speakers. The first element of the diphthong [eu] is in between a close-mid front vowel [e] and an open vowel [a] on the F1/F2 plane. It is located at a lower level of vowel height than the monophthong [e] in the acoustical vowel space and its position resembles an open-mid front unrounded vowel [ɛ]. It shows that the frequency value of F_1 for the first element in [eu] (male: 614 Hz; female: 772 Hz) is greater than that for the target vowel [e] (male: 501 Hz; female: 617 Hz), whereas the F_2 value for the first element in [eu] (male: 1818 Hz; female: 2221 Hz) is lower than that for the monophthong (male: 2076 Hz; female: 2481 Hz). Likewise, [o] in the diphthong [oi] is at a lower level of vowel height than the monophthong [o] in the acoustical vowel space. [o] in [oi] has higher F_1 and F_2 values in comparison with the monophthong [o]. In consideration of the formant pattern, the first element in [oi] should be treated as an open-mid back rounded vowel [ɔ]. The first elements of the diphthongs [eu] and [oi] have similar F_1 values and they are at the same level of vowel height as shown in Figures 4.8 and 4.10. The first element of [eu] is in between the monophthong [e] and the monophthong [a], whereas the first element of [oi] is in between the monophthongs [o] and [a]. For the Meixian Hakka monophthongs, [a] seems to be resistant to the so-called effect of vowel undershoot. Figure 4.10 displays that, for the female speakers, there is no difference between the positions of the first elements of the diphthongs [ai, au] and that of the monophthong [a] within the acoustical vowel space. While for the male speakers in Figure 4.8, on one hand, the mean F_1 and F_2 values for the first element of [ai] are equivalent to those of the monophthong [a]. On the other hand, the first element of [au] is in between the monophthongs [o] and [a], and it might be logical to transcribe the sound segment as a near-open central vowel [ɐ].

As shown in Table 4.4, the two elements of [eu, ui, oi] and the second elements of [ai, au] have higher F_1 values than the F_1 values for the monopthongs [i, e, o, u]. It indicates that these diphthongs have more open articulatory positions than the monophthongal vowels. Besides, F_2 measurements indicate that the second elements of [eu, au] and the first element of [oi] are pronounced in slightly more front articulatory positions than the monophthongs [u, o], whereas the second elements of [ui, oi, ai] and the first element of [eu] are pronounced in slightly more retracted positions than the monophthongs [i, e]. Such acoustic outcomes may be correlated with the activities of genioglossus and styloglossus as stated in the study of Collier and Raphael (1982). As such correlation is related more to physiology of the vocal apparatus than phonetics, it would not be illustrated in detail in this chapter.

The F_2 range, transition duration between the two target positions and F_2 rate of change for the Meixian Hakka diphthongs have been calculated and are displayed in Table 4.5. A comparison of the F_2 rate of change between male and female speakers

Test diphthong	F_2 range (Hz) (s.d.)		Transition duration (ms) (s.d.)		F_2 rate of change (Hz/ms) (s.d.)	
	Male	Female	Male	Female	Male	Female
[ua]	527 (124)	686 (114)	122.7 (30)	129.9 (30)	4.51 (1.2)	5.54 (1.5)
[au]	378 (98)	532 (129)	131.9 (34)	134.9 (46)	2.99 (0.9)	4.48 (2.4)
[ia]	656 (152)	921 (164)	100.9 (17)	111.5 (21)	6.67 (1.7)	8.43 (1.6)
[ai]	724 (182)	1010 (183)	184.7 (44)	183.7 (44)	4.03 (1.0)	5.75 (1.4)
[io]	1265 (247)	1569 (251)	111.2 (21)	107.3 (23)	11.67 (2.7)	15.13 (3.2)
[oi]	1007 (276)	1308 (290)	152.2 (31)	155.1 (42)	6.68 (1.7)	9.04 (3.0)
[ui]	1153 (240)	1486 (233)	130.2 (45)	137.6 (42)	9.83 (3.4)	11.71 (3.5)
[iu]	1315 (231)	1717 (172)	166 (47)	180.8 (42)	8.48 (2.8)	9.94 (2.2)
[uo]	97 (57)	220 (108)	76.9 (24)	84.0 (38)	1.23 (0.6)	3.00 (2.0)
[ie]	169 (102)	319 (133)	87.0 (21)	97.8 (35)	2.15 (1.6)	3.61 (1.9)
[eu]	879 (217)	1170 (214)	118.0 (25)	116.6 (21)	7.64 (1.9)	10.19 (1.9)

Table 4.5 Mean F_2 range (Hz), transition duration (ms) and F_2 rate of change (Hz/ms) for the Meixian Hakka diphthongs produced by male and female speakers.

shows that diphthongs produced by the female speakers have a faster F_2 rate of change than those produced by the male speakers. The difference in F_2 rate of change between the two groups of speakers is consistent across the diphthongs in Meixian Hakka. It may be noticed that, in fact, the duration of transition is similar for both male and female speakers. Therefore, the F_2 rate of change for the female speakers is faster than that for the male speakers as a result of the greater F_2 frequency range for the female speakers than the male speakers. And such difference in F_2 range between the male and female speakers may be explained by a larger acoustical vowel space for the female speakers. The results of this study show that the Meixian Hakka vowels for the female speakers occupy a larger acoustical vowel space on the F1/F2 plane than those for the male speakers. Therefore, it should not be surprised that the F_2 range for the male and female speakers is different.

Table 4.6 presents the ranking order of the Meixian Hakka diphthongs with respect to the F_2 rate of change, ranging from the slowest to fastest rate of change. The F_2 rate of change for the male speakers ranged from the slowest 1.23 Hz/ms to the fastest 11.67 Hz/ms, whereas that for female speakers ranged from 3 Hz/ms to 15.13 Hz/ms. For both

Test diphthong	Male	Test diphthong	Female
	F_2 rate of change (Hz/ms) (s.d.)		F_2 rate of change (Hz/ms) (s.d.)
[uo]	1.23 (0.6)	[uo]	3.00 (2.0)
[ie]	2.15 (1.6)	[ie]	3.61 (1.9)
[au]	2.99 (0.9)	[au]	4.48 (2.4)
[ai]	4.03 (1.0)	[ua]	5.54 (1.5)
[ua]	4.51 (1.2)	[ai]	5.75 (1.4)
[ia]	6.67 (1.7)	[ia]	8.43 (1.6)
[oi]	6.68 (1.7)	[oi]	9.04 (3.0)
[eu]	7.64 (1.9)	[iu]	9.94 (2.2)
[iu]	8.48 (2.8)	[eu]	10.19 (1.9)
[ui]	9.83 (3.4)	[ui]	11.71 (3.5)
[io]	11.67 (2.7)	[io]	15.13 (3.2)

Table 4.6 Mean F_2 rate of change (Hz/ms) and standard deviation for the Meixian Hakka diphthongs. The F_2 rates of change are listed in ascending order for the male and female speakers.

male and female speakers, [uo] has the slowest rate of change and [io] has the fastest rate of change. In terms of the transcription of Meixian Hakka diphthongs, there are many pairs of mirror-image diphthongs in the vowel system of Meixian Hakka and they are [au] – [ua], [ai] – [ia], [oi] – [io] and [iu] – [ui]. The two diphthongs in the same pair of mirror-image diphthongs have the two target elements of which the formant movements are in two directions. As can be noticed, the F_2 rate of change for the diphthongs with a high vowel onset is relatively faster than that for its mirror-image counterpart with a high vowel offset. For examples, [ua] has a counterpart [au] (male: 2.99 Hz/ms, female: 4.48 Hz/ms). The same patterns of F_2 rate of change can be observed in other pairs of mirror-image diphthongs in Meixian Hakka and it is consistent in the data of both male and female speakers. With respect to the vowel height and backness of the two elements, a diphthong with formant movement that undergoes a considerable shift along the front and back axis of the F1/F2 plane has a relatively faster transition rate than other diphthongs. As shown in Table 4.6, there is a general tendency that the diphthongs [oi, io, eu, iu, ui] has a faster F_2 rate of change

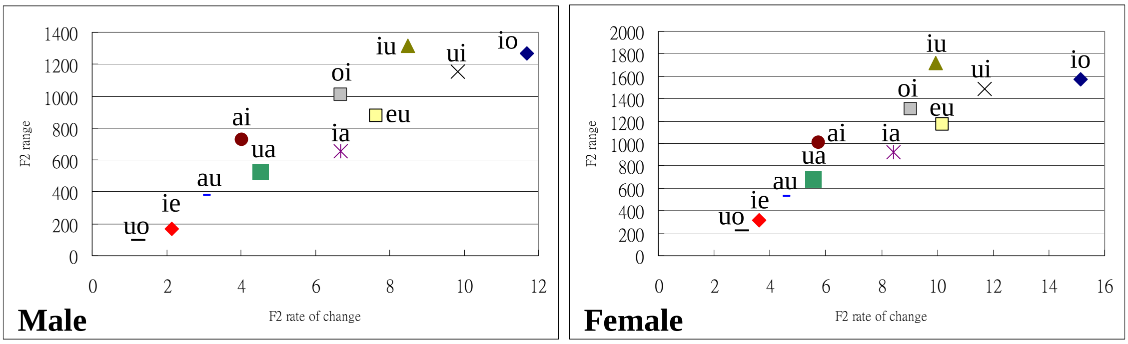


Figure 4.11 Correlation between the F_2 range and F_2 rate of change for the Meixian Hakka diphthongs.

than other diphthongs in Meixian Hakka. In contrast, a diphthong with formant movement that involves only a small vertical shift on the F1/F2 plane results in the slowest transition rate. The results of this study indicate that [uo] and [ie] have the slowest F_2 rate of change. The two elements of [uo] are back vowels and those of [ie] are front vowels. In this case, the two elements in [uo] and [ie] are distinguished from each other by the difference in vowel height. The onsets of the diphthongs are high vowels, whereas the offsets are high-mid vowels.

Figure 4.11 displays the correlation between the F_2 range and F_2 rate of change for the Meixian Hakka diphthongs. The x-axis of the graphs indicates the F_2 rate of change and the y-axis indicates that the mean F_2 difference between the two elements of a diphthong. Results show that there is a positive correlation between the F_2 rate of change and F_2 range in Meixian Hakka. The correlation implies that the smaller the mean F_2 difference between the two elements, the slower the F_2 rate of change, and vice versa. The pattern of correlation between the male and female speakers is similar. [uo] has the smallest distance between the onset and offset of a diphthong and it turns out to have the slowest F_2 rate of change. On the contrary, [io] has the nearly greatest F_2 difference and its F_2 rate of change is the fastest among the Meixian Hakka diphthongs. It seems that the F_2 range and F_2 rate of change may contribute as primary cues for diphthong perception in Meixian Hakka. With the assumption that each diphthong in a language has a relative constant F_2 rate of change, results shows that most diphthongs in a language can be differentiated in terms of F_2 transition rate.

Figures 4.12 and 4.13 show the temporal organizations of the 11 diphthongs [ai, ia, oi, io, au, ua, iu, ui, ie, uo, eu] in Meixian Hakka. The onset duration, offset duration as well as the duration of the transition between the two elements of each diphthong are presented. Result shows that the temporal organizations support the classification of the two categories of the Meixian Hakka diphthongs. Regardless of [iu], the diphthongs in the first category [ie, ia, io, ua, uo], of which the formant patterns of the diphthongs are more or less identical to those of the target vowels, tend to have a relatively short steady state of the first element, and a relatively long steady state of the second element, in comparison with those of the diphthongs in the second category. In contrast, for the diphthongs in the second category [eu, ui, oi, au, ai], the durations of onset state steady, transition and offset steady state are similar. This may imply that the formant patterns and temporal organization of the diphthongs may be different in order to maximize the minimum distance between any pair of diphthongs in the two categories. The temporal organization of the diphthong [iu] is different from other diphthongs in the first category. As can be seen, the temporal organization of [iu] is in fact similar to that of the diphthongs in the second category. The steady state of the first element of [iu] has a long duration and it is sometimes longer than the first elements of other Meixian Hakka diphthongs that produced by most male and female speakers.

Figures 4.12a and 4.13a demonstrate the temporal organizations of the 11 Meixian Hakka diphthongs for every individual male and female speaker. Though between-speaker variation can be found in the temporal organizations of the Meixian Hakka diphthongs, the observable distinction between the two categories of diphthongs is

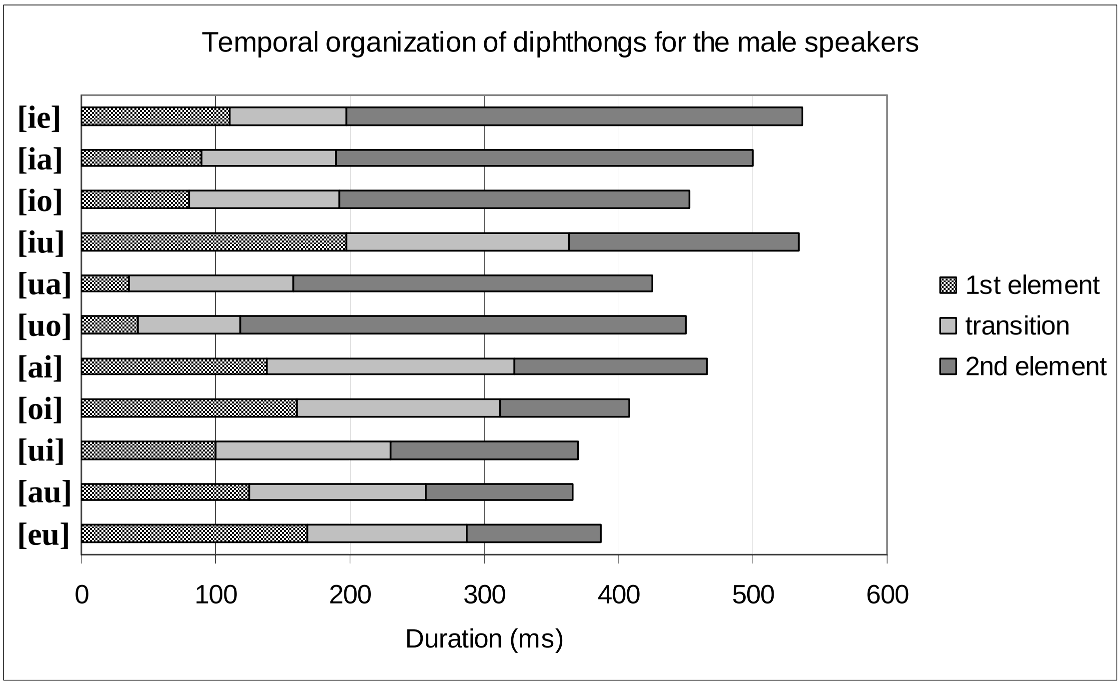


Figure 4.12 Temporal organization of the 11 diphthongs [eu, au, ui, oi, ai, uo, ua, iu, io, ia, ie] in Meixian Hakka for the male speakers.

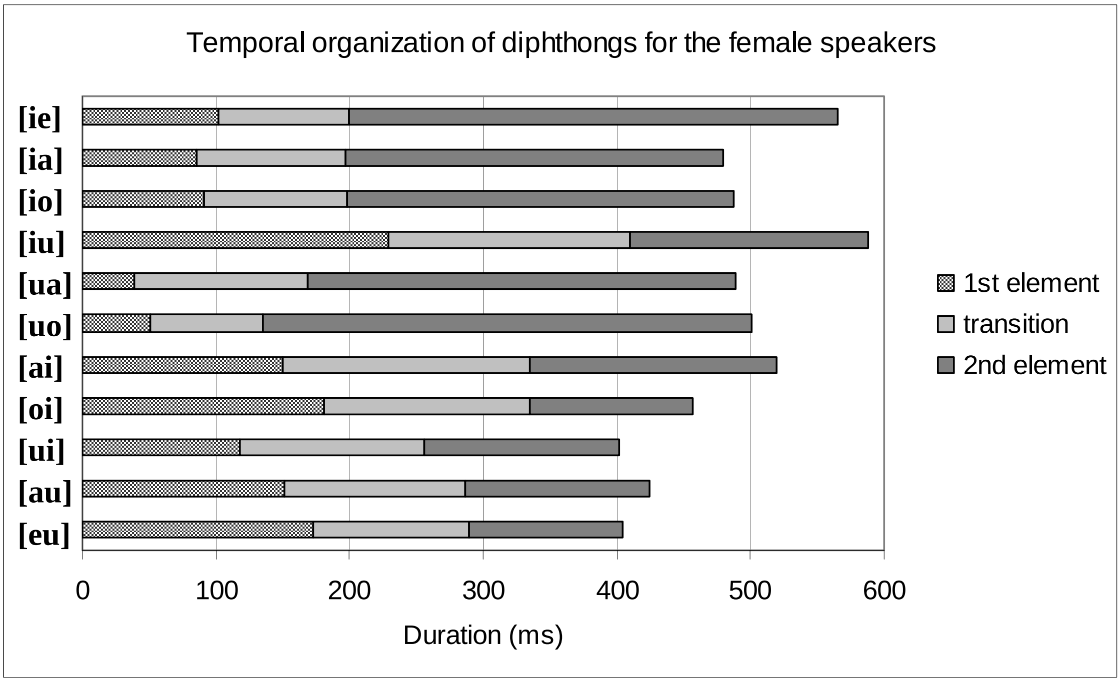


Figure 4.13 Temporal organization of the 11 diphthongs [eu, au, ui, oi, ai, uo, ua, iu, io, ia, ie] in Meixian Hakka for the female speakers.

consistent for the majority of speakers. The mean total duration shows that the diphthongs [ie, iu] are the longest among the Meixian Hakka diphthongs and it applies to the data of both male and female speakers. Besides, the mean total durations of [ui, au, eu] show that they are the shortest diphthongs. The temporal data of every individual speaker indicates that [ie, iu] as the longest diphthongs and [ui, au, eu] as the shortest diphthongs generally holds true for most speakers in this study. In terms of the temporal organization of a diphthong, it is proposed that the so-called diphthongs [ua, uo] should be considered as monophthongs instead. The temporal data of [ua] and [uo] shows that there is no or a very brief steady state in most cases for [u] in [ua] and [uo]. As shown in Figures 4.12a and 4.13a, onset steady state for the first element of [ua, uo] was not observed in the acoustic data of male speaker 1 as well as female speakers 7, 9 and 10. In addition, the onset steady state for [u] in [ua, uo] is very brief for other speakers. The absolute durations of vowels, diphthongs and triphthongs for every individual speaker are listed in Appendix C. Each data in the tables displays the average value of all obtained data points. For some data, there is a number displayed in a small circle. The number inside the circle indicates the total number of data points for the calculation of the average value. For examples, the symbol ② represents that the average value is the mean of two data points and no duration of the specific period was observed in other repetitions of the same test word. In this study, for the majority of speakers, the mean duration of the first element for [ua, uo] is less than 50 ms. The acoustic and perceptual characteristics of initial approximants in English have been investigated in the study of O'Connor et al. (1957). The steady-state duration of the two glides /w, j/ was claimed to be about 30 or 40 ms and it was shorter than the duration of /r, l/ which was about 50 or 60 ms. It was even remarked that the steady-state period of /w, j/ could be dispensed. For the relatively short steady-state period of /w, j/, duration of more than 40 ms would be perceived as a full vowel /u/ or /i/. In this case, the results of this study seem to support the claim that the first element of [ua, uo]

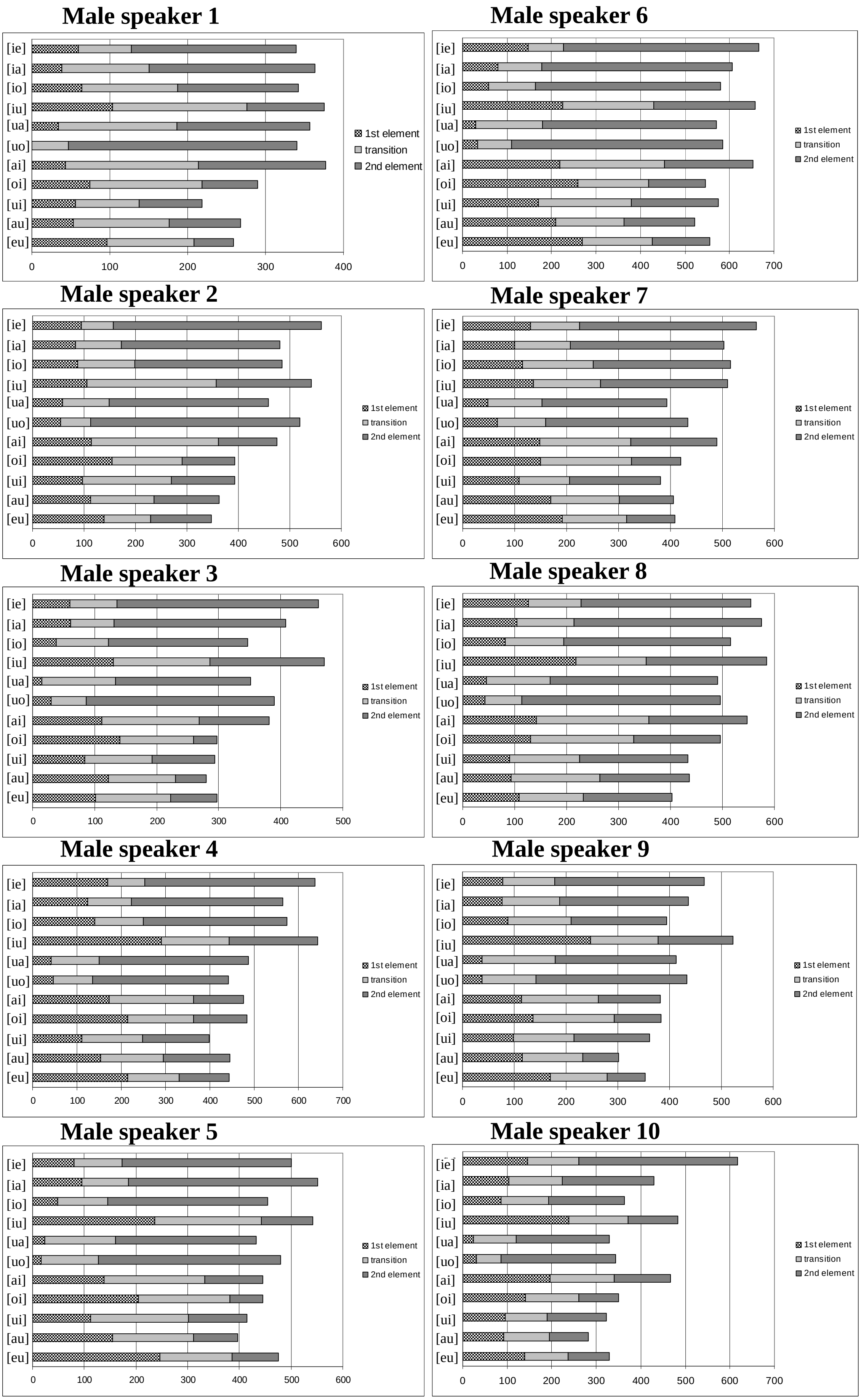


Figure 4.12a Temporal organizations of the 11 Meixian Hakka diphthongs [eu, au, ui, oi, ai, uo, ua, iu, io, ia, ie] that produced by 10 male speakers

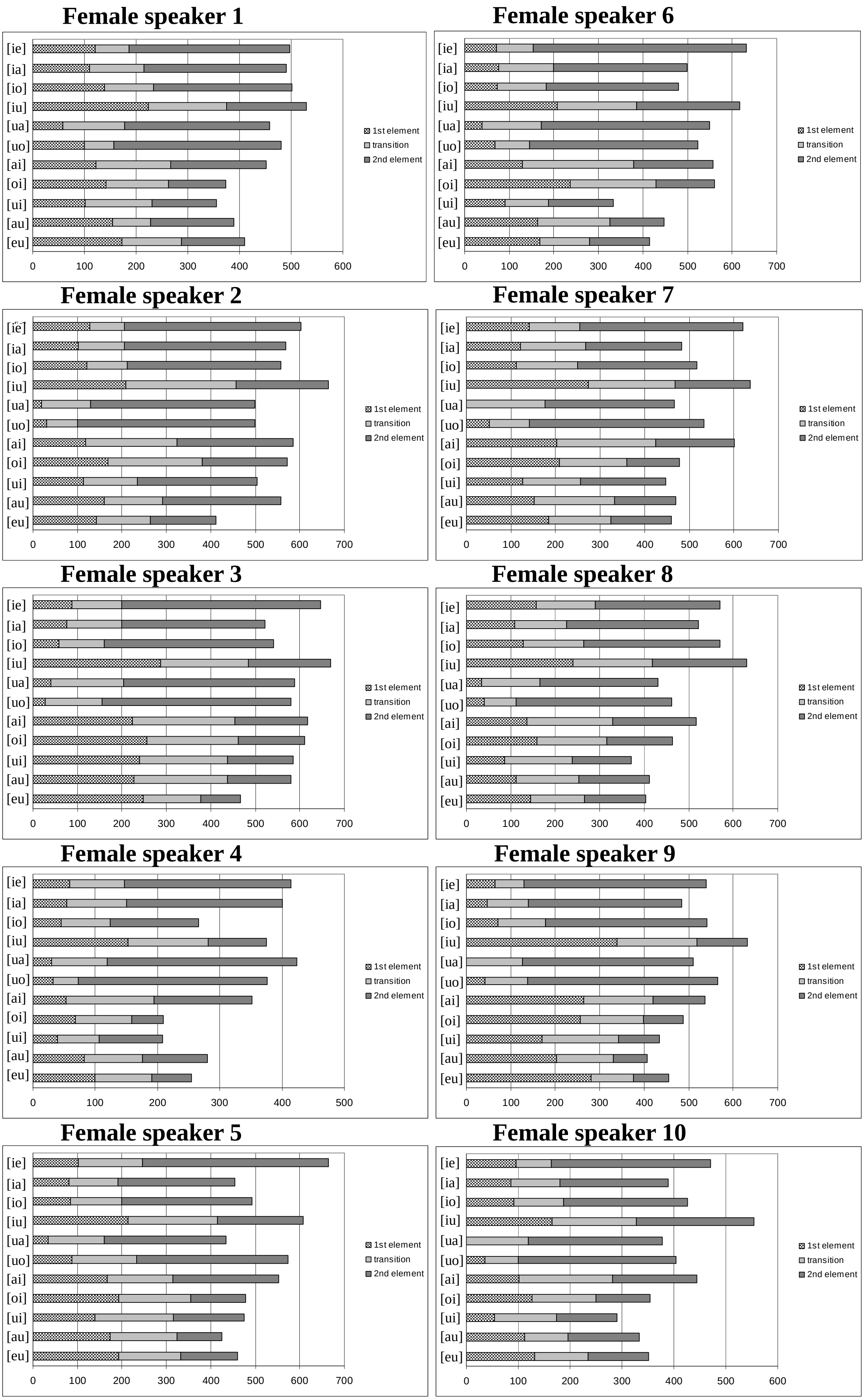


Figure 4.13a Temporal organizations of the 11 Meixian Hakka diphthongs [eu, au, ui, oi, ai, uo, ua, iu, io, ia, ie] that produced by 10 female speakers

might resemble the glide /w/. In consideration of phonology, [ua] and [uo] occur only after velar stops [k] and [k^h] in Meixian Hakka as mentioned in the past studies (Huang, 1995; Wen, 2006; Yuan et al., 2001). In this case, it may be more appropriate to consider the sound segments as the monophthong [a] and [o] which preceded by labialized velar stops [k^w] and [k^{wh}]. It is rational as consonants [k^w] and [k^{wh}] have also been used in transcribing Cantonese to reduce the complexity of the vowel system. In conclusion, based on the spectral data and the temporal organization of the diphthongs, it claims that there should be nine diphthongs in Meixian Hakka and they should be phonetically transcribed as [ie, ia, io, iu, aɪ, ɔɪ, ʊi, ao, ɛʊ]. In this study, there are two phonetic realizations for diphthong /au/. For the female speakers, the formants of [a] in [ao] are similar to those of the monophthong [a] and the diphthong was pronounced as [ao], whereas for the male speakers, [a] in [ao] have lower F_1 and F_2 values than the monophthong [a] and it was pronounced as [ɐo] instead. The so-called diphthongs [ua] and [uo] as described in the past studies should be considered as monophthongs that preceded by consonants [k^w] and [k^{wh}].

4.3.3 Triphthongs

Triphthong is not common in the sounds of the world languages and the number of languages containing triphthongs is comparatively small. As long as linguistics has existed, there seems to be little or no theoretical discussion or phonological investigation upon the definition and classification of triphthongs. As mentioned, there are four triphthongs in Meixian Hakka which have been described in the past literatures (Huang, 1995; Wen, 2006; Yuan et al., 2001). Similar to diphthongs, triphthongs are defined in this study as a sequence of three vocalic elements that are perceptually different. It is supposed that a triphthong would involve a relatively quick change in vowel quality from the onset steady state to the offset steady state by passing over a target vowel in the middle that may or may not have the steady-state portion. In this section, the results of a supplementary study investigating the triphthongs /uoi/ and /ioi/ are presented first, followed by the acoustic data of other triphthongs in Meixian Hakka. Figure 4.14 shows the photographs of the position of the lips during the first sound in /uoi/ that produced by a native female speaker. The two photographs were captured from the video recordings of the test words. Results show that the first sound of the so-called triphthong /uoi/ is articulatorily a labiodental sound. The speaker brings the inner surfaces of the lower lip into contact with the upper teeth and there is no apparent rounding of the lips. It should be noted that lips positions were consistent in all five repetitions of the recordings. Figure 4.15 displays the waveform, spectrogram and formant trajectories for the so-called triphthong [ʊɔɪ] (i.e. [ʋɔɪ]) 'to bake in hot ashes' in Meixian Hakka. The spectrogram shows that the first sound of the so-called triphthong [ʊɔɪ] is a frictionless voiced sound and it is maintainable in duration. It can be noticed that there is no turbulence at the beginning of the syllable and the formant pattern of the first sound in the so-called triphthong [ʊɔɪ] is different from the formant pattern of vowel [ʊ] which should have a low F_1 and F_2. Based on the articulatory and acoustic data in this study, it seems to be more rational to treat the first sound in the so-called triphthong /uoi/ as a voiced labiodental frictionless continuant [ʋ] rather than a vowel [ʊ]. In that case, the so-called triphthong /uoi/ as listed in Hashimoto's study (1973), which was claimed to be acoustically [ʊɔɪ], should be a sequence of an approximant [ʋ] and a diphthong [ɔɪ]. The findings of this study support the phonological description of the identical syllable in the studies of Huang (1992; 1995), Wen (2006) and Yuan et al. (2001), which indicate that the so-called triphthong /uoi/ does not exist in the sound system of Meixian Hakka.

Figure 4.14 The lip positions during the first sound of the so-called triphthong [ʊɔɪ] (i.e., [ʊɔɪ]) ‘to bake in hot ashes’ in Meixian Hakka for a female speaker. The left picture shows the front face view and the right picture shows the side face view of a speaker.

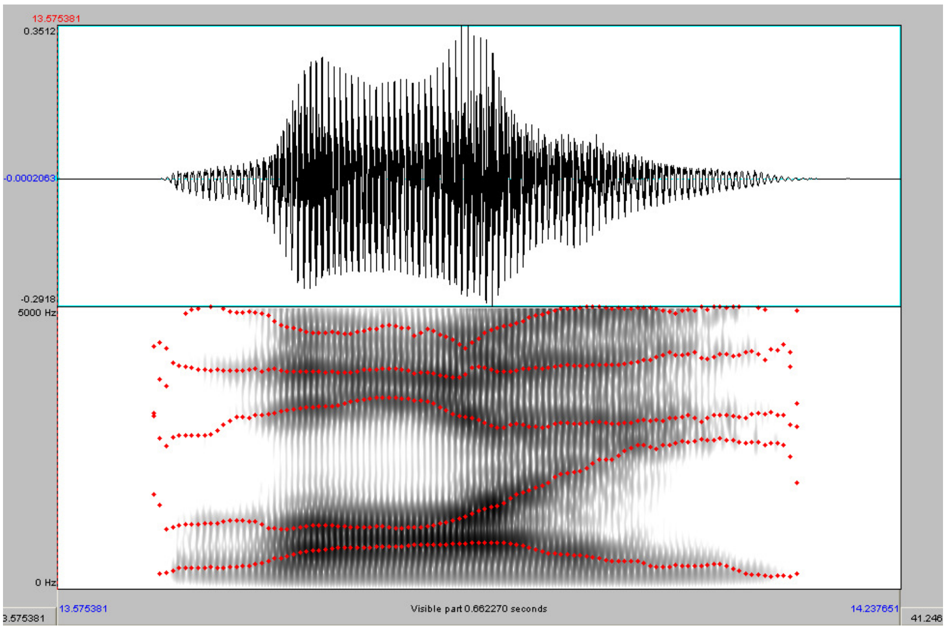


Figure 4.15 Formant trajectories for the so-called triphthong [ʊɔɪ] (i.e., [ʊɔɪ]) ‘to bake in hot ashes’ in Meixian Hakka for a female speaker.

In Hashimoto’s study (1973), the word which meant ‘to be tired’ in Meixian Hakka was transcribed as a CV syllable with a consonant [k^h] followed by a triphthong [iɔɪ], whereas the word which meant ‘to open’ as a sequence of consonant [k^h] and a diphthong [ɔɪ]. Despite their tonal difference, there is no distinction in transcription between these two words in other studies (Huang, 1995; Xie, 1994). In this study, the

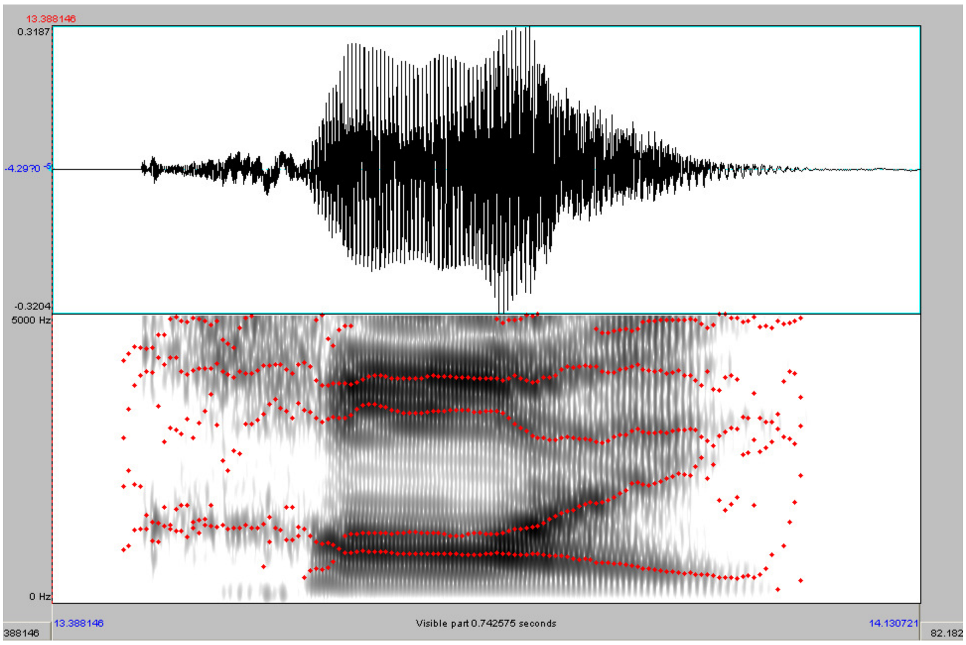


Figure 4.16 Waveform, spectrogram and formant trajectories for [k^hiɔɪ51] (i.e., [k^hɔɪ]) ‘to be tired’ in Meixian Hakka produced by a female speaker.

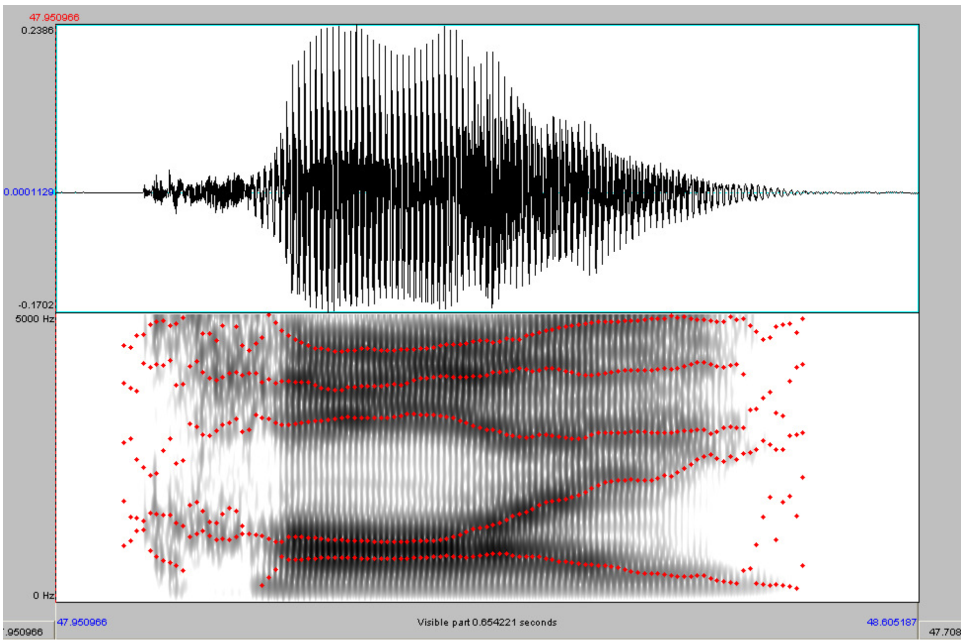


Figure 4.17 Waveform, spectrogram and formant trajectories for [k^hɔɪ44] 开 ‘to open’ in Meixian Hakka produced by a female speaker.

spectral data may provide evidence on proving whether a triphthong /ioi/ exists in the sound system of Meixian Hakka. Figure 4.16 and Figure 4.17 show the waveform, spectrogram and formant trajectories for the CV syllables [k^hiɔɪ] (i.e., [k^hɔɪ]) ‘to be tired’ and [k^hɔɪ] ‘to open’ in Meixian Hakka, respectively. The formant trajectories in

Figure 4.16 show that there is neither a vowel [i]-like sound nor a palatalized [k] preceding the steady-state of vowel [ɔ] in the syllable. Figures 4.16 and 4.17 indicate that the formant trajectories for the two words resemble those for a diphthong rather than a triphthong. The steady state of the vowel [ɔ] begins after the syllable-initial consonant [k^h] and the steady states of F_1, F_2 and F_3 trajectories for vowel [ɔ] in the two words are maintained for nearly 50% of the entire syllable duration. The spectral data indicates that the so-called triphthong [iɔɪ] in Hashimoto's study (1973) should be transcribed as a diphthong [ɔɪ].

Figure 4.18 and Figure 4.19 show the waveform, spectrogram and formant trajectories for the test word 扣 'to buckle' in Meixian Hakka, which was transcribed as [k^hieu] in the studies of Huang (1992) and Xie (1994), but as [k^heu] in the study of Huang (1995) and Yuan et al. (2001). The test word shown in Figure 4.18 was produced by a male speaker and that shown in Figure 4.19 was produced by a female speaker. As can be seen, for both male and female speakers, the formant trajectories for the test word resemble those for a diphthong rather than a triphthong. The formant frequencies of F_1, F_2 and F_3 remain steady after the vowel onset. The steady states of F_1, F_2 and F_3 trajectories are maintained for over 50% of the entire duration of the syllable. After the steady state of the vowel [e], there is an obvious decrease in F_2 and a slight decrease in F_1. The formant trajectories also indicate that there is neither a vowel [i] preceding the vowel [e] nor a palatalized [k] at the beginning of the syllable. The formant pattern of the test word, which resembles a diphthong [eu], is consistent for the 10 male and 10 female speakers in the study. The formant trajectories justify the treatment of so-called triphthong [ieu] in the CV syllables in Meixian Hakka as a diphthong [eu]. The results of this study indicate that the transcription of the so-called [ieu] in the study of Hashimoto (1973), Huang (1992) and Xie (1994) is not supported by the spectral

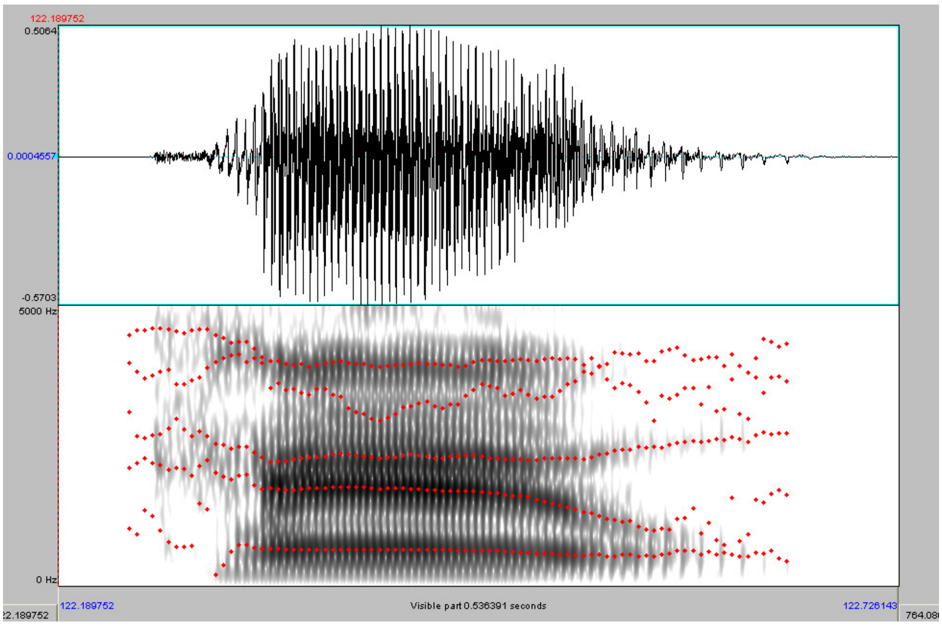


Figure 4.18 Waveform, spectrogram and formant trajectories for [k^{h}ieu] (i.e., [k^{h}eu]) 扣 'buckle' in Meixian Hakka produced by a male speaker.

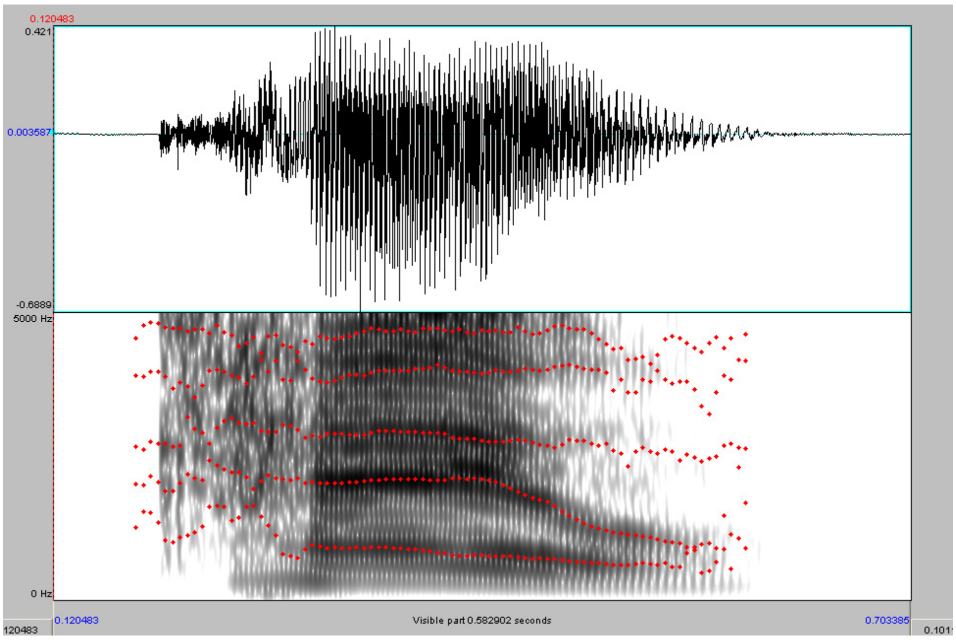


Figure 4.19 Waveform, spectrogram and Formant trajectories for [k^{h}ieu] (i.e., [k^{h}eu]) 扣 'buckle' in Meixian Hakka produced by a female speaker.

evidence. The formant pattern of the so-called triphthong [ieu] was not realized in the syllable production of the target test word. It further indicates that the complementary distribution between [ieu] and [eu] that conditioned by the syllable-initial consonants does not exist in Meixian Hakka. On the contrary, the spectral data support the

Test Vowels	Male speakers			Female speakers		
	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[e]	501 (41)	2076 (174)	2644 (157)	617 (63)	2481 (125)	3161 (231)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[o]	508 (46)	789 (61)	2817 (287)	654 (51)	1002 (88)	3110 (222)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[ɿ]	369 (37)	1293 (122)	2767 (226)	430 (55)	1438 (129)	3116 (194)

Table 4.7 Mean frequency values of F_1, F_2 and F_3 (in Hz) and their standard deviations (s.d.) for the monophthongs [i, e, a, o, u] and the apical vowel [ɿ] in Meixian Hakka that produced by 10 female and 10 male speakers.

	Male speakers			Female speakers		
	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)	F_1 (s.d.)	F_2 (s.d.)	F_3 (s.d.)
[iai] (1st)	469 (39)	1883 (233)	2804 (196)	568 (43)	2228 (117)	3073 (225)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iai] (2nd)	824 (64)	1506 (83)	2472 (233)	1083 (74)	1768 (74)	2705 (264)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[iai] (3rd)	476 (92)	2072 (229)	2723 (222)	538 (129)	2583 (282)	3303 (241)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iui] (1st)	259 (31)	2150 (222)	3034 (198)	285 (34)	2632 (92)	3369 (178)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iui] (2nd)	424 (50)	1100 (147)	2583 (201)	547 (63)	1315 (166)	2958 (135)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[iui] (3rd)	307 (33)	2145 (222)	2803 (263)	373 (45)	2621 (157)	3202 (232)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iau] (1st)	276 (11)	2170 (155)	3300 (214)	297 (38)	2658 (157)	3686 (241)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)
[iau] (2nd)	805 (53)	1203 (83)	2772 (160)	1047 (71)	1533 (97)	2779 (227)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[iau] (3rd)	478 (56)	809 (93)	2921 (238)	586 (110)	1007 (84)	3071 (266)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[uai] (1st)	428 (68)	851 (113)	2475 (259)	513 (47)	942 (82)	2764 (173)
[u]	374 (28)	776 (66)	2653 (228)	442 (48)	916 (78)	2935 (156)
[uai] (2nd)	850 (61)	1379 (78)	2426 (258)	1053 (77)	1604 (80)	2645 (234)
[a]	888 (64)	1414 (77)	2616 (283)	1104 (115)	1662 (76)	2839 (306)
[uai] (3rd)	408 (71)	2156 (236)	2787 (253)	466 (98)	2656 (179)	3317 (201)
[i]	270 (26)	2219 (234)	3194 (199)	298 (44)	2812 (133)	3686 (222)

Table 4.8 Mean values of F_1 and F_2 (in Hz) and their standard deviations for the first, second and third elements of triphthongs [iai, iui, iau, uai] and their target vowels in Meixian Hakka for 10 female and 10 male speakers.

phonological analysis of Hanyu Fangyin Zihui (1989), Huang (1995) and Yuan et al. (2001), which have grouped the relevant syllables under the rime of [eu]. Despite of the triphthong /ieu/, the phonetic data of the remaining four triphthongs in Meixian Hakka, which are [iai], [iui], [iau] and [uai], are presented in this chapter. The average values of F_1, F_2 and F_3 (in Hertz) for the monophthongal vowels [i, e, a, o, u] and the three elements of every triphthong in the (C)V syllables in Meixian Hakka for 10 male and 10 female speakers are listed in Tables 4.7 and 4.8, respectively. In addition, the standard deviations for the F-values are listed in the table. The average formant values show that [iai, iui, iau, uai] have higher F_1 values for the third element of the triphthongs than monophthongs [i, u]. It indicates that all triphthongs have more open articulatory positions at their ending than the monophthongal vowels. The F_1 values also indicate that [a] in [iai, iau, uai] have slight more close articulatory positions than the monophthong [a], whereas [u] in [iui, iau, uai] have more open articulatory positions than the monophthong [u] in general. Besides, the F_2 measurements show that [i] in [iai, iui, iau, uai] have lower F_2 values than the monophthong [i], indicating that [i] as an element of a triphthongs may have a more retracted articulatory position and longer oral resonating cavity. Similarly, for both male and female speakers, [u] in [iui, iau] have more front articulatory positions than the vowel [u].

Figures 4.20 and 4.21 show the positions of the first, second and third elements of the triphthongs [iai, iui, uai, iau] in Meixian Hakka on the F1/F2 plane for 10 male and 10 female speakers, respectively. The positions of the monophthongal vowels are also shown on the F1/F2 plane, which are displayed by the corresponding IPA symbols in dotted circles. The angled triphthong arrows on the F1/F2 plane indicate the formant movements for the 4 triphthongs [iai, iui, uai, iau]. The end point of an arrow without arrow head shows the first element of a triphthong; the middle point shows the second element of a triphthong; and the end point of an arrow with arrow head shows the third element of a triphthong. As can be seen, even though there is some discrepancy between

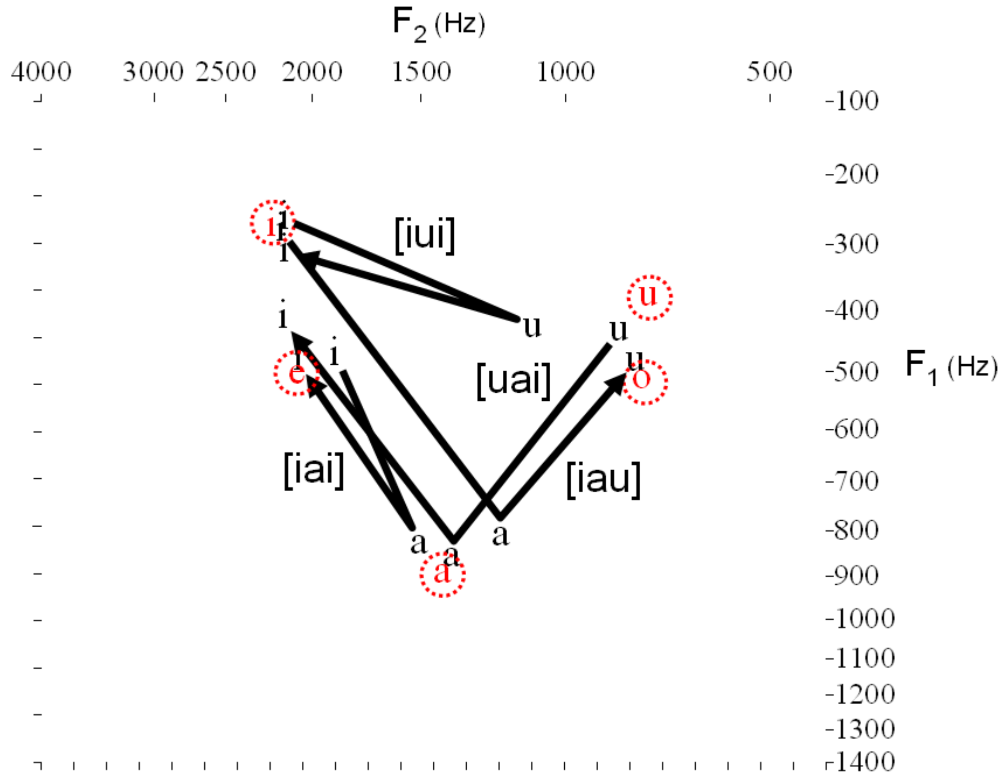


Figure 4.20 Formant movements for the triphthongs [iai, iui, iau, uai] and the positions of the 5 monophthongal vowels [i, e, a, o, u] in Meixian Hakka on the F1/F2 plane for 10 male speakers.

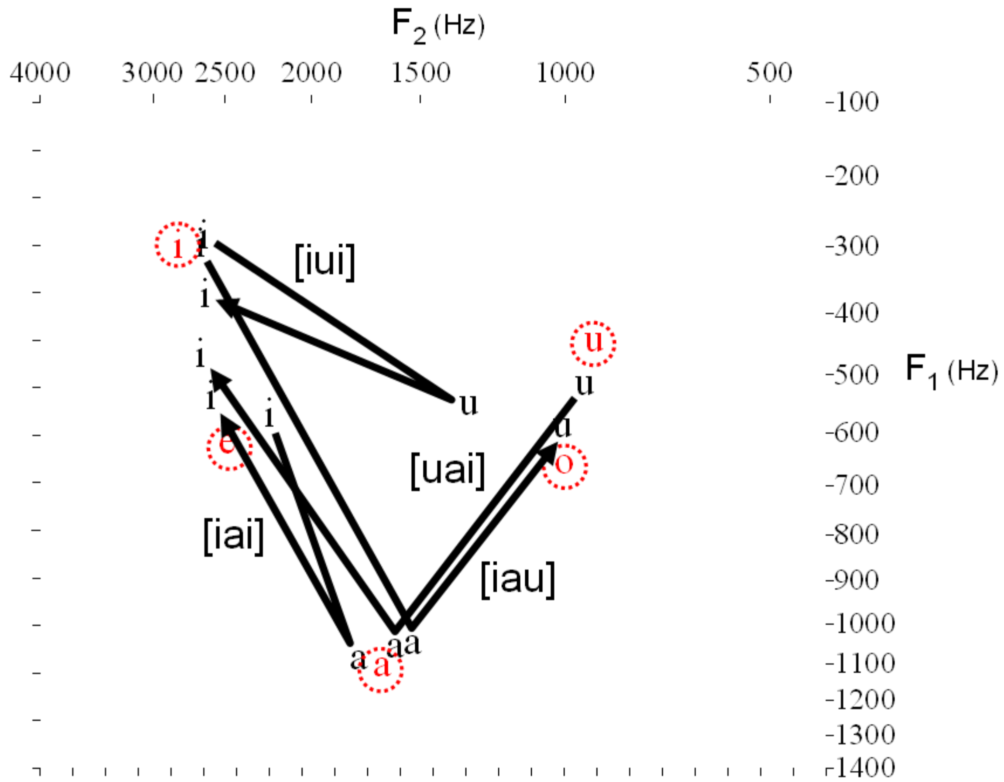


Figure 4.21 Formant movements for the triphthongs [iai, iui, iau, uai] and the positions of the 5 monophthongal vowels [i, e, a, o, u] in Meixian Hakka on the F1/F2 plane for 10 female speakers.

the male and female speakers in the formant pattern of particular elements of each triphthong, the overall formant patterns of the Meixian Hakka triphthongs on the F1/F2 plane for the male speakers are similar to those for the female speakers. For both male and female speakers, vowel undershoot occurs in every triphthong in Meixian Hakka but not necessarily occurs in each of the three elements of a triphthong. Results of the Meixian Hakka triphthongs show that vowel undershoot can be observed in the first and third elements of [iai], the second element of [iui], the first and third elements of [uai] and the third element of [iau]. Lindblom (1963) indicated that formant pattern of a vowel shifted away from its target frequencies as vowel duration reduced. He found that the phenomenon of vowel undershoot might result from a decrease in vowel duration. As syllables were produced in a short period of time, the articulators would be unable to move fast enough to reach the target positions of the vowels. In this study, vowel undershoot is found in [u] in [iui], [uai] and [iau]. Results show that the third element of [iau] and the first element of [uai] have higher F_1 and F_2 values than the monophthong [u]. Besides, the position of [u] in [iui] is close to the center of the acoustical vowel space and it is far from the position where the target vowel is. It appears that [u] in Meixian Hakka triphthongs is more central than the monophthong [u] on the F1/F2 plane. The acoustical positions of the third element of [iau] for both male and female speakers locate near to the position of the monophthongal vowel [o] rather than that of [u]. Therefore, acoustically, the third element of [iau] should be considered as [o]. The second element of [iui] has an F_1 value in between monophthongs [u] and [o], and an F_2 value that assimilates a central vowel. As shown in Figures 4.20 and 4.21, the acoustical positions of the second element of [iui] locate near to the centre of the acoustical vowel space. The second element of [iui] has higher F_2 and slight higher F_1 values than the monophthong [u]. Based on the acoustic data, the second element of [iui] should be treated as a centralized near-close near-back vowel [ʊ̈]. Such great deviation in formant pattern of the second element of [iui] may be due to the lack of time in speech production and it is supported by the temporal data in this study. As shown in Figures 4.22a and 4.23a, the duration of the second element of [iui] is extremely short for most

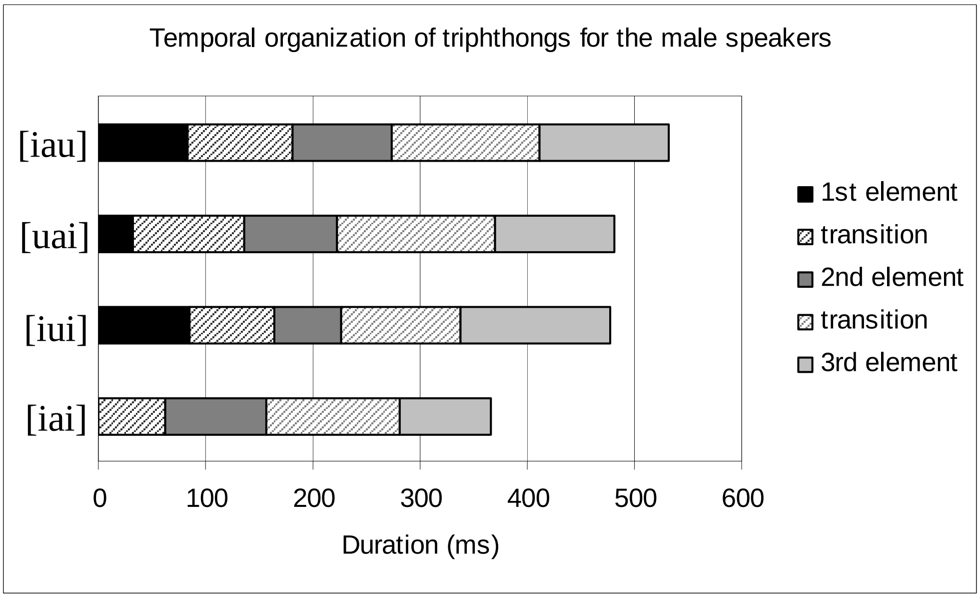


Figure 4.22 Temporal organizations of the 4 triphthongs [iau, uai, iui, iai] in Meixian Hakka for 10 male speakers

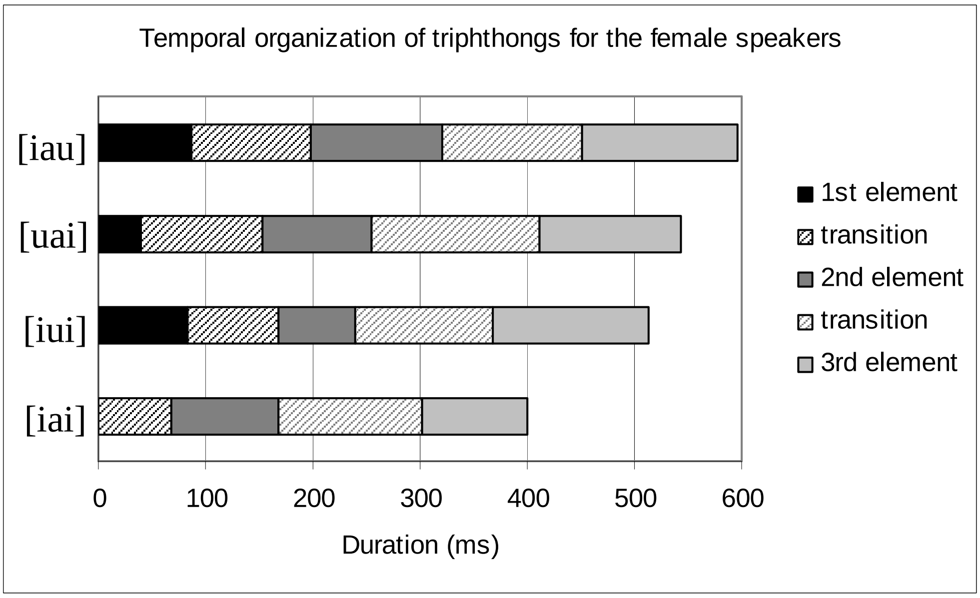


Figure 4.23 Temporal organizations of the 4 triphthongs [iau, uai, iui, iai] in Meixian Hakka for 10 female speakers

speakers in this study. It shows that throughout the production of a sound segment, the articulator does not have enough time to reach the sound target. For the first and third elements of [iai] and the third element of [uai], their acoustical positions are close to the position of the monophthongal vowel [e] rather than [i]. It may be justified to transcribe them acoustically as [e] instead of [i]. On the contrary, the first and third elements of [iui] and the first element of [iau] generally match the target of the monophthongal vowels [i]. For vowel [a] in Meixian Hakka triphthongs, [a] in [iai, iau, uai] has a slightly lower F_1 value than the value of the monophthong [a]. It shows that [a] in triphthongs is slightly centralized in the acoustical vowel space. However, such influence on [a] in the triphthongs is relatively small in comparison with that on vowels [i] and [u].

Figures 4.22 and 4.23 show the temporal organizations of Meixian Hakka triphthongs that produced by 10 male and 10 female speakers, respectively. In addition, Figures 4.22a and 4.23a exhibit the temporal organizations of the triphthongs for each individual male speaker as well as female speaker. The temporal data of the triphthongs show that, for both male and female speakers, there is no observable steady state for the first element in [iai]. The temporal data in Figures 4.22a and 4.23a further prove that the absence of steady state for the first element of [iai] is consistent for every individual speaker in this study. Taking the spectral and temporal data into consideration, it is justified to treat the first element of [iai] as a palatal glide [j] rather than a vowel [i]. It should be noted that, in Meixian Hakka, triphthong [iai] can only exist either in isolation or with a syllable-initial consonant [k]. In this study, [kiai] was used as the test word. It may be questionable whether the first element of [iai] is under the effect of a palatalized [k]. A supplementary study has been conducted and the recordings of [iai] in isolation were analyzed. The formant trajectories of [iai] that produced by one male and

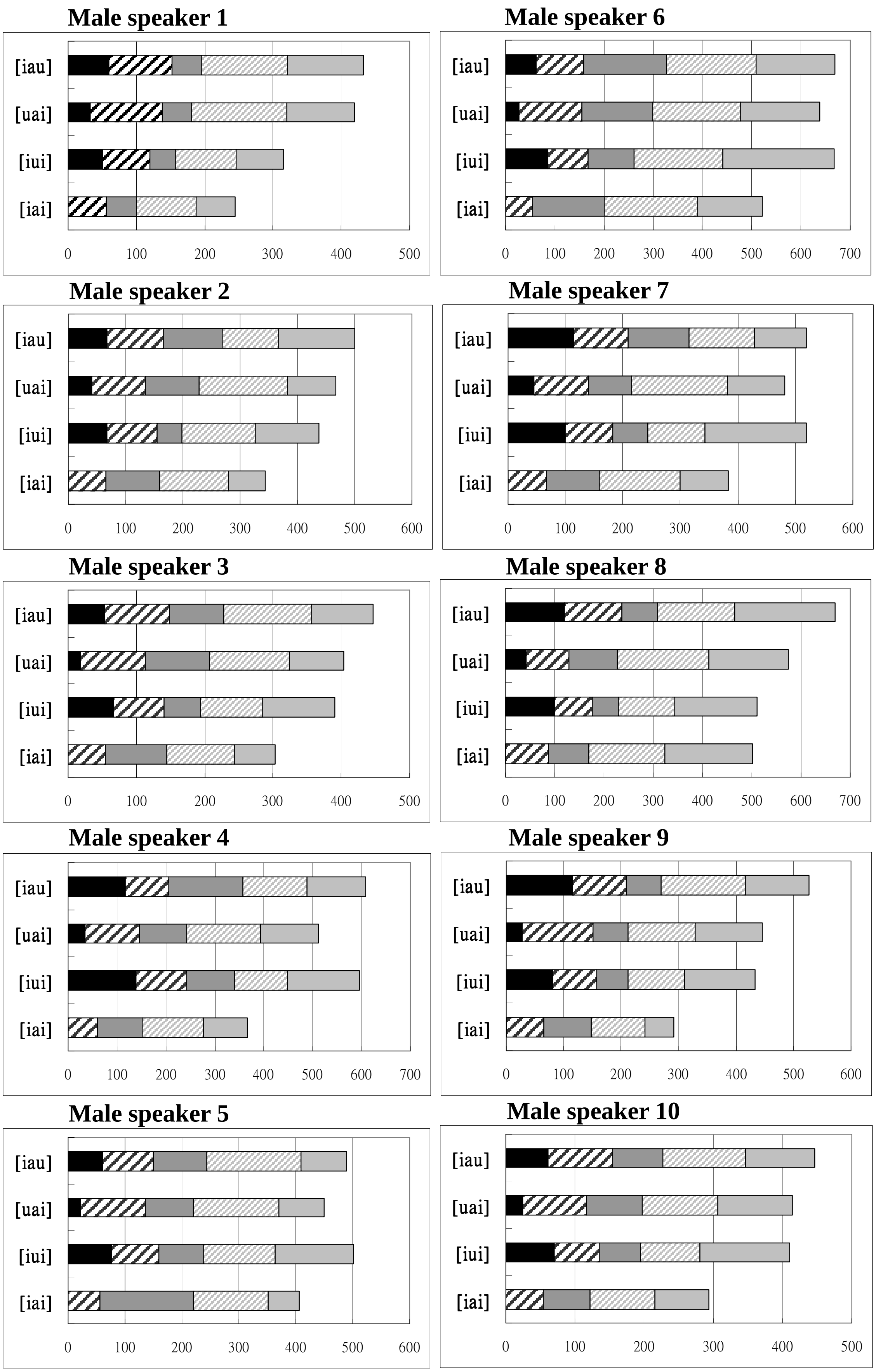


Figure 4.22a Temporal organizations of the 4 triphthongs [iau, uai, iui, iai] in Meixian Hakka that produced by the 10 male speakers

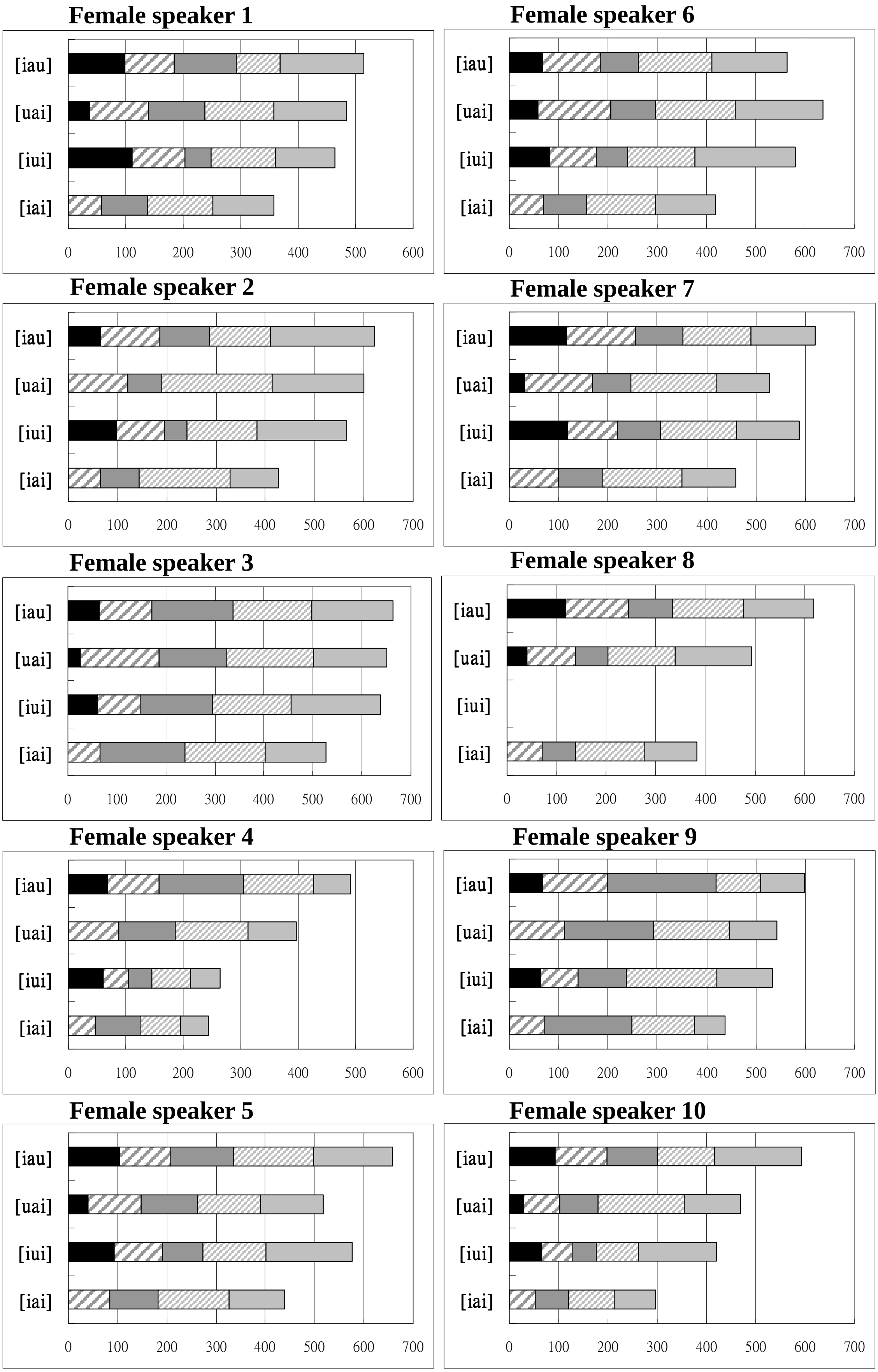


Figure 4.23a Temporal organizations of the 4 triphthongs [iau, uai, iui, iai] in Meixian Hakka that produced by the 10 female speakers

two female speakers are displayed in Figure 4.24. As described in the study of Huang (1995), there are two pronunciations for the word, which means ‘push’, in Meixian Hakka. The IPA transcriptions of the two pronunciations are [iai33] and [ai33], respectively. In this study, only one of the ten male and two of the ten female speakers accepted that the word has two alternative pronunciations, whereas for the other speakers, [ai33] is the only possible pronunciation for the word. The formant trajectories of [iai] that produced in isolation show that there are observable steady states for the first element of [iai]. It is different from [iai] in CV syllables of which the steady state of the first element is absent. After the steady state of the first element, there is an increase in F_1 and a decrease in F_2 to form a transition in between the first and second elements of [iai]. Results indicate that when triphthong [iai] is preceded by a syllable-initial stop [k], the steady state of the first element of [iai] would be lost. It further proves that the formant transition at the beginning of [iai] in CV syllables is the acoustic characteristic of a triphthong rather than the influence of a palatalized [k^j]. In addition, for both male and female speakers, triphthong [uai] only has a brief or no steady state for the first element. The temporal data of female speakers 2, 4 and 9 show that there is no steady state for the first element of [uai] (Figure 4.23a). It may be rational to transcribe the first element of [uai] acoustically as a labiovelar glide [w] due to the short duration or absence of steady state and the occurrence of formant undershoot. In Meixian Hakka, [uai] can only exist in CV syllables with a syllable-initial consonant of [k] or [k^h]. Therefore, it is proposed in this study to treat triphthong [uai] as a diphthong [ai] which is preceded by a labiovelar consonant [k^w] or [k^{wh}]. Though there is only one test word for [iui] in Meixian Hakka, both the spectral and temporal data of [iui] show that [iui] is a triphthong of which each element has an apparent steady state and the formants of the first and second elements which are transcribed as [i] reach the target of

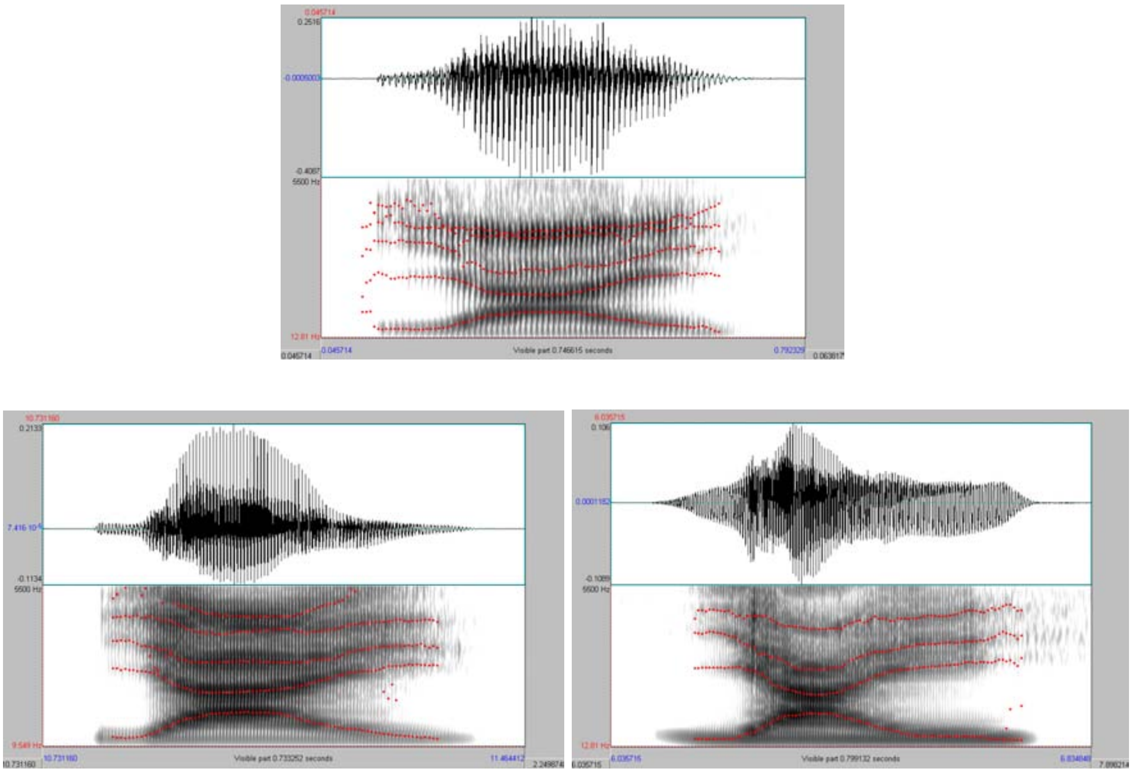


Figure 4.24 Formant trajectories for [iai] ‘push’ in Meixian Hakka for a male speaker (the upper graph) and two female speakers (the lower graphs)

monophthong [i]. The test word of triphthong [iui] in Meixian Hakka is [iui51] 銳 which means ‘sharp’ and it is supposed that it is one of the basic words in the Chinese language. Unlike other triphthongs, [iau] can occur in various phonological environments. The temporal data show that there is a clear steady state for each element of triphthong [iau]. Comparing the results of [iau] and [uai] in this study with the data of Beijing Mandarin (Zee, 2001), the centralization of the second element [a] in Beijing Mandarin is much greater than that in Meixian Hakka. In conclusion, the spectral data and the temporal data in this study show that there are only two triphthongs in Meixian Hakka and they are [iui] and [iau]. It is proposed that the so-called triphthong [iai] should be considered as palatal glide [j] plus diphthong [ai], whereas in consideration of the phonological constraint of the vowel, the so-called triphthong [uai] should be treated as diphthong [ai] which is preceded by either a labialized stop [k^w] or [k^{wh}].

4.4 Discussion

In the traditional understanding, vowel is characterized by a static articulatory configuration and the frequencies of the first three formants of the steady-state portion. Experimental data further show that vowels can be highly identifiable in terms of the frequencies of the first two formants only (Delattre et al., 1952, Peterson & Barney, 1952). Acoustic data of different languages such as American English, Banglam, Greek, Italian and Yoruba also displayed that the frequencies of F_1 and F_2 can be sufficient to distinguish among vowels in a sound system (Fourakis et al., 1999; Hawkins & Midgley, 2005; Hossain et al., 2007; Ladefoged & Maddieson, 1990). Meixian Hakka is a five-vowel system and the five vowels are [i, e, a, o, u]. The results of this study are in agreement to the claim that the first two formants are the most important for the classification of vowels. The vowel ellipses in Meixian Hakka show that there are five non-overlapping spectral categories indicating five distinctive vowels. Such results are consistent for both male and female speakers. Duration seems not to be an important cue for the distinction of Meixian Hakka vowels.

Two approaches have been proposed to explain the cross-linguistic tendencies regarding the phonetic structure of vowel systems. On one hand, Liljencrants and Lindblom (1972) first attempted to define the vowel boundaries within the acoustic space universally and the principles of maximal perceptual contrast and minimal articulatory effort were involved. The Dispersion Theory predicted that the distribution of vowels in natural language is closely related to the size of vowel system. Within the available acoustic space, vowel categories disperse in order to maximize the distance from its neighbours (Liljencrants & Lindblom, 1972; Lindblom, 1986). The vowel system with an increase in the number of point vowel should result in the expansion of overall vowel system. Under this theory, it suggests that the locations of point vowels within the acoustic space should be consistent for languages having similar phonemic vowel categories. On the other hand, Quantal Theory proposed by Steven (1972; 1989) has taken the correlation between articulation and acoustic output into account. This theory proposes that the locations of vowel categories in the acoustic space can be explained as an account of the non-monotonic relation between variation in the articulatory configuration and its corresponding acoustic output. In other words, there are certain regions of the vocal tract that a relatively great variation in articulatory configurations does not result in the expected variation in acoustic consequences. Instead, within the certain regions, the acoustic consequences of the point vowels are relatively stable and they are perceived as an identical point vowel in the acoustic space. As articulatory configuration of vocal tract should be similar across languages, this theory predicted that the stability regions of acoustics should be approximately the same across all languages. That is, it was proposed that the locations of the point vowels such as /i/, /a/, and /u/ within the phonetic space should be consistent in languages of the world. Five-vowel system is particularly favored in the world languages and Meixian Hakka is one of them.

This paper here will test the two theories presented above for the vowel system of Meixian Hakka. It is supposed that, on one hand, if the results of this study are in agreement to the Dispersion Theory, the locations of the five Meixian Hakka vowels in the acoustic space should be approximately identical to the locations of the vowels in other languages with a five-vowel system. On the other hand, if the results agree with the Quantal Theory, the locations of the point vowels in Meixian Hakka should be similar to those in other languages of the world regardless of the size of the vowel inventory. With regard to the Dispersion Theory, the results of this study in comparison to the acoustic data of different languages such as Greek (Fourakis et al., 1999), Hebrew

(Most et al., 2000) and Spanish (Guirao & Manrique, 1975) indicate that the distribution of vowels within the acoustic space for different five-vowel systems is not identical and it seems not to support the prediction of the Dispersion Theory. For examples, though the locations of [i, e, o] in this study are similar to those of the vowels in Greek and Hebrew, it shows that the [u] and [a] in Meixian Hakka are at lower vertical positions, i.e., higher F_1, than those in Greek and Hebrew. In contrast, [i] and [u] in Spanish are at lower positions in comparison to those in the other three languages. Moreover, if Dispersion Theory holds true, the expansion of the vowel system for the four languages with five point vowels should be more or less the same. However, it shows that the areas of F1 x F2 space for the four languages vary in progression Meixian Hakka > Spanish > Greek ≈ Hebrew. Contradictory evidence for the Dispersion Theory was also stated in the study of Bradlow (1995). In the cross-linguistic study of Bradlow, it has been stated that the expansions of vowels are similar for English, Spanish and Greek, even though the size of English vowel system with 11 vowels is much larger than that of Spanish and Greek systems with 5 vowels only. This indicates that the boundaries of vowel categories in the vowel space should be defined on a language-specific basis in spite of the structure of the vowel inventory. Quantal Theory predicted that the point vowels of languages in the world should occupy similar positions in the acoustic space and the regions of stability should be consistent across languages. Cross-linguistic comparison of Meixian Hakka with other languages provides the indication that there is a certain degree of variation in the locations of the five vowels in the acoustic space. For examples, [o] in Spanish has an F_1 value that is similar to that of [a] in Greek and Hebrew. In addition to the lower positions of [i] and [u] in Spanish, it shows that Spanish vowels have a more open articulatory configuration than Meixian Hakka, Greek and Hebrew. The phonetic differences between similar segments of different languages with identical phonemic system imply that the data in Meixian Hakka do not support the claim of certain acoustically stable regions in the Quantal Theory.

The spectral characteristics of diphthongs propose that diphthongs in Meixian Hakka can be separated into two categories. Traditional classification of diphthongs indicates that diphthongs can be phonologically classified into two categories, namely, falling (offgliding) diphthongs and rising (ongliding) diphthongs. This classification of diphthongs is related to the variation in prominence of the two elements of a diphthong. In general speaking, falling diphthong starts with a more prominence syllabic nucleus and ends with a less prominence high vowel /i, u/ or semivowel (e.g., /ai/, /au/). In contrast, rising diphthong starts with a less prominence high vowel or semivowel and ends with a more prominence syllabic nucleus (e.g., /ia/, /ua/).). It has been long recognized that the prominence of a syllable nucleus is determined by three factors, namely length, relative pitch and loudness, of which the acoustic correlates are duration, fundamental frequency and intensity, respectively. With respect to the Chinese dialects, Hsu (2008) declared that, acoustically, duration rather than intensity of a syllable is a more reliable indicator of prominence. In addition to the spectral properties, the classification of Meixian Hakka diphthongs is also supported by the temporal relationship among onset, offset and transition of a diphthong. Results of this study show that acoustically there are in total 9 diphthongs in Meixian Hakka which are [ie, ia, io, iu, aɪ, ɔɪ, ʊi, ao, ɛʊ]. [ie, ia, io] are characterized as rising diphthongs, whereas [aɪ, ɔɪ, ao, ɛʊ] are characterized as falling diphthongs. In this study, it may be more appropriate to treat [ʊi] and [iu] separately as both elements in the diphthongs are high and others factors such as temporal organization has to be considered as well. In consideration of the spectral properties, the two elements of rising diphthongs in Meixian Hakka have F1/F2 patterns that resemble the formants of target vowel in monophthong. The formant frequencies of the two elements of falling diphthongs, however, drift away from the locations of the corresponding targets in the acoustic space. The F_1 values for all vowel except [a] increase in falling diphthongs.

The temporal organization of diphthongs also supports the categorization of Meixian Hakka diphthongs. The overall durations of rising diphthongs are relatively longer than those of falling diphthongs in most cases for all speakers. In addition, rising diphthong in Meixian Hakka has a longer steady-state for the second element than the falling diphthong, while falling diphthong has a longer transition in between the two elements than the rising diphthong. Huang (1995) stated that the onset and offset elements of [iu] in Meixian Hakka had similar duration. Results of this study show that [iu] is one of the longest diphthong in Meixian Hakka. It composes of two elements with similar duration and a varying transition in between. It should be noted that the durations of the two elements are not consistently similar. For most of the speakers, the duration of the first element is slightly longer than that of the second element. Statistic analysis further shows that the onset of [iu] has a duration which is significantly greater than the offset for both male [$t = -6.545$, $p > .00$] and female speakers [$t = 3.782$, $p > .00$].

This study has also investigated the dynamic properties of diphthongs in Meixian Hakka and results show that F_2 range and F_2 rate of change are crucial for the classification of diphthongs in Meixian Hakka. Chan and Ren (1988) have investigated the dynamic properties of diphthongs in Shanghai, Cantonese and Mandarin. The findings on the three Chinese dialects of their study are found to be very similar to the results of this study. In their study, it was declared that the F_2 rate of change in the three dialects corresponded to the classification of diphthongs and ongliding diphthongs tend to have a faster F_2 transition rate than their offgliding counterpart. Results of Meixian

Hakka diphthongs in this study support the generalization in their study. For examples, [ia], which starts with a high vowel, has a faster rate of change than [ai], which ends with a high vowel. The findings of the three Chinese dialects in Chan and Ren's study were also compared with the findings of English (Gay, 1968), Spanish (Manrique, 1979) and Maithili (Jha, 1985). Results of this study indicate that the generalization that ongliding diphthong has a faster rate of change than its offgliding counterpart seems to be language-universal rather than language-specific. In addition, it is proposed that there is another language-universal characteristic of diphthongs with respect to the relevance of vowel height and backness to the F_2 rate of change. The acoustic data of this study show that diphthong that involves large movement along the front-back axis tends to have the fastest rate of change. On the contrary, diphthong that involves small movement along the high-low axis tends to have the slowest rate of change. Such findings are also observed in the three Chinese dialects (Chan & Ren, 1988) as well as Spanish (Manrique, 1979) and English (Gay, 1968). Last but not the least, the correlation between F_2 range and F_2 rate of change is positive for the Meixian Hakka diphthongs. Kent and Moll (1972) proposed that the transition duration was relatively constant and the rate of change of F_2 transition should correspond to the acoustic distance between the onset and offset of a diphthong. In other words, diphthong of which the two elements in greater distance was expected to have a faster rate of change, whereas diphthong of which the two elements in shorter distance should have a slower transition rate. This proposal is supported by the results of this study as well as the findings of Shanghai, Cantonese, English, Maithili and Spanish as described in the study of Chan and Ren (1988). Furthermore, results show that there are similarities on the dynamic nature of diphthongs between Meixian Hakka and Spanish. For the majority of Meixian Hakka diphthongs, there are apparent steady states for both onset and offset of a diphthong. Similarly, prominent steady states for both onset and offset were found in Spanish diphthongs. With respect to the F_2 range and F_2 rate of change, both Spanish and Meixian Hakka diphthongs involve a large range of F_2 transition rate and there is a positive correlation between F_2 range and F_2 rate of change for the two languages. The F_2 rate of change for the male speakers ranges from 2.3 Hz/ms to 16.6 Hz/ms for Spanish diphthongs (Chan & Ren, 1988) and it ranges from 1.2 Hz/ms to 11.6 Hz/ms for Meixian Hakka diphthongs. It is proposed that the similarity of dynamic nature of diphthongs between the two languages may due to the fact that both languages have a large inventory of diphthongs in comparison to languages with a small inventory such as Mandarin.

CHAPTER FIVE

CONCLUSION

The research described in this dissertation investigates tones and vowels in Meixian Hakka from acoustic and perceptual perspectives. It is worthwhile to research from acoustic and perceptual perspectives as no or little relevant data have been presented in the literature. It is believed that results of this study would provide valuable information regarding vowels and tones in Meixian Hakka for further research and some insights into the correlation between acoustic data and tone perception.

The acoustic characteristics of tones in isolation and in different phonetic contexts were investigated. It was found that there are contemporary six citation tones in Meixian Hakka which correspond to the six historical tonal categories. Results showed that Yin Ping is acoustically a level tone at the middle frequency range; Yang Ping is a slight falling tone at the low frequency range; Shang Sheng is a simple falling tone from the middle frequency range to the lowest part of an individual's frequency range; Qu Sheng is a delayed falling tone from the high frequency range to the low frequency range; Yin Ru is a short tone with a falling contour that is similar to the F_0 contour of Shang Sheng; and Yang Ru is a short tone with a steady contour at the high frequency range. Normalization strategy was employed to extract acoustic information that is linguistically significant. Pitch values were then assigned to the six citation tones based on the patterns of normalized F_0 contour. Yin Ping is transcribed as the mid level tone with a pitch value of [33]; Yang Ping is transcribed as the low level tone with a pitch value of [11]; Shang Sheng is transcribed as the mid-high to low falling tone with a pitch value of [41]; Qu Sheng is transcribed as the high to low falling tone with a pitch value of [51]; Yin Ru is transcribed as the mid-high to low level short tone with a pitch value of [41]; Yang Ru is transcribed as the high level short tone with a pitch value of [55] or [5]. The non-Ru tones in Mexian Hakka are shown to be significantly longer than the Ru tones. The comparison of F_0 contour between Shang Sheng and Yin Ru showed that the overall contour shapes of the two tones are similar and both tones have a mid-high to low falling contour. The two tones distinguish from one another as the duration of Shang Sheng is greater than that of Yin Ru.

In addition to the acoustic properties of tones in isolation, the F_0 contours of Meixian Hakka tones in different phonetic contexts were investigated. The comparison of F_0 between vowels [i, a, u] showed that iF0 differences were found in Meixian Hakka for all tones. Tones on high vowels [i, u] tend to have a higher F_0 than tones on low vowel [a]. With support of the results of this study, such correlation between F_0 and vowel height seems to be universal as it was found in other contour tone languages, register tone languages, and even non-tone languages from different language families (Walen & Levitt, 1995; Zhu, 1999; Connell, 2002; Gonzales, 2009). Results of this study further showed that patterns of iF0 may vary across tones in accordance with the contour shape. The iF0 differences between high and low vowels exist over the whole duration of the F_0 contour for tones with a level contour. For tones with a falling contour, the iF0 differences are conspicuous in the first half of F_0 contour. Such differences decrease gradually and may finally diminish towards the end of the fall. This finding is not only found in Meixian Hakka, but also can be observed in the study of Shanghai tones (Zhu, 1999).

In this study, the citation tones in CV(C) syllables were also compared with the citation tones in V(C) syllables. Results exhibited that the mean durations of tones in V(C) syllables are significantly greater than those in CV(C) syllables. In consideration of contour shape and tonal height, the effect of neighboring consonant on the F_0 onset of the following tone was observed. The F_0 onsets of tones after voiceless consonants in CV(C) syllables are higher in pitch than the F_0 onsets of tones in V(C) syllables and a fall in F_0 tends to occur at the beginning of the tones in CV(C) syllables. This finding holds true across the six citation tones in Meixian Hakka. There is a tendency for all citation tones except Yin Ping [33] in CV(C) syllables to start at a frequency which is higher than that of the corresponding tones in V(C) syllables. The difference in F_0 between tones in CV(C) syllables and V(C) syllables decreases throughout the overall duration. The two F_0 curves intersect at some point and then the F_0 curve of tones in CV(C) syllables is lower than that in V(C) syllables.

Discrepancies in the descriptions of tone sandhi patterns were found between the past studies. The acoustic realization of the citation tones in bisyllabic words and tone sandhi patterns of bisyllabic combinations in Meixian Hakka were investigated. It showed that tone sandhi appears in a two-syllable structure and the citation form of tone on the first syllable changes to its sandhi form when the sandhi environment is fulfilled. Results indicated that there are three tone sandhi rules in Meixian Hakka. Only the mid level tone [33], the mid-high to low falling tone [41] and the high to low falling tone [51] may undergo tone sandhi. The sandhi environments for [33] and [41] are the same and the tone changes to its sandhi form when it is followed by [11, 41, <u>41</u>]. On the other hand, the sandhi environment for [51] is similar to that for [33] and [41] and the tone is realized as its sandhi form when it is followed by [11, 41, <u>41</u>, 51]. Two sandhi forms that are different from the six citation tones in Meixian Hakka were observed. The high level tone [55] is the sandhi form of [51], whereas the mid to high rising tone [35] is the sandhi form of [33]. The tone sandhi patterns observed in this study are consistent with the patterns of Lü (2006), but not with the tone sandhi patterns of Hashimoto (1973), Huang (1995) and Yuan et al. (2001). It shows that the discrepancies between the past studies may due to the lack of sufficient speakers in the study of Hashimoto (1973) and the impressionistic analysis of Huang (1995) and Yuan et al. (2001).

Based on the acoustic data of tones obtained in the study, a perceptual experiment was conducted to investigate the perceptual properties of the four long citation tones in Meixian Hakka and to examine the potency of acoustic properties for tone identification. The perceptual results showed that there is a strong correlation between the perception and the acoustic characteristics of the four Meixian Hakka tones. Tonal height is the potent cue for the identification of the two level tones [33, 11] which are acoustically distinct from each other in frequency range. It showed that a level tone is perceived by the Meixian Hakka speakers when the F_0 trajectory of stimuli is either level or falling with a drop of less than 30 Hz. The potent cues for the distinction between [51] and [41] were found to be the tonal height and the slope. On one hand, the identification of [51] displayed that a high F_0 onset value may provide sufficient cue for the distinction between [51] and other tones. The value of F_0 offset is rather insignificant for the identification of [51]. The potency of tonal height seems to be correlated with the acoustic characteristics of [51] as the highest point of the F_0 contour of [51] is acoustically at the highest part of an individual's frequency range and its value is the highest in comparison with that of other tones [33, 11, 41].

In addition to tonal height, slope was also shown to be the potent cue for the identification of [51, 41, 11]. Results displayed that for stimuli with a relatively low F_0 offset value, [51] is identified when there is a steep falling slope of the contour, [41] is identified when the falling slope is less steep, and [11] is identified when the F_0 contour has a slight falling slope. The potency of slope for the distinction between the three tones [51, 41, 11] relates closely to the difference in F_0 onset between the three tones. The F_0 contours of [51, 41, 11] are acoustically falling. The F_0 offsets of the three tones are all at the low frequency range and the three tones are distinct from each other in F_0 onset value. The steep falling slope of [51] results from the high F_0 onset of the contour at the high frequency range. In comparison, the falling slope of [41] is less steep as the F_0 onset of [41] is at the middle frequency range. The slope of [11] is very gentle as the F_0 onset and offset of the contour are both at the low frequency range. The perceptual results further indicated that the perceptual matrix of the Meixian Hakka tones is perceptually divided into three levels in terms of F_0 onset value, which are high, mid and low, and into two levels in terms of F_0 offset value, which are high and low. It showed that the perceptual patterns of the four long tones in the perceptual space generally match the distribution of the tones in the acoustic tone space.

The spectral characteristics and temporal organization of monophthongs, diphthongs and triphthongs were investigated and discussed in chapter four. Acoustic data suggested that there are 5 vowels, 1 apical vowel, 9 diphthongs and 2 triphthongs in the sound system of Meixian Hakka. The apical vowel [ɿ] in Meixian Hakka is a high central vowel with F_1 at the low frequency and F_2 at the middle frequency range of the acoustical vowel space. Based on the formant patterns of vowels, the 5 monophthongs are transcribed as [i, e, a, o, u]. The transcription of vowels in this study based on the acoustic data is consistent with the phonological transcription described in the past studies. The difference in distribution of vowels in the acoustical vowel space is found between front and back vowels. Results show that the distance between high and mid vowels is larger for the front vowels than that for the back vowels. That is to say, the distance between [i] and [e] on the F1/F2 plane is much greater than that between [u] and [o]. This seems to be resulted from the relatively low F_1 of [u] in comparison to the F_1 of [i]. Even though [i] and [u] are both classified as high vowels in Meixian Hakka, the position of [u] on the acoustic space in terms of F_1 is significantly lower than that of

[i].

It is suggested that the diphthongs in Meixian Hakka in general can be divided into two categories according to the spectral characteristics and temporal organization. The rising diphthongs [ie, ia, io] in Meixian Hakka distinguish from the falling diphthongs [ai, oi, au, eu, ui] with regard to their relative distance from the target monophthongs on the F1/F2 plane. Acoustic results show that the positions of the two elements of the rising diphthongs resemble those of the corresponding monophthongs, whereas the positions of the two elements of the falling diphthongs drift away from the positions of the corresponding monophthongs. In comparison to formant patterns of the corresponding monophthongs, F_1 values increase and F_2 values drift towards the centre of the acoustic space for the two elements of the falling diphthongs. However, such difference in formant pattern between monophthong and vowels of a falling diphthong is not found in the case of low vowel [a]. The distinction between the two types of diphthongs is also supported by the temporal organization. The overall duration of the falling diphthongs is generally shorter than that of the rising diphthongs for all speakers. By comparing the duration of the two elements of a diphthong, the steady-state duration of the second element of the rising diphthongs is much longer than that of the first element, whereas the duration of the first element of the falling diphthongs is relatively longer than that of the second element. In addition, the transition duration of the falling diphthongs is longer than that of the rising diphthongs for most speakers. The case of diphthongs [iu] in Meixian Hakka is complicated and it is difficult to be categorized into either one of the two types of diphthongs. In terms of formant pattern, [iu] is similar to the pattern of rising diphthongs and the two elements of the diphthong resemble the corresponding monophthongs. However, in terms of the temporal organization, [iu] is similar to the pattern of falling diphthongs and duration of the second element of the diphthong is much shorter than that of the rising diphthongs.

Besides, F_2 range and F_2 rate of change have been examined in this study. Results show that the dynamic properties of diphthongs are optimal to differentiate among diphthongs in Meixian Hakka. Diphthongs end with a more prominent vowel (e.g. [ia]) tend to have a faster F_2 rate of change than its mirror-image counterpart (e.g. [ai]). F_2 rate of change is also shown to be determined by the degree of formant movement between the two elements of a diphthong along the front-back axis and high-low axis of the F1/F2 plane. The comparison of F_2 rate of change between Meixian Hakka diphthongs and diphthongs of other languages indicates that the dynamic nature of diphthongs in terms of F_2 rate of change seems to be language-universal. It is concluded that there are 9 diphthongs in Meixian Hakka and their narrow transcriptions are [ie, ia, io, iu, aɪ, ɔɪ, ʊi, ao, ɛʊ] based on the acoustic data in this study. Regarding triphthongs in Meixian Hakka, it shows that vowel undershoot occurs in all triphthongs but not in all elements of a triphthong. Acoustic and production data suggest that there are two triphthongs in Meixian Hakka and they are phonetically transcribed as [iʊ̈i] and [iao].

Although the findings in the present study enable us to have a better understanding on the acoustic properties of vowels and tones as well as the correlation between perceptual responses and acoustic characteristics of tones in Meixian Hakka, the acoustic and perceptual data in this study are limited to only the elderly. The old age group is chosen in the present study as the native speakers in the old age group appear to have least influence from Mandarin Chinese, which is the official language of China, and it is assumed that the elders may retain the original dialect of Meixian Hakka. Moreover, the data of the elderly in this study may be directly comparable to the descriptions of the past studies. The limitations of this study are the limited number of subjects and the lack of concentration of the elderly. The number of subjects is limited as many old people were not willing to participate in the experiment. Besides, more time was needed to give instructions and do the recordings as the elders tend to be less able to concentrate on a task. Regardless of the limited data of the elderly, it is believed that this study may provide useful information about the acoustic and perceptual properties of tones and vowels in Meixian Hakka. Further cross-linguistic investigation remains to be done to study the difference in acoustics of the Meixian Hakka sound system between the old and young age group under the influence of Mandarin Chinese.

REFERENCES

Ainsworth, W. A. (1972). Duration as a cue in the recognition of synthetic vowels. *Journal of the Acoustical Society of America*, 51, 648-651.

Anderson, S.R. (1978). Tone features. In V.A. Fromkin (ed.), *Tone: a linguistic survey*. New York: Academic Press, 133-176.

Arlund, P. S. (2006). *An acoustic, historical, and developmental analysis of Sarikol Tajik diphthongs*. Doctoral dissertation, University of Texas at Arlington, Arlington, Texas, US. Retrieved December 17, 2008 from http://hdl .handle.net/10106/438

Bennett, D. C. (1968). Spectral form and duration as cues in the recognition of English and German vowels. *Language and Speech*, 11, 65-85.

Blevins, J. (1995). The syllable in phonological theory. In J. A. Goldsrnith (ed.), *The Handbook of Phonological Theory*. Carnbridge, MS: Blackwell, 206-244.

Bond, Z. S. (1978). The effects of varying glide durations on diphthong identification. *Language and Speech*, 21, 253-263.

Bradlow, A. (1995). A comparative study of English and Spanish vowels. *Journal of the Acoustical Society of America*, 97, 1916-1924.

Catford, J. C. (1977). *Fundamental problems in phonetics*. Edinburgh: Edinburgh University Press.

Chan, M. K. M., and Ren, H. (1988, December). Acoustic analysis of the diphthongs in Shanghai, Cantonese and Mandarin. Paper presented at *International Conference on Wu Dialects*, Chinese University of Hong Kong, Hong Kong.

Chao, Y. R. (1930). A system of “tone letter”. *Le Maittre Phonetique*, 30, 24 – 27. (Reprinted in《*方言*》, 2, 281-283, 1980)

Chao, Y. R. (1948). *Mandarin Primer: An intensive course in spoken Chinese*. Cambridge: Harvard University Press.

Chen, M. Y. (2000). *Tone Sandhi: Patterns across Chinese dialects*. Cambridge University Press.

Chen, X. (陈修). (1993).《*梅县客方言研究*》. Meizhou: Jinan Daxue Chubanshe.

Cheung, Y. M. (张玉敏). (2006). 梅县客家话连读变调的语音分析. *The Proceedings of the 7th Phonetic Conference of China and International Forum on Phonetic Frontiers*, 1-5. Beijing: Peking University.

Collier, R., and ‘t Hart, H. (1983). The perceptual relevance of the formant trajectories in Dutch diphthongs. In M. van den Broecke, V. van Heuven and W. Zonneveld (eds.), *Sound structures: Studies for Antonie Cohen*. Dordrecht: Foris, 31-45.

Collier, R., and Raphael, L. (1982). Some acoustic and physiological observations on diphthongs. *Language and Speech*, 25(4), 305-323.

Connell, B. (2002). Tone languages and the universality of intrinsic F0: evidence from Africa. *Journal of Phonetics*, 30, 101-129.

Delattre, P., Liberman, A. M., Cooper, F. S., and Gerstman, L. (1952). An experimental study of the acoustic determinants of vowel color: observations on one- and two-formant vowels synthesized from spectrographic patterns. *Word*, 8(3), 195-210.

Disner, S. F. (1980). Evaluation of vowel normalization procedures. *Journal of Acoustical Society of America*, 67(1), 256–261.

Dofs, E. (2008). *Onomatopoeia and iconicity: A comparative study of English and Swedish animal sound*. Retrieved December 13, 2008 from http://urn.kb.se/resolve?urn=urn:nbn:se:kau:diva-1746

Erickson, M. L. (2000). Simultaneous effects on vowel duration in American English: A covariance structure modeling approach. *Journal of the Acoustical Society of America*, 108, 2980-2995.

Ericsdotter, C. and Ericssibm, A. M. (2001). Gender differences in vowel duration in read Swedish: Preliminary results. T*he Proceedings of Fonetik 2001, the XIVth Swedish Phonetics Conference. Working Papers of the Department of Linguistics, Lund University*, 49, 34-37.

Fourakis, M., Botinis, A., and Katsaiti, M. (1999). Acoustic characteristics of Greek vowels. *Phonetica*, 56, 28-43.

Francis, A. L., Ciocca, V., Wong, K. M., Chan, K. L. (2006). Is fundamental frequency a cue to aspiration in initial stops? *Journal of the Acoustical Society of America*, 120 (5), 2884-2895.

Fujisaki, H. (1983). Dynamic characteristics of voice fundamental frequency in speech and singing. In P. F. MacNeilage (ed.), *The Production of Speech.* New York: Springer-Verlag, 39-55.

Gandour, J. (1974). Consonant types and tone in Siamese. *Journal of Phonetics*, 2, 337-350.

Gandour, J. (1983). Tone perception in Far Eastern languages. *Journal of Phonetics*, 11, 149-175.

Gandour, J., and Harshmann, R. (1978). Crosslanguage differences in tone perception: a multidimensional scaling investigation. *Language and Speech*, 21, 1-33.

Gao, Y. (高永安). (2007). 驻马店方言象声词.《*北京广播电视大学学报*》, 2007 年第 4 期, 50-53.

Gay, T. (1968). Effect of speaking rate on diphthong formant movements. *Journal of the Acoustical Society of America*, 44, 1570-1573.

Gay, T. (1970). A perceptual study of American English diphthongs. *Language and Speech*, 13, 65-88.

Gleason, H. A. (1961). *An introduction to descriptive linguistics* (revised ed.). New York: Holt, Rinehart and Winston.

Gonzales, A. (2009). Intrinsic F0 in Shona vowels: A descriptive study. *Selected Proceedings of the 39th Annual Conference on African Linguistics: Linguistic Research and Languages in Africa*, 145-155. Retrieved June 30, 2010 from www.lingref.com/cpp/acal/39/paper2194.pdf

Gordon, M., Barthmaier, P., and Sands, K. (2002). A cross-linguistic acoustic study of voiceless fricatives. *Journal of the International Phonetic Association*, 32, 141-174.

Gottfried, M., Miller, J. D., and Meyer, D. J. (1993). Three approaches to the classification of American English diphthongs. *Journal of Phonetics*, 21, 205-229.

Guirao, M., and Manrique, A. M. B. (1975). Identification of Argentine Spanish vowels. *Journal of Psycholinguistic Research*, 4(1), 17-25.

Guo, J. (郭锦桴). (1993). *《汉语声调语调阐要与探索》*. Beijing: Beijing Language and Culture University Press.

Han, M. S. and Weitzman, R. S. (1970). Acoustic Features of Korean /P, T, K/, /p, t, k/ and /p^h, t^h, k^h/. *Phonetica*, 22, 112-128.

Hamann, S., and Sennema, A. (2005). Acoustic differences between German and Dutch labiodentals. In C. Geng, J. Brunner, & D. Pape (eds.), *ZAS Papers in Linguistics*, 42, 33-41.

Hamdi, R., Ghazali, S., and Barkat-Defradas, M. (2005) Syllable structure in spoken Arabic: a comparative investigation. *Proceedings of the INTERSPEECH 2005*, 2245-2248.

Hanson, H. M. (2009). Effects of obstruent consonants on fundamental frequency at vowel onset in English. *Journal of the Acoustical Society of America*, 125(1), 425–441.

Hanyu Fangyin Zihui (2nd ed.)《*汉语方言字汇(第二版)*》. (1989). Beijing: Wenzi Gaige Chubanshe.

Hashimoto, M. (1973). *The Hakka Dialect: A linguistic study of its phonology, syntax, and lexicon*. Cambridge, England: Cambridge University Press.

Hawkins, S., and Midgley, J. (2005). Formant frequencies of RP monophthongs in four age groups of speakers. *Journal of the International Phonetic Association*, 35(2), 183-199.

Hillenbrand, J., Getty, L. A., Clark, M. J., and Wheeler, K. (1995). Acoustic characteristics of American English vowels. *Journal of the Acoustical Society of America*, 97, 3099–3111.

Holbrook, A., and Fairbanks, G. (1962). Diphthong formants and their movements. *Journal of Speech and Hearing Research*, 5(1), 38-58.

Hombert, J. M. (1978). Consonant types, vowel quality and tone. In V.A. Fromkin (ed.), *Tone: a linguistic survey*. New York: Academic Press, 77-111.

Hossain, S. A., Rahman, M. L., and Ahmed, F. (2007). Acoustic classification of Bangla vowels. *Proceedings of World Academy of Science, Engineering and Technology*, 20, 321-326.

House, A. S., and Fairbanks, G. (1953). The influence of consonantal environment upon the secondary acoustic characteristics of vowels. *Journal of the Acoustical Society of America*, 25, 105-113.

Howie, J. M. (1976). *Acoustical studies of Mandarin vowels and tones*. Cambridge: Cambridge University Press.

Huang, X. (黄雪贞). (1988). 客家方言声调的特点.《*方言*》 4, 241-246.

Huang, X. (黄雪贞). (1989). 客家方言声调的特点续论.《*方言*》 2, 121-124.

Huang, X. (黄雪贞). (1992). 梅县客家话的语音特点.《*方言*》4, 275-289.

Huang, X. (黄雪贞). (1995). *《梅县方言词典》*. Nanjing: Jiangsu Jiaoyu Chubanshe.

Hsieh, F. (2007). ***Relational Correspondence in Tone Sandhi***. PhD thesis. Massachusetts Institute of Technology.

Hsu, H. (2008). /iu/ and /ui/ in Chinese Dialects. In Y. E. Hsiao, H. Hsu, L. H. Wee, & D. Ho (eds.), *Interfaces in Chinese Phonology*. Taiwan: Institute of Linguistics, Academic Sinica, 121-138.

Idson, W. L., and Massaro, D. W. (1980). The role of perceived duration in the identification of vowels. *Journal of Phonetics*, 8, 407-425.

Ishihara, S. (2004). *An Acoustic-Phonetic Descriptive Analysis of Kagoshima Japanese Tonal Phenomena*. PhD thesis. The Australian National University.

Jassem, W. (1975). Normalization of F_0 curves. In G. Fant and M.A.A. Tatham, *Auditory Analysis and Perception of Speech*. London: Academic Press.

Jassem, W. (1979). Classification of fricative spectra using statistical discriminant functions. In B. Lindblom & S. Öhman (eds.), *Frontiers of Speech Communication Research*. New York: Academic Press, 77-91.

Jha, S. K. (1985). Acoustic analysis of the Maithili diphthongs. *Journal of Phonetics*, 13, 107-115.

Johnson, K., and Martin, J. (2001). Acoustic vowel reduction in Creek: Effects of distinctive length and position in the word. *Phonetica*, 58, 81–102.

Karlgren, B. J. (1915-1926). *Etudes sur le phonologie Chinoise*. Chinese translation: 趙元任, 羅常培, 李方桂合譯 《中國音韻學研究》, 北京 : 商務印書館, 1994 年.

Karlgren, B. J. (1954) Compendium of phonetics in Ancient and Archaic Chinese. *Bulletin of the Museum of Far Eastern Antiquities*, 26, 211-367. Chinese translation: 張洪年譯 《中國聲韻學大綱》, 台北 : 國立編譯館, 1972 年.

Kent, R. D., and Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Australia: Thomson Learning.

Kent, R. D., and Moll, K. L. (1972). Tongue body articulation during vowel and diphthong gestures. *Folia Phoniatrica*, 24, 278-300.

Klatt, D. H. (1976). Linguistic uses of segmental duration in English: Acoustic and perceptual evidence. *Journal of the Acoustical Society of America*, 59, 1208-1221.

Ladd, D. R., and Silverman, K. E. A. (1984). Vowel intrinsic pitch in connected speech. *Phonetica*, 41, 31-40.

Ladefoged, P. (2001). *A course in phonetics* (4th ed.). US: Harcourt College Publishers.

Ladefoged, P., and Maddieson, I. (1990). Vowels of the world's languages. *Journal of Phoentics*, 18, 93-122.

Lee, W.-S., and Zee, E. (2003). Standard Chinese (Beijing). *Journal of the International Phonetic Association*, 33, 109-112.

Lehiste, J., and Peterson, G. E. (1961a). Transitions, glides, and diphthongs. *Journal of the Acoustical Society of America*, 33, 268-277.

Lehiste, J., and Peterson, G. E. (1961b).Some basic considerations in the analysis of intonation. *Journal of the Acoustical Society of America*, 33, 419-425.

Liljencrants, J., and Lindblom, B. (1972). Numerical simulation of vowel quality system: The role of perceptual contrast. *Language*, 48, 839-862.

Lindblom, B. (1963). Spectrographic study of vowel reduction. *Journal of the Acoustical Society of America*, 35, 1773-1781.

Lindblom, B. (1986). Phonetic universals in vowel systems. In J. Ohala & J. Jaeger (eds.), *Experimental Phonology*. New York: Academic Press, 13-44.

Lindau, M., Norlin, K., and Svantesson, J.-O. (1990). Some cross-linguistic differences in diphthongs. *Journal of the International Phonetic Association*, 20(1), 10-14.

Lü, J.G. (吕建国). (2006). 梅縣客家方言單字調和連讀變調實驗分析. 《*嘉應學院學報(哲學社會科學)*》, 24(4), 31-36.

Lisker, L., and Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20, 384–422.

Maddieson, I. (1978). Universals of tone. In J.H. Greenberg (ed.), ***Universals of human language (Phonology)***, 2. Stanford, California: Stanford University Press, 335-365.

Manrique, A. M. B. de (1979). Acoustic analysis of the Spanish diphthongs. *Phonetica*, 36, 194-206.

Mohr, B. (1971). Intrinsic variations in speech signal. *Phonetica*, 23, 65-93.

Moore, B. C. (1997). Aspects of auditory processing related to speech perception. In W. Hardcastle and J. Laver (eds.), *The Handbook of Phonetic Sciences*. Oxford: Blackwell, 539-565.

Most, T., Amir, O., and Tobin, Y. (2000). The Hebrew vowel system: Raw and normalized acoustic data. *Language and Speech*, 43(3), 295-308.

Nearey, T. M. (1978). *Phonetic Feature Systems for Vowels*. Indiana: Indiana University Linguistics Club.

Norman, J. (1988). *Chinese.* . Cambridge: Cambridge University Press.

O'Connor, J. D., Gerstman, L. J., Liberman, A. M., Delattre, P. C., & Cooper, F. S. (1957). Acoustic cues for the perception of initial /w, j, r, l/ in English. *Word*, 13, 24-43.

Peterson, G.. E., and Barney, H. L. (1952). Control methods used in a study of the vowels. *Journal of the Acoustical Society of American*, 24, 175-184.

Peterson, G. E., and Lehiste, I. (1960). Duration of syllable nuclei in English. *Journal of the Acoustical Society of America*, 32(6), 693-703.

Ren, H. (1986). On the acoustic structure of diphthongal syllables. *UCLA working papers in phonetic 65*.

Rose, P. (1982). Acoustic characteristics of the Shanghai-Zhenhai syllable types. In D. Bradley (ed.), *Papers in Southeast Asian Linguistics, No.8, Tonation*, *Pacific Linguistics*, *A-62*, 1-54.

Rose, P. (1987). Considerations in the normalization of the fundamental frequency of linguistic tone. *Speech Comunication*, 6, 343–351.

Rose, P. (1990). Acoustics and phonology of complex tone sandhi. *Phonetica* 47, 1-35.

Rose, P. (2002). *Forensic Speaker Identification*. London; New York: Taylor & Francis.

Saussure, F. de (1983). *Course in general linguistics*. LaSalle, Ill.: Open Court. English translation by Roy Harris

Simpson, A. P. (2009). Phonetic differences between male and female speech. *Language and Linguistic Compass*, 3/2, 621-640.

Simpson, A. P., and Ericsdotter, C. (2003). Sex-specific durational differences in English and Swedish. *The Proceedings of the 15th ICPhS*, Barcelona, 1113-1116.

Stevens, K. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In E. David & P. Denes, *Human Communication: A Unified View* (pp. 51-66). New York: McGraw-Hill.

Stevens, K. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3-46.

Stevens, S. S., Volkmann J., and Newman, E. B. (1937). A scale for the measurement of the psychological magnitude pitch. *Journal of the Acoustical Society of America*, 8, 185-190.

Stevens, S. S., and Volkmann, J. (1940). The relation of pitch to frequency: A revised scale. *The American Journal of Psychology*, 53 (3), 329-353.

Trager, G. L., and Smith, H. L. (1957). *An Outline of English Structure*. Washington: American Council of Learned Societies.

Traunmüller, H., and Eriksson, A (1995, manuscript). *The frequency range of the voice fundamental in the speech of male and female adults*. Retrieved May 27, 2010 from http://www.ling.su.se/staff/hartmut/f0_m&f.pdf

Thurgood, G. (2003). A subgrouping of the Sino-Tibetan languages: The interaction between language contact, change, and inheritance. In G. Thurgood & R. J. LaPolla (eds.), *The Sino-Tibetan languages*. London: Routledge, 3-21.

Wang, L. (王力). (1956).《*汉语音韵学*》. Beijing: Zhonghua Shuju.

Wen, C. (温昌衍). (2006).《*客家方言*》. Guangzhou: Huanan Ligong Daxue Chubanshe.

Whalen, D. H., and Levitt, A. G. (1995). The universality of intrinsic F_0 of vowels. *Journal of Phonetics*, 23, 349-366.

Wong, Y. W., and Xu, Y. (2007). Consonantal perturbation of F_0 contours of Cantonese tones. *The Proceedings of the 16th ICPhS*, Saarbrücken, 1293-1296.

Xie, L. (谢留文). (2003).《*客家方言语音研究*》. Beijing: China Social Sciences Press.

Xie, Y. (谢永昌). (1994).《*梅县客家方言志*》. Guangzhou: JiNan University Press.

Xu, C. X., and Xu, Y. (2003). Effects of consonant aspiration on Mandarin tones. *Journal of the International Phonetic Association*, 33(2), 165-181.

Yadav, R. (1979). The influence of aspiration on vowel duration in Maithili. *South Asian Languages Analysis*, 1, 157-165.

Yang, C. S., and Kasuya, H. (1996). Speaker individualities of vocal tract shapes of Japanese vowels measured by magnetic resonance images. *Proceedings of The 4th International Conference on Spoken Language Processing*, Philadelphia, USA, 949-952.

Yip, M. (2001). Tonal features, tonal inventories, and phonetic targets. *UCL Working Papers in Linguistics*, 13, 161-188.

Yip, M. (2007) Tone. In P. de Lacy (ed.), *The Cambridge Handbook of Phonology*. Cambridge: Cambridge University Press, 229-252.

Yuan, J. H., et al. (袁家骅等). (2001).《*汉语方言概要*》. Beijing: Wenzi Gaige Chubanshe.

Zee, E. (1999). An acoustical analysis of the diphthongs in Cantonese. *Proceedings of the 14th International Congress of Phonetics Sciences*, San Francisco, USA, 1101-1105.

Zee, E. (2000). Phonetic value of the vowels in Cantonese. *Proceedings of the 2nd International Symposium on Chinese Spoken Language Processing*.

Zee, E. (2001). The phonetic value of the vowels, diphthongs, and triphthongs in Beijing Mandarin. *Proceedings of the 5th National Conference on Modern Phonetics*, 54-60.

Zee, E. (2002). Apicality in Beijing Mandarin. *Symposium conjoint sur les langues et al cognition*, April 2002.

Zee, E., and Lee, W. S. (2001). An acoustical analysis of the vowels in Beijing Mandarin. *Proceedings of Eurospeech*, 1, 643-646.

Zee, E., and Maddieson, I. (1980). Tone and tone sandhi in Shanghai: phonetic evidence and phonological analysis. *Glossa*, 14:1, 45-88.

Zhou, R. (周日健). (1998). 客家方言的声调特点补论. In 李如龙 & 周日健 (eds.),《*客家方言研究：第二届客方言研讨会论文集*》. Guangzhou, China: Jinan University Press, 243-256.

Zhu, X. (1999). *Shanghai Tonetics*. Muenchen, Germany: Lincom Europa.

Appendix A

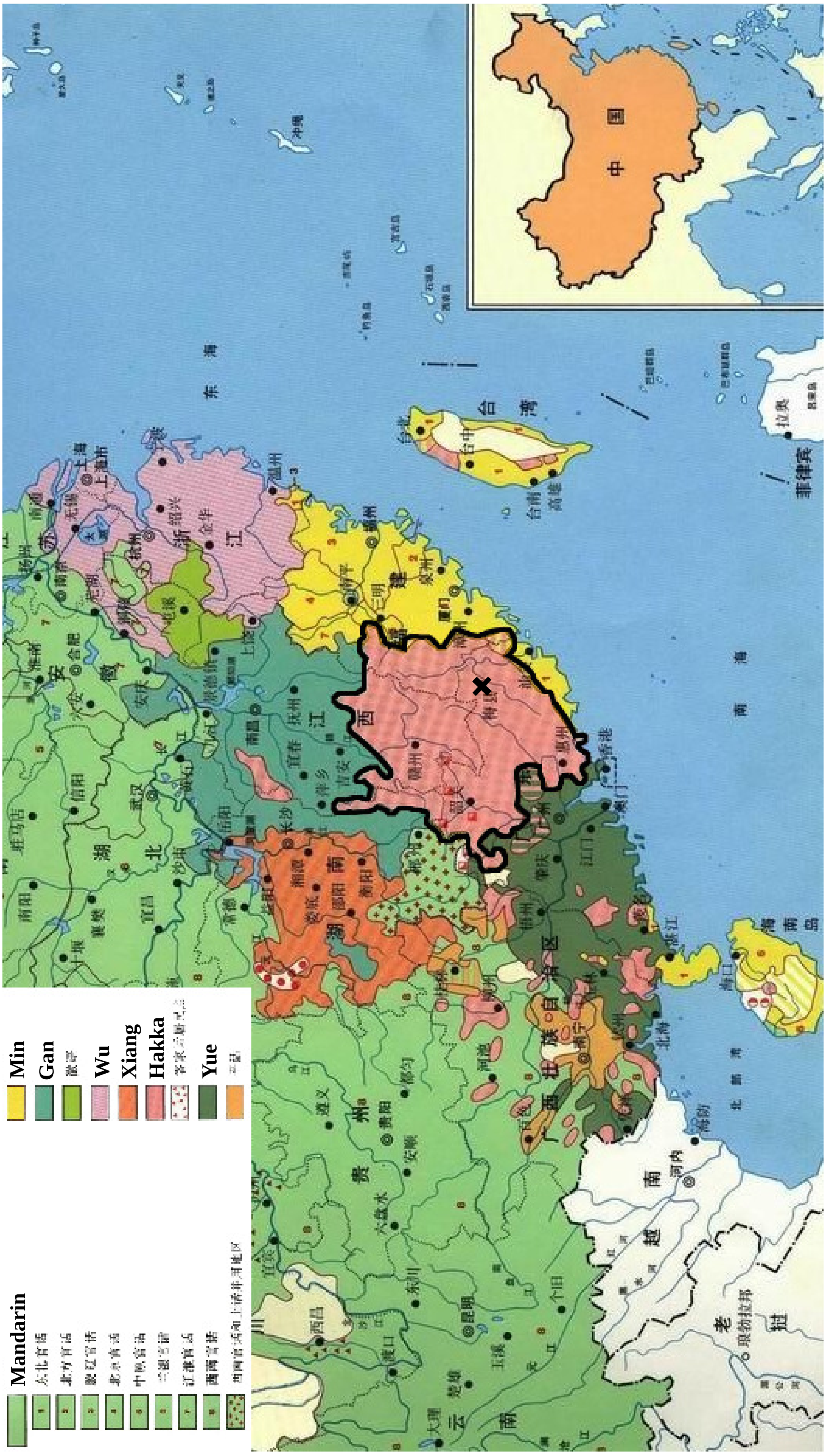
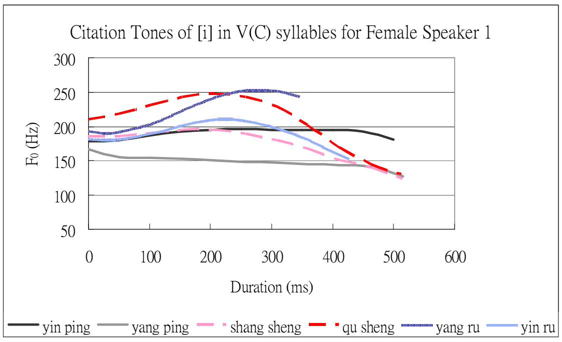
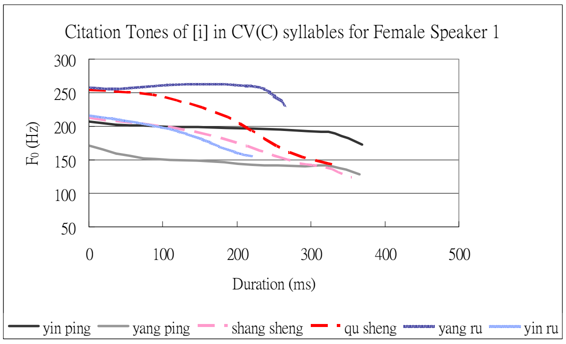


Figure 1 Detailed dialect map of southern China in which the location of the county of Meixian is marked by a cross

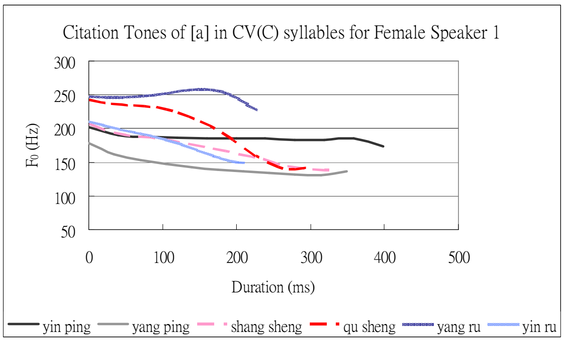
Appendix B



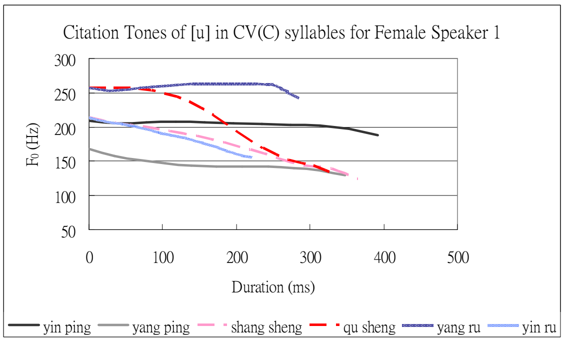
1a. Citation tones associated with [i] in V(C) syllables



1b. Citation tones associated with [i] in CV(C) syllables

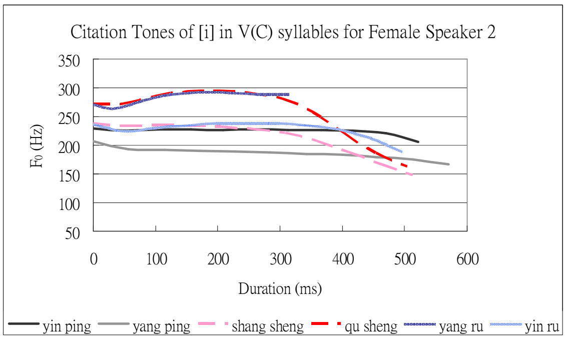


1c. Citation tones associated with [a] in CV(C) syllables

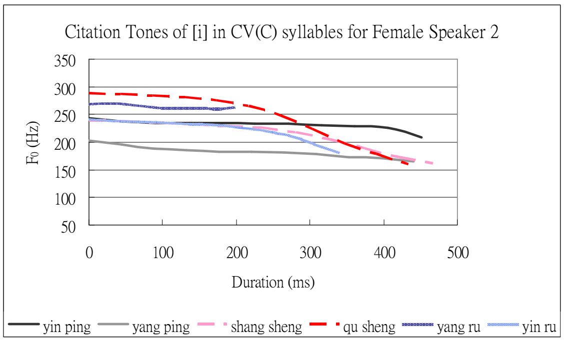


1d. Citation tones associated with [u] in CV(C) syllables

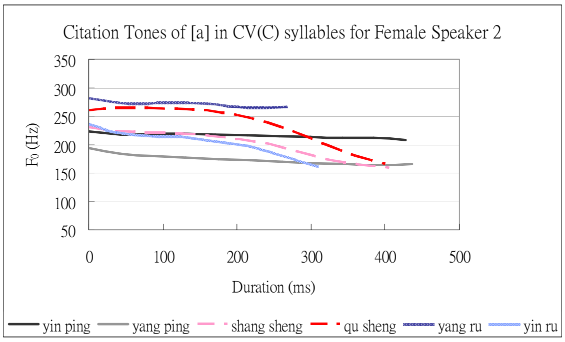
Figure 2a Mean F_0 contours of the six citation tones in Meixian Hakka produced by Female Speaker 1 in different phonetic contexts



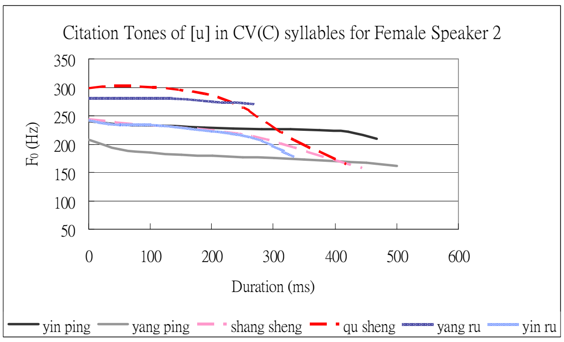
2a. Citation tones associated with [i] in V(C) syllables



2b. Citation tones associated with [i] in CV(C) syllables

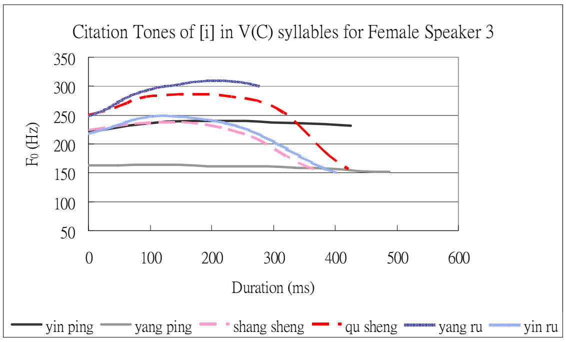


2c. Citation tones associated with [a] in CV(C) syllables

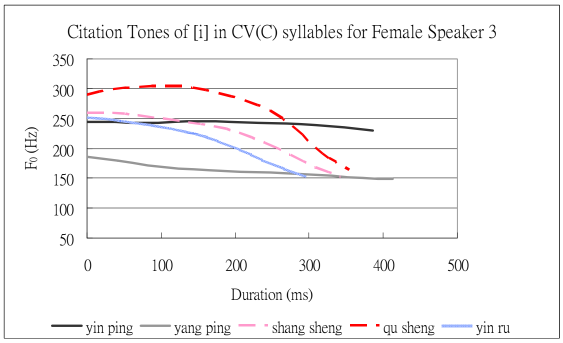


2d. Citation tones associated with [u] in CV(C) syllables

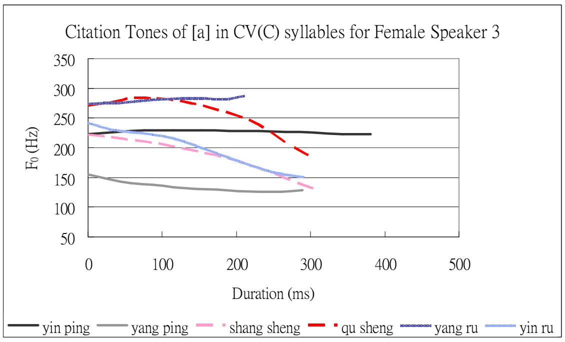
Figure 2b Mean F_0 contours of the six citation tones in Meixian Hakka produced by Female Speaker 2 in different phonetic contexts



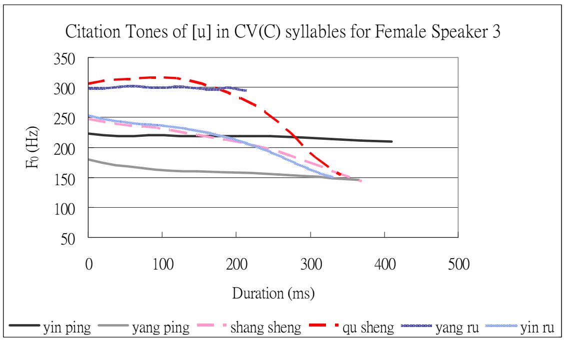
3a. Citation tones associated with [i] in V(C) syllables



3b. Citation tones associated with [i] in CV(C) syllables

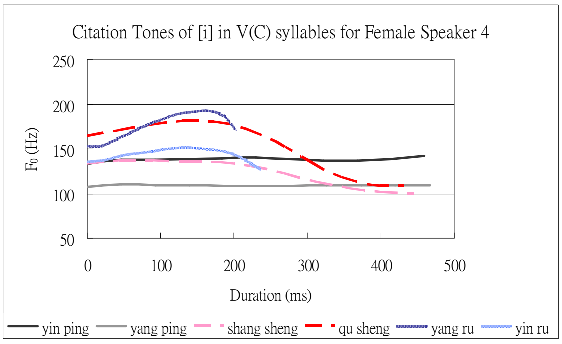


3c. Citation tones associated with [a] in CV(C) syllables

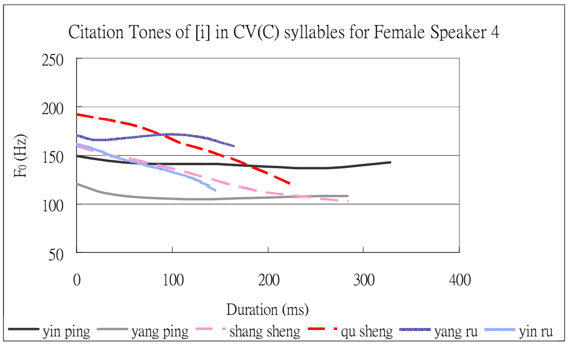


3d. Citation tones associated with [u] in CV(C) syllables

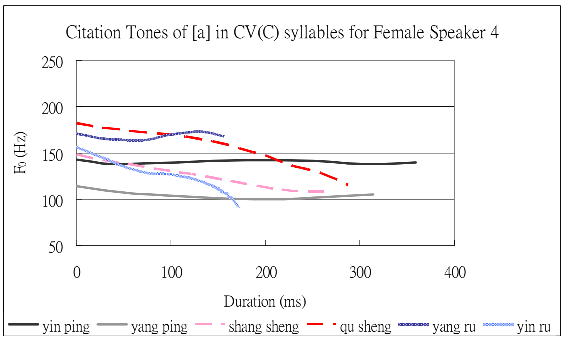
Figure 2c Mean F_0 contours of the six citation tones in Meixian Hakka produced by Female Speaker 3 in different phonetic contexts



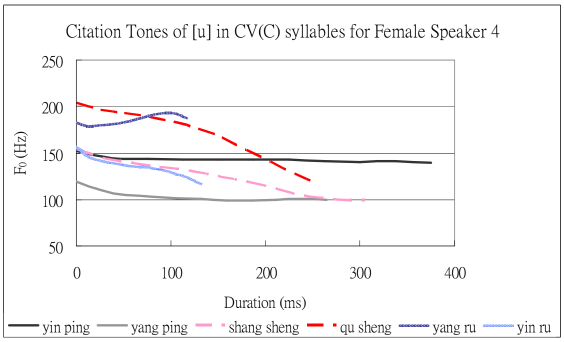
4a. Citation tones associated with [i] in V(C) syllables



4b. Citation tones associated with [i] in CV(C) syllables

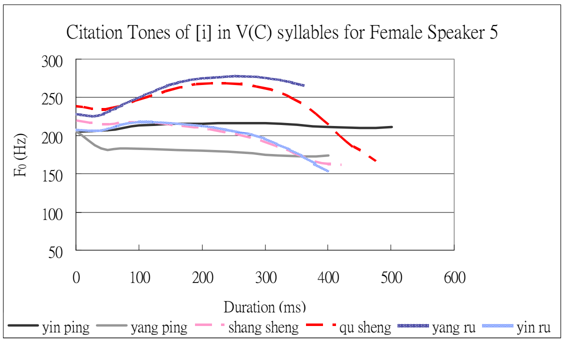


4c. Citation tones associated with [a] in CV(C) syllables

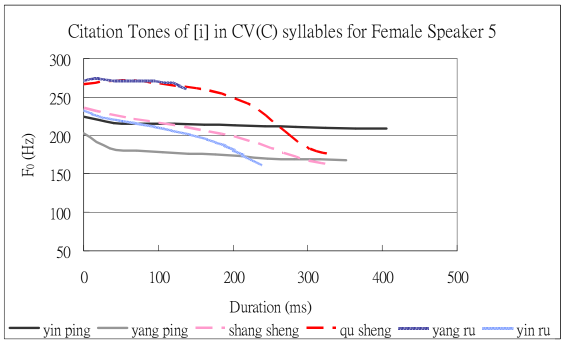


4d. Citation tones associated with [u] in CV(C) syllables

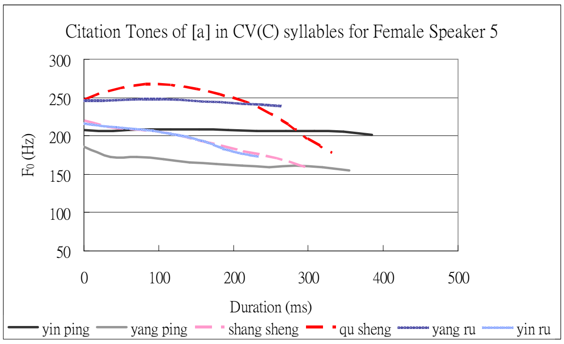
Figure 2d Mean F_0 contours of the six citation tones in Meixian Hakka produced by Female Speaker 4 in different phonetic contexts



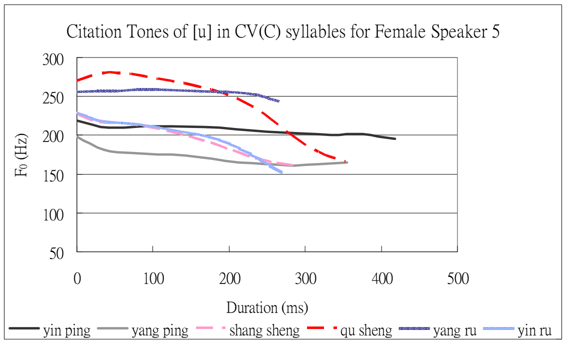
5a. Citation tones associated with [i] in V(C) syllables



5b. Citation tones associated with [i] in CV(C) syllables

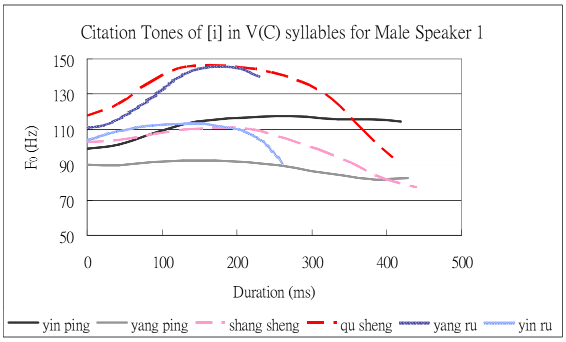


5c. Citation tones associated with [a] in CV(C) syllables

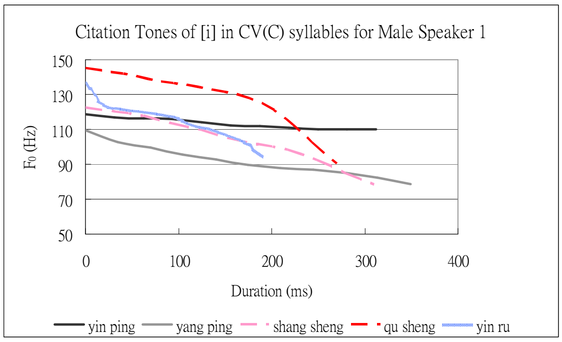


5d. Citation tones associated with [u] in CV(C) syllables

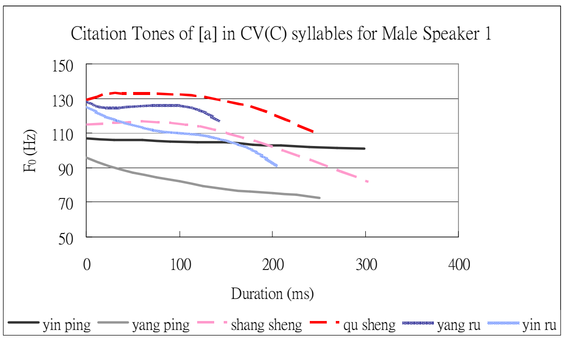
Figure 2e Mean F_0 contours of the six citation tones in Meixian Hakka produced by Female Speaker 5 in different phonetic contexts



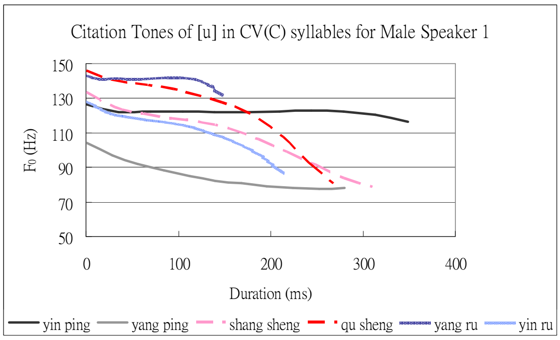
6a. Citation tones associated with [i] in V(C) syllables



6b. Citation tones associated with [i] in CV(C) syllables

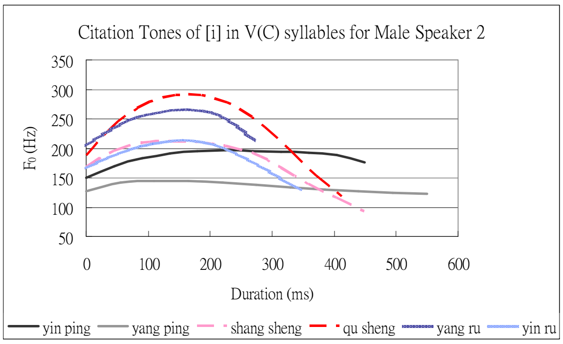


6c. Citation tones associated with [a] in CV(C) syllables

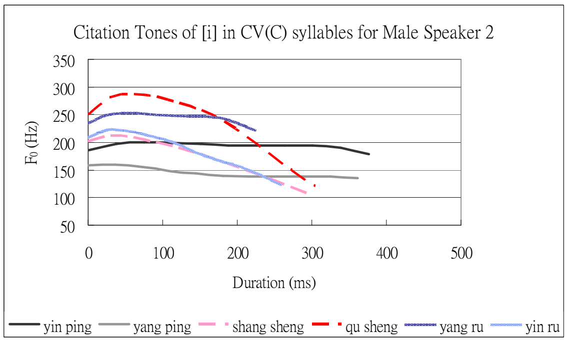


6d. Citation tones associated with [u] in CV(C) syllables

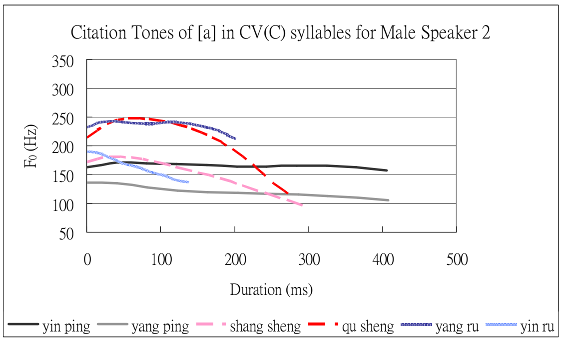
Figure 2f Mean F_0 contours of the six citation tones in Meixian Hakka produced by Male Speaker 1 in different phonetic contexts



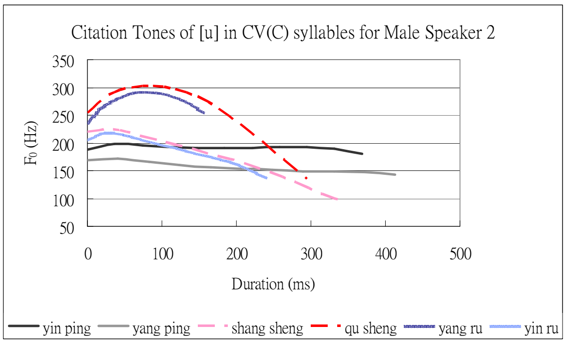
7a. Citation tones associated with [i] in V(C) syllables



7b. Citation tones associated with [i] in CV(C) syllables

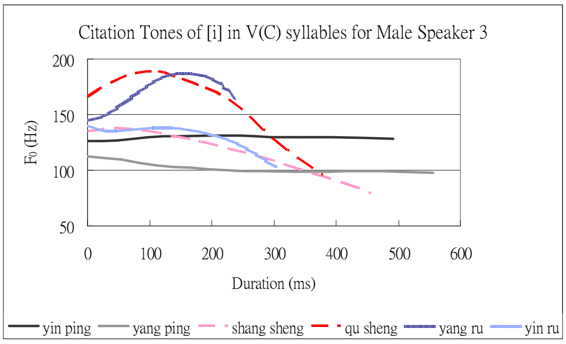


7c. Citation tones associated with [a] in CV(C) syllables

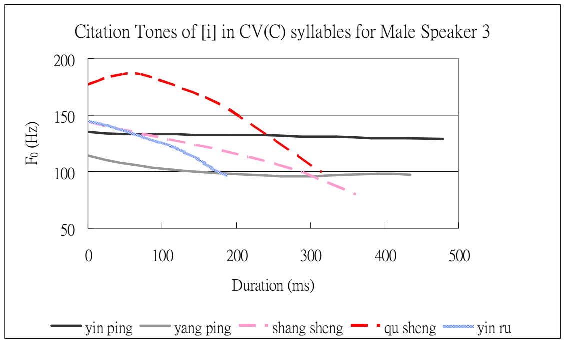


7d. Citation tones associated with [u] in CV(C) syllables

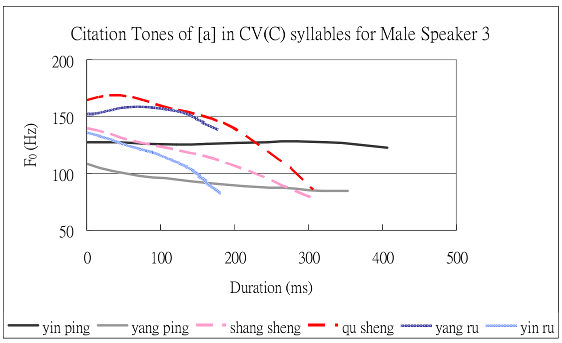
Figure 2g Mean F_0 contours of the six citation tones in Meixian Hakka produced by Male Speaker 2 in different phonetic contexts



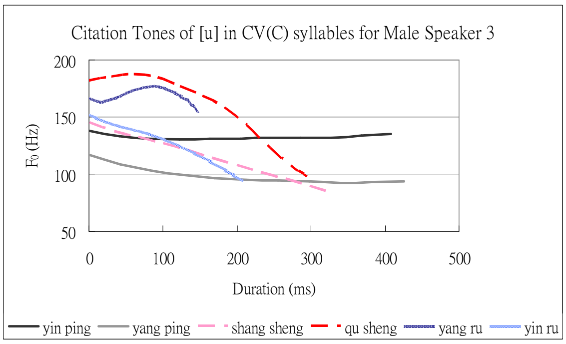
8a. Citation tones associated with [i] in V(C) syllables



8b. Citation tones associated with [i] in CV(C) syllables

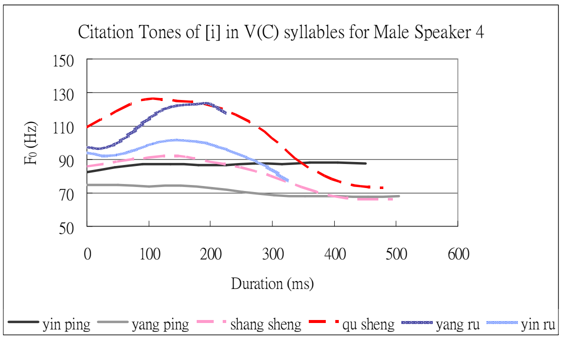


8c. Citation tones associated with [a] in CV(C) syllables

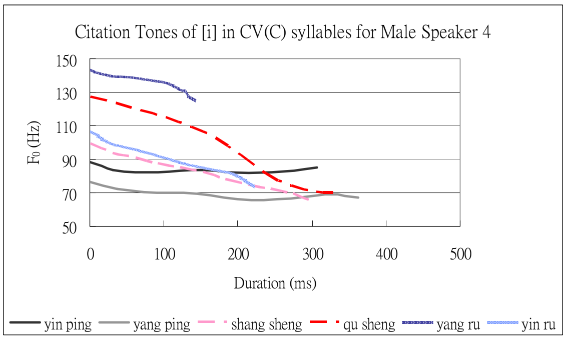


8d. Citation tones associated with [u] in CV(C) syllables

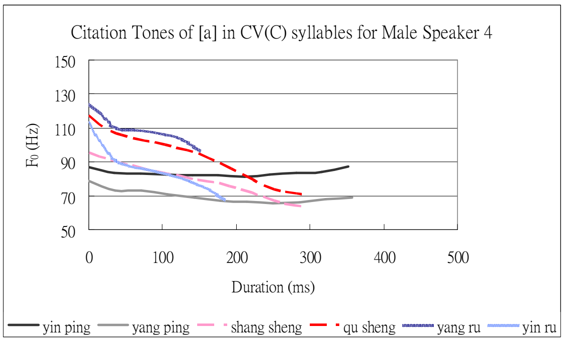
Figure 2h Mean F_0 contours of the six citation tones in Meixian Hakka produced by Male Speaker 3 in different phonetic contexts



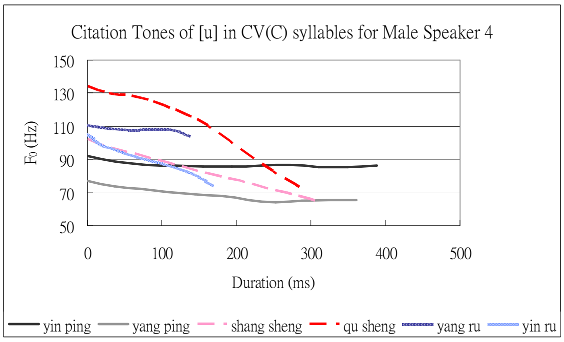
9a. Citation tones associated with [i] in V(C) syllables



9b. Citation tones associated with [i] in CV(C) syllables

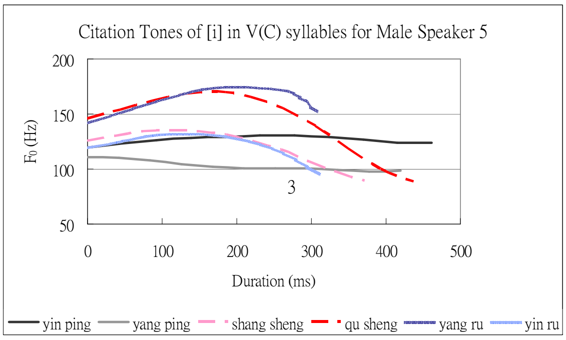


9c. Citation tones associated with [a] in CV(C) syllables

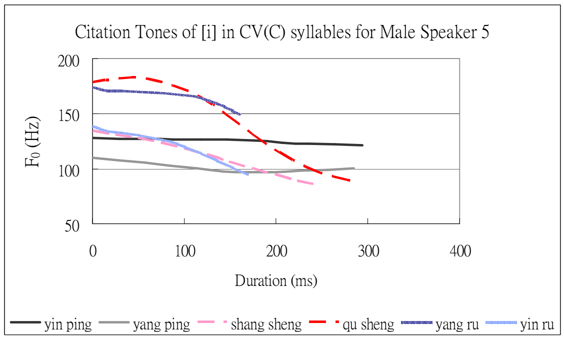


9d. Citation tones associated with [u] in CV(C) syllables

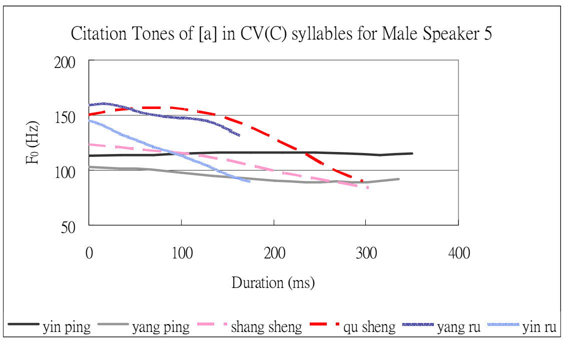
Figure 2i Mean F_0 contours of the six citation tones in Meixian Hakka produced by Male Speaker 4 in different phonetic contexts



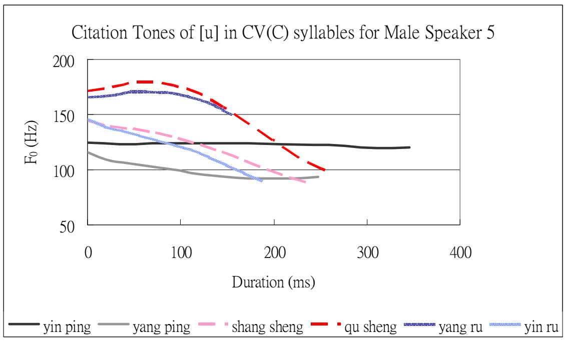
10a. Citation tones associated with [i] in V(C) syllables



10b. Citation tones associated with [i] in CV(C) syllables



10c. Citation tones associated with [a] in CV(C) syllables



10d. Citation tones associated with [i] in CV(C) syllables

Figure 2j Mean F_0 contours of the six citation tones in Meixian Hakka produced by Male Speaker 5 in different phonetic contexts

Appendix C

Vowel	Male speaker 1							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	238.4	16.6	2307.4	45.2	3173.2	32.9	355.6	10.8
[ɛ]	445.6	4.8	2114.3	72.3	2641.1	132.2	322.7	18.2
[a]	888.3	59.7	1373.6	40.3	2907.7	251.4	322.0	31.5
[o]	464.8	12.6	747.9	22.8	3143.4	54.1	322.3	22.4
[u]	376.6	24.2	827.8	10.8	2840.6	43.5	282.9	25.7
[ɿ]	328.9	15.0	1467.0	21.3	2775.4	92.8	268.0	18.8

Diphthong	Male speaker 1					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	862.8	47.5	1395.5	48.3	2733.7	384.6
[ai] (2nd)	432.8	32.8	1993.3	165.0	2837.8	277.8
[ia] (1st)	277.0	19.9	2185.1	143.4	3143.6	302.0
[ia] (2nd)	909.5	18.2	1469.5	24.4	2728.6	55.3
[ui] (1st)	423.5	9.6	1235.4	66.9	2701.4	35.0
[ui] (2nd)	338.5	8.4	2221.9	127.3	2760.2	107.2
[iu] (1st)	271.4	11.8	2193.3	46.6	2894.7	90.4
[iu] (2nd)	373.2	29.4	849.9	57.9	2520.5	226.4
[oi] (1st)	608.1	31.3	959.6	4.9	2332.3	488.5
[oi] (2nd)	399.0	41.5	1966.4	171.6	2561.9	137.7
[io] (1st)	238.2	28.3	2250.7	63.0	3001.7	136.9
[io] (2nd)	504.0	19.2	825.7	58.4	3192.1	81.9
[au] (1st)	707.7	37.5	1119.8	51.1	3039.9	77.6
[au] (2nd)	464.7	19.6	831.5	26.6	3244.1	61.6
[ua] (1st)	371.3	27.3	899.7	31.5	2843.8	132.3
[ua] (2nd)	895.2	38.2	1311.5	64.3	2934.1	87.2
[eu] (1st)	533.2	19.9	1798.0	75.1	2594.3	71.6
[eu] (2nd)	424.1	47.0	1028.7	116.2	2141.0	293.2
[ie] (1st)	259.4	33.7	2238.7	134.7	3343.0	276.5
[ie] (2nd)	439.5	39.5	2064.7	105.9	2653.8	85.2
[uo] (1st)	336.0	4.2	738.6	23.9	1872.0	493.5
[uo] (2nd)	493.3	15.8	766.7	18.9	3027.0	123.1

Triphthong	Male speaker 1					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	416.7	31.4	1985.1	180.8	3020.6	176.4
[iai] (2nd)	728.9	24.4	1548.2	86.0	2603.1	54.5
[iai] (3rd)	520.5	12.1	1931.8	106.1	2716.6	106.1
[iui] (1st)	232.5	28.0	2194.2	75.5	2882.5	184.8
[iui] (2nd)	413.5	3.9	1234.1	29.6	2623.7	62.6
[iui] (3rd)	289.3	12.3	2246.0	91.7	2905.0	214.7
[iau] (1st)	261.5	8.4	2230.1	73.6	3237.3	257.5
[iau] (2nd)	793.4	61.8	1256.9	69.3	2954.4	103.9
[iau] (3rd)	465.7	27.0	806.2	39.7	3161.9	94.8
[uai] (1st)	382.6	38.6	868.5	19.5	2916.0	224.2
[uai] (2nd)	831.7	65.5	1403.6	90.9	2393.6	166.7
[uai] (3rd)	423.5	31.1	2032.2	151.7	2649.3	72.9

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 1			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	42.7 (7.3)	170.7 (36.0)	163.8 (51.8)	377.3 (49.3)
[ia]	38.6 (4.5)	112.1 (14.9)	212.4 (31.0)	363.3 (42.4)
[ui]	55.7 (13.6)	82.0 (17.2)	80.5 (29.2)	218.3 (28.0)
[iu]	103.6 (20.2)	172.7 (16.6)	98.6 (13.0)	375.0 (26.7)
[oi]	74.1 (16.9)	143.8 (21.2)	72.2 (19.3)	290.3 (27.6)
[io]	64.6 (20.1)	122.7 (12.6)	155.2 (18.9)	342.6 (37.9)
[au]	53.2 (11.1)	122.9 (20.7)	91.2 (29.2)	267.4 (39.5)
[ua]	33.6 (3.6)	152.9 (36.8)	170.6 (35.2)	357.2 (25.8)
[eu]	96.7 (8.7)	111.7 (25.3)	50.4 (10.5)	259.0 (38.9)
[ie]	59.7 (17.4)	67.3 (17.5)	212.7 (29.1)	339.8 (45.7)
[uo]	0.0 (0.0)	47.2 (15.6)	292.8 (20.1)	340.0 (24.0)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 1					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	59.6 (19.0)	92.6 (8.6)	43.4 (6.9)	126.1 (15.0)	110.7 (15.4)	432.7 (44.0)
[uai]	32.0 (4.2)	105.6 (12.0)	42.8 (4.2)	139.8 (20.6)	98.5 (12.0)	418.8 (36.3)
[iui]	50.7 (9.2)	69.2 (8.7)	37.6 (7.1)	89.1 (16.3)	68.1 (8.3)	314.8 (24.9)
[iai]	0 (0.0)	55.5 (5.5)	44.6 (15.0)	86.9 (15.4)	57.6 (22.4)	244.6 (15.1)

Vowel	Male speaker 2							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	301.8	7.0	2408.2	37.5	3036.2	82.1	423.6	26.5
[ɛ]	482.5	28.0	2200.9	54.2	2662.2	81.5	378.6	32.2
[a]	836.3	21.3	1418.4	43.1	2724.3	149.8	442.4	21.8
[o]	478.9	62.0	771.2	122.4	2593.7	65.6	417.3	23.1
[u]	395.6	5.8	862.4	45.9	2450.8	42.2	398.7	56.6
[ʅ]	404.2	7.6	1125.1	30.0	2806.3	52.0	361.3	42.8

Diphthong	Male speaker 2					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	823.7	30.0	1456.8	29.0	2660.5	89.4
[ai] (2nd)	407.2	28.3	2256.6	21.8	2910.9	144.9
[ia] (1st)	282.0	27.5	2162.5	57.4	3487.3	163.7
[ia] (2nd)	854.8	22.4	1380.8	41.7	2573.5	43.7
[ui] (1st)	491.9	15.0	1078.9	35.0	2557.8	40.2
[ui] (2nd)	388.6	8.0	2033.0	14.7	2483.5	20.2
[iu] (1st)	290.0	24.0	2361.3	43.4	2929.1	46.1
[iu] (2nd)	389.4	42.9	952.4	71.1	2492.0	45.2
[oi] (1st)	615.0	10.2	1027.9	45.6	2745.6	63.2
[oi] (2nd)	491.4	14.4	1789.0	353.9	2273.6	216.7
[io] (1st)	267.3	11.9	2244.1	70.4	3248.6	249.5
[io] (2nd)	529.4	16.2	863.2	50.2	2632.9	50.5
[au] (1st)	789.2	27.4	1267.7	35.7	2832.0	204.7
[au] (2nd)	529.7	27.0	856.2	33.9	2642.7	26.4
[ua] (1st)	467.5	42.6	913.2	54.6	2559.0	148.9
[ua] (2nd)	814.0	19.0	1297.3	24.4	2458.6	183.6
[eu] (1st)	663.2	11.8	1761.3	38.3	2574.8	37.8
[eu] (2nd)	507.8	33.9	998.3	85.0	2503.0	28.2
[ie] (1st)	298.3	6.9	2409.9	42.8	3340.8	87.0
[ie] (2nd)	544.0	15.7	2150.1	61.3	2794.5	48.1
[uo] (1st)	374.1	21.1	844.5	50.4	2407.7	141.0
[uo] (2nd)	532.5	18.2	895.0	40.3	2692.6	63.8

Triphthong	Male speaker 2					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	474.3	20.7	1943.6	67.6	2393.9	148.5
[iai] (2nd)	782.1	8.0	1507.1	16.9	2484.5	71.3
[iai] (3rd)	513.5	25.3	2055.2	36.7	2610.3	85.0
[iui] (1st)	285.9	15.7	2333.3	30.9	2829.9	290.0
[iui] (2nd)	496.1	23.6	1260.3	35.2	2525.3	83.5
[iui] (3rd)	354.7	18.4	2157.8	73.5	2587.0	53.9
[iau] (1st)	278.5	6.8	2215.8	73.6	3306.4	194.2
[iau] (2nd)	746.7	28.8	1158.1	77.3	2754.5	37.2
[iau] (3rd)	544.1	35.7	920.6	55.9	2724.5	93.0
[uai] (1st)	475.9	39.0	931.7	54.0	2485.2	35.8
[uai] (2nd)	787.7	28.4	1357.1	39.7	2581.1	212.0
[uai] (3rd)	426.7	31.0	2143.4	67.9	2590.5	142.4

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 2			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	114.2 (19.4)	247.1 (32.5)	114.0 (16.8)	475.4 (59.2)
[ia]	83.6 (21.4)	88.8 (10.7)	308.4 (23.6)	480.9 (25.6)
[ui]	96.7 (10.8)	173.0 (21.0)	123.0 (22.6)	392.7 (35.5)
[iu]	106.1 (15.2)	250.7 (31.1)	184.6 (11.8)	541.4 (26.6)
[oi]	154.4 (15.4)	136.4 (20.4)	102.5 (27.9)	393.4 (54.2)
[io]	87.6 (12.7)	111.2 (23.9)	285.8 (41.6)	484.7 (37.9)
[au]	113.1 (20.2)	123.1 (22.9)	126.4 (17.4)	362.6 (30.0)
[ua]	58.3 (2.7) ③	89.6 (13.4)	310.7 (17.5)	444.2 (31.3)
[eu]	138.3 (16.2)	90.5 (16.8)	118.9 (10.4)	347.8 (24.9)
[ie]	93.9 (19.6)	63.5 (25.1)	403.6 (8.0)	561.1 (21.5)
[uo]	54.3 (20.0)	58.6 (12.0)	406.6 (50.0)	519.6 (39.2)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 2					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	67.2 (7.6)	98.4 (7.4)	102.2 (9.4)	99.1 (20.5)	133.5 (37.0)	500.7 (36.0)
[uai]	40.2 (16.4)	94.3 (27.7)	94.1 (14.9)	154.5 (24.7)	83.9 (18.9)	467.3 (14.2)
[iui]	66.7 (17.1)	88.5 (5.6)	42.9 (12.8)	128.5 (24.6)	109.9 (23.4)	436.7 (75.3)
[iai]	0.0 (0.0)	65.8 (6.4)	94.0 (8.7)	119.4 (12.6)	64.1 (24.0)	343.5 (29.1)

Vowel	Male speaker 3							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	287.3	13.6	2280.4	42.6	3381.0	62.3	448.6	30.2
[ɛ]	539.5	18.8	2243.9	40.8	2622.8	43.5	362.6	28.8
[a]	1000.5	12.3	1445.5	28.9	2665.0	54.1	349.2	19.5
[o]	548.0	21.1	832.6	17.9	3082.5	134.1	388.4	33.4
[u]	382.9	22.5	834.6	47.1	2698.4	292.9	358.7	11.6
[ɿ]	423.4	28.7	1380.7	77.4	2939.1	92.1	351.4	37.5

Diphthong	Male speaker 3					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1010.6	17.8	1523.0	40.4	2759.4	92.2
[ai] (2nd)	549.7	51.8	2240.0	33.5	2629.3	75.9
[ia] (1st)	311.4	15.3	2127.1	67.4	3225.6	46.8
[ia] (2nd)	954.7	25.3	1518.4	22.5	2455.9	58.4
[ui] (1st)	468.3	21.3	1001.7	50.8	2653.3	53.2
[ui] (2nd)	386.0	33.4	2273.4	35.2	2693.2	53.2
[iu] (1st)	275.3	9.7	2273.9	31.2	3151.9	198.1
[iu] (2nd)	395.0	22.1	824.4	30.4	2747.1	132.7
[oi] (1st)	688.2	46.2	971.8	30.3	3051.2	178.4
[oi] (2nd)	592.4	27.9	1966.7	46.1	2640.4	31.9
[io] (1st)	288.9	10.6	2242.8	20.8	3100.6	42.5
[io] (2nd)	638.0	7.5	915.2	26.4	3071.9	59.2
[au] (1st)	875.2	29.3	1242.4	47.8	2290.3	268.2
[au] (2nd)	622.9	23.2	932.6	55.3	3005.3	160.3
[ua] (1st)	457.0	39.3	919.1	14.5	2402.8	38.1
[ua] (2nd)	1003.7	17.8	1447.1	37.9	2606.2	122.8
[eu] (1st)	679.1	34.9	1939.6	27.2	2690.5	39.6
[eu] (2nd)	521.3	34.1	923.9	44.2	2065.0	187.3
[ie] (1st)	304.5	26.0	2194.5	100.1	3183.0	76.1
[ie] (2nd)	553.6	16.4	2109.7	36.7	2542.3	44.8
[uo] (1st)	423.9	17.8	833.7	56.7	2754.9	269.7
[uo] (2nd)	562.3	12.7	916.7	44.0	3081.0	31.0

Triphthong	Male speaker 3					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	497.5	18.8	2035.8	40.2	2820.5	52.3
[iai] (2nd)	908.4	25.7	1502.8	31.3	2431.8	21.1
[iai] (3rd)	684.0	18.3	2054.5	54.2	2602.4	101.1
[iui] (1st)	290.6	13.4	2252.8	49.4	3046.7	94.1
[iui] (2nd)	464.5	14.5	1132.8	81.2	2600.8	18.4
[iui] (3rd)	355.3	14.5	2266.7	67.6	2592.7	49.9
[iau] (1st)	279.4	5.6	2111.3	47.9	3218.9	213.5
[iau] (2nd)	844.4	25.9	1120.4	27.4	2781.8	99.7
[iau] (3rd)	500.0	32.4	873.0	32.5	2966.6	93.5
[uai] (1st)	458.4	9.7	940.3	4.9	2404.8	59.7
[uai] (2nd)	973.3	32.1	1444.5	79.0	2661.6	63.5
[uai] (3rd)	573.7	43.5	2264.8	64.1	2642.7	33.3

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 3			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	111.4 (27.6)	156.8 (29.5)	113.0 (13.4)	381.2 (21.4)
[ia]	60.8 (8.4)	70.1 (5.8)	276.4 (28.8)	407.4 (34.5)
[ui]	83.7 (17.6)	108.6 (18.6)	101.1 (23.9)	293.6 (35.4)
[iu]	130.1 (29.8)	155.3 (17.6)	184.0 (11.5)	469.5 (40.2)
[oi]	140.1 (42.4)	119.5 (9.2)	37.0 (21.5)	296.7 (42.5)
[io]	37.9 (27.3)	83.7 (10.7)	225.1 (33.7)	346.8 (59.6)
[au]	121.8 (18.2)	108.9 (15.0)	49.2 (16.3)	280.0 (28.0)
[ua]	15.5 (1.7)	118.2 (21.3)	217.5 (11.6)	351.3 (13.2)
[eu]	101.7 (40.8)	120.3 (27.7)	74.9 (29.2)	297.0 (52.9)
[ie]	59.9 (28.3)	76.2 (10.5)	325.0 (68.0)	461.2 (72.6)
[uo]	29.7 (5.9)	56.5 (7.2)	303.1 (20.3)	389.3 (16.0)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 3					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	53.0 (9.6)	94.9 (6.2)	79.8 (15.4)	128.5 (12.2)	90.4 (13.6)	446.9 (19.2)
[uai]	18.1 (2.9)	94.6 (22.9)	94.0 (13.7)	117.2 (9.7)	79.1 (2.9)	403.1 (21.3)
[iui]	66.4 (21.3)	74.0 (14.7)	54.0 (8.0)	90.6 (18.7)	105.9 (24.0)	391.2 (59.5)
[iai]	0.0 (0.0)	55.1 (7.3)	89.2 (13.7)	98.6 (13.4)	59.9 (9.8)	302.9 (5.9)

Vowel	Male speaker 4							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	268.1	10.4	2438.8	22.8	3406.9	170.0	549.6	35.6
[ɛ]	580.1	16.0	2171.3	14.9	2763.8	49.4	429.8	41.4
[a]	966.5	10.1	1459.4	56.2	2745.5	93.5	457.0	45.7
[o]	589.2	6.8	861.4	26.9	3047.6	74.4	422.4	37.6
[u]	386.9	8.9	740.3	16.3	2915.4	21.5	432.9	50.5
[ʅ]	393.4	14.2	1278.5	27.6	3013.8	45.4	414.5	38.5

Diphthong	Male speaker 4					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	964.8	47.0	1475.7	53.1	2797.0	178.8
[ai] (2nd)	357.2	9.1	2540.9	29.8	3199.3	53.8
[ia] (1st)	281.0	8.7	2366.0	54.0	3550.0	103.4
[ia] (2nd)	982.0	15.5	1599.0	30.6	2758.8	58.5
[ui] (1st)	407.7	24.8	923.8	40.1	2973.8	78.1
[ui] (2nd)	287.0	4.8	2391.5	52.9	3011.4	186.7
[iu] (1st)	310.9	14.5	2334.4	16.4	3112.3	76.1
[iu] (2nd)	349.7	32.2	803.5	15.8	2865.5	38.2
[oi] (1st)	650.6	29.8	1005.6	46.4	2914.8	46.2
[oi] (2nd)	377.6	10.5	2335.3	23.7	2876.2	132.4
[io] (1st)	239.6	14.3	2412.0	50.2	2947.6	212.6
[io] (2nd)	567.9	18.0	885.4	22.8	3097.5	97.1
[au] (1st)	878.2	43.4	1305.5	38.8	2943.2	111.5
[au] (2nd)	491.7	51.2	758.6	58.9	3131.7	334.0
[ua] (1st)	449.6	21.9	730.6	57.4	2750.0	219.3
[ua] (2nd)	941.2	11.1	1466.0	23.0	2808.9	277.5
[eu] (1st)	700.2	11.2	1935.6	30.8	2644.9	108.8
[eu] (2nd)	420.3	30.0	878.1	88.7	2743.8	113.8
[ie] (1st)	292.3	23.6	2468.4	106.5	3547.0	158.6
[ie] (2nd)	522.9	43.0	2276.9	77.0	3095.4	246.7
[uo] (1st)	382.0	14.4	655.2	39.6	2800.1	160.9
[uo] (2nd)	601.9	7.0	840.5	27.4	3053.7	121.5

Triphthong	Male speaker 4					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	451.5	30.7	2203.6	61.8	2796.6	143.7
[iai] (2nd)	903.2	22.3	1584.2	29.1	2879.5	197.2
[iai] (3rd)	381.8	18.7	2511.0	48.6	3079.0	75.0
[iui] (1st)	235.8	14.1	2344.5	64.4	2892.3	107.1
[iui] (2nd)	412.3	8.3	964.3	89.0	2953.4	82.1
[iui] (3rd)	284.1	10.5	2349.2	34.6	2682.2	68.7
[iau] (1st)	282.6	11.9	2373.0	16.9	3572.3	69.6
[iau] (2nd)	860.0	21.3	1285.2	27.9	2793.8	119.0
[iau] (3rd)	392.6	41.1	668.9	45.1	3104.6	84.7
[uai] (1st)	380.4	48.5	658.1	67.7	2747.2	166.3
[uai] (2nd)	843.5	74.5	1335.5	68.9	2562.4	187.7
[uai] (3rd)	392.7	35.7	2487.4	85.9	3178.9	83.4

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 4			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	173.1 (19.5)	189.9 (21.4)	112.7 (25.2)	475.8 (34.2)
[ia]	124.8 (49.8)	98.3 (13.6)	340.6 (35.6)	563.8 (77.0)
[ui]	110.6 (33.2)	138.1 (20.5)	149.4 (50.1)	398.1 (88.5)
[iu]	290.6 (35.0)	153.4 (11.4)	199.0 (50.5)	643.1 (68.5)
[oi]	214.7 (49.6)	147.9 (18.7)	121.4 (28.1)	484.1 (83.2)
[io]	140.8 (33.0)	109.2 (12.8)	323.4 (39.0)	573.5 (75.5)
[au]	152.7 (18.6)	142.6 (36.4)	150.5 (28.2)	445.9 (63.0)
[ua]	42.0 (19.5) ②	107.8 (22.1)	337.5 (46.5)	466.3 (53.0)
[eu]	215.0 (34.7)	114.8 (20.2)	113.2 (21.1)	443.1 (34.9)
[ie]	168.9 (85.6)	83.9 (27.2)	383.8 (58.4)	636.6 (155.2)
[uo]	46.3 (3.8)	89.7 (16.7)	305.8 (41.7)	441.8 (42.3)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 4					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	116.1 (17.9)	89.3 (13.8)	151.2 (66.7)	132.3 (37.7)	119.2 (20.6)	608.2 (114.7)
[uai]	34.8 (8.5)	110.7 (8.1)	96.6 (8.5)	150.9 (26.5)	119.3 (7.8)	512.4 (33.6)
[iui]	139.2 (25.4)	103.5 (17.7)	97.4 (37.4)	108.4 (26.0)	147.3 (33.6)	595.9 (126.9)
[iai]	0.0 (0.0)	60.6 (5.6)	90.6 (14.3)	125.6 (29.2)	89.7 (30.7)	366.8 (68.5)

Vowel	Male speaker 5							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	304.5	5.5	1894.0	16.4	3087.9	34.8	501.5	35.4
[ɛ]	537.1	12.5	1888.9	2.2	2449.6	39.9	434.6	54.2
[a]	860.9	20.6	1408.4	31.1	2551.4	46.3	429.0	14.5
[o]	523.2	6.0	778.5	8.5	2788.1	148.7	441.7	10.4
[u]	383.8	18.6	719.4	24.9	2515.4	53.6	408.9	43.2
[ʅ]	387.9	18.2	1281.8	16.1	2652.2	37.3	375.8	35.7

Diphthong	Male speaker 5					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	886.5	14.0	1398.9	30.9	2464.8	36.1
[ai] (2nd)	367.5	14.7	2005.9	29.3	3010.8	179.1
[ia] (1st)	272.3	1.9	1972.7	28.5	3283.4	69.7
[ia] (2nd)	857.2	8.6	1464.7	13.8	2559.9	52.1
[ui] (1st)	451.6	10.6	910.5	37.7	2711.4	98.2
[ui] (2nd)	294.6	10.2	1933.5	23.3	2966.0	199.8
[iu] (1st)	280.3	6.9	1939.4	18.1	3192.5	153.1
[iu] (2nd)	358.2	18.8	793.2	33.5	2416.2	64.1
[oi] (1st)	668.2	15.6	991.5	28.7	2712.6	18.3
[oi] (2nd)	397.9	35.2	1914.7	19.5	2487.7	54.7
[io] (1st)	288.6	10.3	1911.5	58.2	3169.8	67.4
[io] (2nd)	624.3	23.9	926.1	20.1	2702.5	53.6
[au] (1st)	800.0	32.7	1128.6	54.7	2693.6	53.9
[au] (2nd)	574.0	31.8	775.5	50.8	2519.1	99.7
[ua] (1st)	427.7	27.5	727.1	26.7	2268.0	117.0
[ua] (2nd)	859.6	22.4	1399.3	16.9	2389.5	29.3
[eu] (1st)	650.4	19.2	1699.6	24.8	2286.7	9.1
[eu] (2nd)	499.9	20.7	970.4	58.1	2215.1	278.2
[ie] (1st)	266.1	6.2	1977.7	29.1	3203.3	26.3
[ie] (2nd)	527.8	8.1	1872.1	26.1	2444.4	52.3
[uo] (1st)	362.6	27.7	731.3	60.7	2339.3	98.5
[uo] (2nd)	607.5	13.5	862.5	32.8	2837.2	70.0

Triphthong	Male speaker 5					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	503.2	13.4	1740.6	15.0	2945.8	45.1
[iai] (2nd)	842.3	21.4	1477.5	30.9	2419.5	46.6
[iai] (3rd)	497.9	52.7	1954.3	24.0	2563.1	51.4
[iui] (1st)	288.7	5.5	1899.8	47.8	3116.7	30.1
[iui] (2nd)	468.3	5.6	964.0	32.3	2556.9	48.0
[iui] (3rd)	302.8	14.0	1890.4	24.9	2938.0	165.6
[iau] (1st)	283.8	7.9	1920.1	20.1	3205.9	24.8
[iau] (2nd)	778.6	22.7	1203.8	84.2	2531.3	46.2
[iau] (3rd)	486.3	14.1	762.0	14.6	2580.5	85.4
[uai] (1st)	504.0	61.6	809.9	82.2	2329.1	44.1
[uai] (2nd)	866.0	26.8	1390.8	26.3	2329.6	58.1
[uai] (3rd)	384.7	27.1	2027.1	25.5	2660.2	219.9

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 5			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	138.1 (25.2)	195.2 (41.3)	111.1 (14.7)	444.5 (22.2)
[ia]	95.0 (25.2)	90.6 (6.1)	365.6 (11.0)	551.3 (33.1)
[ui]	112.2 (23.9)	189.5 (24.5)	112.9 (12.7)	414.7 (39.2)
[iu]	236.4 (35.1)	205.5 (62.9)	100.4 (13.6)	542.4 (73.3)
[oi]	204.6 (42.9)	176.9 (39.3)	64.2 (12.7)	445.7 (56.3)
[io]	48.0 (13.9)	97.1 (17.9)	309.4 (45.4)	454.6 (66.5)
[au]	155.5 (23.3)	155.0 (42.8)	86.7 (9.1)	397.3 (32.5)
[ua]	23.7 (2.8) ③	137.0 (7.9)	271.5 (44.2)	426.4 (37.1)
[eu]	246.2 (18.9)	139.1 (16.4)	90.7 (7.3)	476.1 (33.8)
[ie]	80.1 (2.2)	92.5 (20.6)	327.7 (39.7)	500.4 (35.8)
[uo]	16.1 (3.7)	110.3 (14.5)	352.7 (51.2)	476.1 (55.2)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 5					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	61.1 (15.5)	88.3 (2.95)	94.8 (15.3)	164.2 (30.2)	80.4 (12.8)	489.0 (17.7)
[uai]	21.5 (9.1) ③	114.7 (22.1)	83.2 (14.7)	150.5 (34.0)	79.4 (22.3)	441.0 (35.9)
[iui]	77.0 (12.1)	82.3 (11.5)	77.5 (12.4)	126.3 (24.4)	137.9 (16.4)	501.2 (34.0)
[iai]	0.0 (0.0)	55.8 (4.4)	164.5 (37.0)	130.8 (23.5)	55.2 (12.9)	406.4 (55.3)

Vowel	Male speaker 6							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	260.7	4.0	1988.3	43.7	2896.9	38.4	583.9	43.2
[ɛ]	519.5	16.0	1802.1	56.3	2568.1	48.6	592.6	32.6
[a]	827.6	10.0	1287.7	17.3	2556.0	65.6	546.3	21.7
[o]	494.6	12.8	720.0	16.6	2466.8	46.4	608.3	43.5
[u]	358.0	26.6	682.3	24.0	2537.0	57.0	537.1	31.2
[ʅ]	344.8	15.8	1195.5	17.7	2385.0	38.1	523.1	31.1

Diphthong	Male speaker 6					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	831.3	53.9	1297.9	23.6	2360.7	106.7
[ai] (2nd)	358.4	25.3	1964.2	66.6	2684.8	240.7
[ia] (1st)	226.9	19.9	1829.9	44.1	2808.0	118.4
[ia] (2nd)	846.5	37.5	1329.7	18.7	2510.6	81.9
[ui] (1st)	447.8	14.1	897.6	44.0	2565.6	123.9
[ui] (2nd)	318.5	19.4	1833.5	55.2	2828.2	78.4
[iu] (1st)	287.6	15.5	1841.8	32.3	2782.3	91.2
[iu] (2nd)	354.4	19.9	787.0	32.1	2377.1	90.9
[oi] (1st)	634.6	20.1	953.8	32.3	2409.9	153.0
[oi] (2nd)	429.2	19.9	1711.7	26.3	2514.7	44.4
[io] (1st)	245.3	3.4	1756.7	39.4	2741.9	194.9
[io] (2nd)	563.9	39.6	849.1	52.3	2303.4	81.1
[au] (1st)	750.4	23.2	1135.7	21.6	2382.8	193.0
[au] (2nd)	513.7	20.5	756.5	14.2	2461.7	123.4
[ua] (1st)	424.0	68.0	789.1	73.2	2323.1	102.6
[ua] (2nd)	826.3	33.0	1302.7	22.9	2504.2	139.7
[eu] (1st)	611.6	33.1	1688.7	54.5	2579.1	68.1
[eu] (2nd)	437.9	17.4	790.2	26.1	2605.2	127.7
[ie] (1st)	242.4	17.5	1946.8	26.2	2927.0	147.1
[ie] (2nd)	489.2	17.2	1813.5	48.6	2544.2	57.7
[uo] (1st)	329.0	29.0	685.5	24.9	2243.0	151.6
[uo] (2nd)	538.3	16.1	799.0	1.7	2419.9	119.9

Triphthong	Male speaker 6					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	481.5	24.4	1555.1	38.1	2659.8	73.5
[iai] (2nd)	818.7	27.6	1380.6	29.4	2516.6	86.8
[iai] (3rd)	421.1	19.1	1889.6	50.3	2517.0	50.3
[iui] (1st)	243.7	4.6	1730.9	16.1	2850.3	62.2
[iui] (2nd)	457.9	22.7	913.4	18.5	2434.3	72.5
[iui] (3rd)	312.0	15.2	1788.2	20.0	2411.5	90.4
[iau] (1st)	228.5	15.9	1861.1	30.9	2828.5	167.0
[iau] (2nd)	769.5	38.0	1113.7	27.7	2219.7	94.9
[iau] (3rd)	443.3	23.9	663.0	13.7	2374.3	39.1
[uai] (1st)	423.2	37.9	803.2	48.5	2239.6	120.0
[uai] (2nd)	819.0	26.0	1302.9	36.6	2359.1	113.2
[uai] (3rd)	363.3	39.0	1962.5	51.1	2718.3	106.2

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 6			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	218.3 (39.3)	235.6 (43.2)	199.8 (48.2)	653.8 (29.9)
[ia]	78.9 (44.3)	99.3 (19.1)	428.3 (33.4)	606.6 (72.3)
[ui]	170.2 (19.8)	209.1 (53.8)	196.5 (50.4)	575.9 (65.8)
[iu]	225.0 (34.3)	204.8 (31.4)	228.3 (8.0)	658.3 (52.2)
[oi]	259.2 (27.6)	158.5 (14.5)	127.9 (37.4)	545.6 (24.7)
[io]	58.3 (14.8)	105.5 (10.0)	416.3 (58.5)	580.1 (44.9)
[au]	208.4 (11.6)	153.7 (46.6)	158.8 (24.5)	521.0 (39.8)
[ua]	29.5 (7.0) ③	149.8 (37.7)	390.8 (54.4)	562.8 (54.4)
[eu]	269.6 (41.8)	156.4 (21.0)	129.5 (16.6)	555.6 (47.0)
[ie]	147.5 (21.4)	79.7 (8.7)	438.5 (22.1)	665.8 (26.7)
[uo]	34.0 (9.4) ③	76.7 (20.0)	474.0 (37.5)	576.3 (11.8)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 6					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	61.1 (16.7)	97.6 (12.3)	168.0 (22.8)	182.2 (41.1)	160.2 (20.8)	669.4 (76.5)
[uai]	27.6 (9.5)	127.6 (21.1)	144.1 (34.9)	178.7 (28.0)	159.1 (55.4)	637.3 (78.4)
[iui]	85.5 (38.8)	81.4 (11.0)	93.7 (9.6)	180.1 (27.4)	226.1 (48.2)	666.9 (36.4)
[iai]	0.0 (0.0)	55.3 (16.1)	144.9 (22.4)	189.4 (32.7)	132.0 (23.9)	521.7 (67.4)

Vowel	Male speaker 7							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	254.3	3.8	2222.6	24.9	3150.3	75.2	487.3	18.5
[ɛ]	477.5	3.3	2076.4	24.4	2625.3	77.5	406.9	26.5
[a]	872.9	21.7	1376.7	12.1	2409.7	53.2	395.5	26.8
[o]	527.8	18.3	843.7	28.1	2608.0	83.7	423.2	52.2
[u]	368.3	5.4	747.9	18.4	2556.3	153.5	366.8	31.0
[ʅ]	387.5	13.1	1243.2	29.5	2719.3	64.7	338.3	35.6

Diphthong	Male speaker 7					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	872.6	48.9	1426.5	44.4	2479.7	155.7
[ai] (2nd)	345.0	25.5	2223.1	28.8	3073.2	65.5
[ia] (1st)	266.2	10.2	2117.8	12.2	3328.6	57.0
[ia] (2nd)	884.1	10.3	1410.0	29.8	2380.0	82.1
[ui] (1st)	410.6	10.6	907.2	20.7	2694.3	70.2
[ui] (2nd)	316.3	5.9	2101.8	35.2	2724.2	91.7
[iu] (1st)	259.7	16.8	2220.2	25.2	3202.1	58.4
[iu] (2nd)	334.8	7.2	698.0	18.7	2686.5	66.0
[oi] (1st)	617.3	18.7	1004.6	19.0	2449.6	48.7
[oi] (2nd)	341.3	12.7	2122.2	60.3	2869.4	74.2
[io] (1st)	254.4	10.5	2166.5	10.3	3326.1	29.0
[io] (2nd)	549.1	8.1	866.5	18.7	2482.7	28.9
[au] (1st)	820.2	21.0	1214.4	35.2	2308.8	38.7
[au] (2nd)	439.6	34.1	742.7	54.1	2634.2	94.2
[ua] (1st)	407.8	24.0	829.4	33.9	2490.1	87.7
[ua] (2nd)	850.6	24.1	1371.1	15.5	2202.0	38.1
[eu] (1st)	576.3	7.8	1865.2	22.9	2648.3	67.3
[eu] (2nd)	392.6	11.3	813.2	11.1	2480.8	125.4
[ie] (1st)	253.1	11.8	2147.4	46.9	3300.6	72.7
[ie] (2nd)	480.4	14.4	2057.0	32.5	2655.3	70.4
[uo] (1st)	356.2	6.0	723.7	39.0	2524.7	50.3
[uo] (2nd)	514.3	7.8	820.8	21.5	2495.3	79.2

Triphthong	Male speaker 7					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	499.8	14.6	1871.4	85.2	2894.2	192.7
[iai] (2nd)	858.6	42.0	1467.0	41.3	2294.5	36.9
[iai] (3rd)	388.9	20.3	2189.4	35.1	2953.8	76.0
[iui] (1st)	255.6	15.1	2227.0	48.9	3185.2	81.0
[iui] (2nd)	389.4	12.0	1271.4	190.4	2451.0	41.7
[iui] (3rd)	314.9	12.3	2176.5	34.2	2910.1	92.0
[iau] (1st)	266.0	18.0	2076.7	62.6	3408.9	77.3
[iau] (2nd)	792.9	13.7	1214.8	37.5	2266.2	76.7
[iau] (3rd)	415.3	12.6	711.7	35.4	2610.2	129.3
[uai] (1st)	397.3	16.6	824.3	45.7	2549.8	38.5
[uai] (2nd)	804.6	22.9	1333.4	21.5	2216.5	22.0
[uai] (3rd)	381.4	21.2	2155.3	66.5	2927.1	70.1

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 7			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	148.2 (17.0)	175.0 (26.0)	165.3 (31.8)	488.6 (40.5)
[ia]	100.2 (16.0)	107.0 (3.2)	295.2 (32.4)	502.6 (35.7)
[ui]	108.5 (14.0)	96.9 (18.0)	174.7 (31.0)	380.2 (40.1)
[iu]	135.6 (27.4)	129.6 (18.5)	245.0 (28.4)	510.3 (46.2)
[oi]	149.6 (8.7)	175.5 (27.5)	93.4 (17.3)	418.7 (27.0)
[io]	115.3 (23.0)	136.4 (13.8)	262.8 (15.4)	514.6 (16.3)
[au]	168.9 (11.5)	132.2 (29.1)	104.3 (19.4)	405.5 (22.7)
[ua]	48.0 (23.4)	104.8 (5.8)	240.7 (27.4)	393.7 (49.0)
[eu]	192.0 (30.5)	123.0 (9.1)	92.8 (21.7)	407.9 (27.7)
[ie]	130.8 (21.6)	94.1 (11.7)	340.6 (25.8)	565.5 (23.9)
[uo]	67.3 (18.0)	92.4 (13.9)	274.0 (25.6)	433.8 (39.0)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 7					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	113.2 (23.3)	95.6 (8.5)	105.1 (17.2)	114.0 (12.9)	90.7 (13.8)	518.9 (37.6)
[uai]	44.5 (8.4)	95.3 (13.9)	74.9 (13.5)	167.6 (28.8)	99.7 (27.5)	482.1 (56.2)
[iui]	100.5 (14.1)	81.7 (14.8)	60.3 (6.0)	100.0 (37.3)	176.9 (27.6)	519.6 (43.8)
[iai]	0.0 (0.0)	66.8 (5.8)	92.8 (32.3)	139.2 (23.2)	84.7 (9.1)	383.7 (15.9)

Vowel	Male speaker 8							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	231.2	2.7	2559.7	25.2	3510.0	32.7	558.7	84.2
[ɛ]	483.0	25.9	2261.6	48.7	2885.9	130.7	438.0	34.2
[a]	959.7	10.0	1566.0	22.8	3013.1	64.0	516.6	63.5
[o]	457.3	9.5	775.2	35.0	3112.4	33.1	442.0	14.2
[u]	328.7	17.6	724.4	54.3	2931.3	44.3	369.9	50.3
[ʅ]	330.0	16.9	1459.1	22.7	3037.4	28.7	404.9	31.3

Diphthong	Male speaker 8					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	954.3	9.8	1556.8	16.4	2965.5	27.6
[ai] (2nd)	376.3	43.5	2544.3	33.2	3203.7	117.2
[ia] (1st)	230.9	4.2	2341.5	55.6	3405.1	46.6
[ia] (2nd)	953.1	20.7	1538.1	13.1	2869.9	30.2
[ui] (1st)	340.5	14.5	881.2	40.3	3032.7	38.7
[ui] (2nd)	275.9	2.6	2465.0	69.3	3264.1	37.5
[iu] (1st)	238.5	26.6	2436.1	14.4	3288.9	68.9
[iu] (2nd)	312.7	8.0	912.7	11.8	2759.1	59.7
[oi] (1st)	544.4	13.1	937.4	13.3	3186.0	32.4
[oi] (2nd)	379.4	10.4	2365.8	53.6	2995.7	120.9
[io] (1st)	232.4	12.9	2290.4	60.4	3362.7	78.2
[io] (2nd)	517.2	11.0	875.6	17.4	3183.9	30.5
[au] (1st)	798.3	36.8	1239.7	22.9	3251.2	36.2
[au] (2nd)	529.8	39.6	922.0	47.7	3203.4	69.2
[ua] (1st)	398.0	36.2	1018.7	58.4	2610.9	57.2
[ua] (2nd)	988.5	18.1	1552.5	21.0	2974.4	85.8
[eu] (1st)	532.0	6.4	2109.7	40.6	2755.4	88.4
[eu] (2nd)	438.4	9.1	891.0	17.9	2841.8	140.2
[ie] (1st)	223.0	7.5	2422.6	18.5	3518.9	69.7
[ie] (2nd)	528.4	22.7	2096.7	43.8	2773.8	48.8
[uo] (1st)	328.0	4.6	813.5	30.5	2782.6	72.3
[uo] (2nd)	558.9	33.4	872.6	38.7	3168.4	71.0

Triphthong	Male speaker 8					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	432.6	24.2	2221.8	63.2	2858.7	50.3
[iai] (2nd)	864.1	31.4	1672.1	30.7	2751.7	70.8
[iai] (3rd)	430.7	11.3	2409.5	66.2	2940.2	101.6
[iui] (1st)	213.6	6.7	2361.6	27.3	3449.3	53.3
[iui] (2nd)	326.1	24.3	1028.2	80.2	2865.2	64.2
[iui] (3rd)	266.9	7.4	2469.2	37.0	3235.5	23.9
[iau] (1st)	225.3	3.9	2276.9	45.1	3379.3	63.5
[iau] (2nd)	834.1	9.8	1275.7	22.2	3085.5	101.8
[iau] (3rd)	442.0	23.8	887.9	30.1	3093.5	88.2
[uai] (1st)	370.4	10.5	964.9	64.1	2640.8	34.2
[uai] (2nd)	890.6	26.6	1502.8	77.3	2866.3	70.9
[uai] (3rd)	363.3	16.9	2570.1	39.4	3207.3	90.8

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 8			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	141.3 (13.8)	216.6 (37.4)	189.8 (31.5)	547.8 (48.1)
[ia]	104.6 (24.3)	109.8 (5.6)	360.5 (56.0)	575.0 (60.0)
[ui]	90.0 (21.8)	134.7 (25.5)	208.4 (21.1)	433.2 (28.5)
[iu]	217.6 (20.5)	135.2 (27.1)	232.2 (28.2)	585.1 (64.3)
[oi]	130.4 (6.9)	198.2 (6.0)	167.1 (19.3)	495.8 (15.7)
[io]	82.0 (5.8)	112.4 (19.0)	321.1 (49.6)	515.7 (47.6)
[au]	93.2 (33.0)	170.4 (44.4)	172.7 (27.4)	436.4 (52.7)
[ua]	46.2 (11.8) ③	121.4 (36.5)	322.3 (24.7)	478.5 (37.6)
[eu]	108.3 (24.5)	123.0 (10.3)	171.7 (30.9)	403.1 (38.1)
[ie]	126.9 (16.8)	100.6 (11.9)	326.8 (48.1)	554.4 (48.6)
[uo]	43.7 (11.2)	70.3 (7.1)	382.0 (51.0)	496.1 (53.2)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 8					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	119.1 (20.6)	116.5 (24.6)	73.4 (10.5)	155.8 (5.0)	204.9 (26.0)	669.9 (41.5)
[uai]	41.2 (12.2)	87.2 (15.5)	97.9 (19.0)	185.4 (26.2)	163.5 (26.5)	575.4 (10.1)
[iui]	100.0 (6.9)	75.7 (14.3)	53.4 (16.3)	113.8 (19.1)	166.8 (27.6)	510.0 (41.9)
[iai]	0.0 (0.0)	86.4 (4.5)	82.9 (13.8)	153.8 (20.6)	178.2 (14.6)	501.5 (28.5)

Vowel	Male speaker 9							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	246.0	5.5	2274.0	79.1	3231.3	43.0	444.3	45.9
[ɛ]	469.7	16.2	2107.8	81.6	2762.6	38.1	380.5	34.3
[a]	802.1	16.7	1411.8	32.4	2254.1	34.4	395.7	29.8
[o]	489.0	25.2	731.8	4.1	2694.5	77.9	391.3	43.8
[u]	338.7	12.3	731.3	31.1	2675.0	25.4	343.2	44.9
[ʅ]	316.3	24.3	1103.4	22.4	2912.7	15.4	294.6	23.9

Diphthong	Male speaker 9					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	777.4	10.9	1458.8	21.9	2295.7	208.3
[ai] (2nd)	361.7	21.8	2034.6	53.6	2728.6	55.5
[ia] (1st)	292.2	18.7	2217.6	93.9	2935.3	62.7
[ia] (2nd)	836.0	13.3	1483.9	29.6	2073.6	99.7
[ui] (1st)	412.6	26.9	1017.3	42.0	2695.5	104.4
[ui] (2nd)	306.9	18.5	2217.8	132.7	2844.7	56.7
[iu] (1st)	291.9	10.2	2175.3	27.5	3016.2	46.1
[iu] (2nd)	312.2	25.0	782.8	44.4	2589.9	35.8
[oi] (1st)	612.0	9.0	933.2	19.1	2505.9	101.7
[oi] (2nd)	269.2	15.5	2023.9	60.9	2786.0	38.0
[io] (1st)	264.9	6.6	2308.5	57.7	3078.4	156.3
[io] (2nd)	562.3	14.4	864.0	27.1	2581.8	41.1
[au] (1st)	794.7	8.9	1203.7	29.9	2241.0	250.8
[au] (2nd)	418.3	63.0	748.9	45.7	2485.1	45.4
[ua] (1st)	373.8	21.6	696.3	43.8	2618.4	123.0
[ua] (2nd)	770.4	8.0	1298.6	37.6	1973.9	68.7
[eu] (1st)	556.6	17.5	1825.8	40.1	2627.8	57.1
[eu] (2nd)	324.8	47.0	946.3	86.2	2510.4	42.0
[ie] (1st)	273.5	14.5	2245.7	68.1	2975.5	195.2
[ie] (2nd)	413.4	8.1	2013.7	12.3	2860.5	71.1
[uo] (1st)	309.0	11.2	690.8	42.8	2496.7	45.1
[uo] (2nd)	559.5	41.3	862.3	5.9	2607.2	88.1

Triphthong	Male speaker 9					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	429.8	21.9	1652.0	30.7	2828.9	119.9
[iai] (2nd)	763.4	17.1	1482.0	15.4	2208.1	36.8
[iai] (3rd)	413.4	35.7	1934.0	54.9	2721.3	261.1
[iui] (1st)	267.3	15.5	2271.6	63.8	3009.1	80.0
[iui] (2nd)	426.6	10.2	1098.7	71.5	2516.9	114.1
[iui] (3rd)	278.2	26.8	2193.4	58.5	2866.4	129.9
[iau] (1st)	277.4	21.2	2292.8	76.7	3075.9	77.0
[iau] (2nd)	796.8	18.7	1327.1	77.6	2126.4	132.8
[iau] (3rd)	399.2	28.4	718.1	58.5	2721.0	218.9
[uai] (1st)	341.7	26.0	700.9	30.1	2499.2	83.3
[uai] (2nd)	807.7	40.9	1375.1	50.5	2108.9	106.5
[uai] (3rd)	334.1	67.6	2112.2	36.0	2677.5	86.3

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 9			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	113.5 (21.4)	148.5 (30.0)	120.0 (12.0)	382.1 (30.8)
[ia]	76.1 (18.4)	111.8 (20.7)	248.0 (33.6)	436.0 (41.9)
[ui]	98.5 (7.0)	117.3 (13.8)	144.6 (26.0)	360.5 (31.2)
[iu]	246.5 (55.0)	131.3 (20.2)	143.8 (19.8)	521.8 (72.1)
[oi]	136.1 (11.9)	156.3 (17.8)	91.5 (25.3)	383.9 (51.3)
[io]	88.1 (23.8)	121.5 (17.0)	184.6 (33.3)	394.2 (54.6)
[au]	115.2 (33.0)	116.6 (18.1)	69.7 (18.5)	301.6 (55.8)
[ua]	38.1 (4.8)	141.4 (19.7)	233.0 (50.6)	412.6 (51.9)
[eu]	169.6 (35.9)	109.3 (26.1)	74.1 (35.3)	353.1 (83.4)
[ie]	78.2 (30.0)	99.9 (7.4)	288.7 (30.0)	466.8 (48.5)
[uo]	37.8 (8.2)	103.9 (10.9)	291.1 (35.3)	433.0 (32.5)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 9					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	115.0 (21.9)	94.4 (15.7)	60.8 (6.1)	145.8 (9.1)	110.8 (9.8)	527.0 (41.7)
[uai]	27.8 (2.4)	123.3 (15.7)	60.9 (19.9)	117.0 (24.6)	115.9 (6.6)	445.0 (48.2)
[iui]	81.7 (18.8)	75.6 (18.9)	54.9 (11.0)	98.2 (22.2)	123.3 (35.6)	433.9 (42.8)
[iai]	0.0 (0.0)	64.8 (11.0)	83.3 (22.6)	93.3 (25.1)	50.2 (25.1)	291.7 (29.3)

Vowel	Male speaker 10							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	295.8	9.1	1832.2	35.8	2998.0	38.6	427.3	75.6
[ɛ]	504.3	16.6	1774.8	6.5	2391.9	46.7	362.8	47.7
[a]	868.5	34.2	1363.4	29.7	2205.5	68.9	370.0	30.6
[o]	541.1	9.9	842.5	28.9	2353.9	39.4	411.7	27.6
[u]	416.7	9.4	844.6	30.5	2248.7	68.1	339.2	53.5
[ʅ]	389.1	5.3	1331.4	32.4	2382.2	31.7	308.6	37.3

Diphthong	Male speaker 10					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	868.9	11.5	1346.7	23.6	2278.8	103.9
[ai] (2nd)	403.4	12.2	1873.4	27.4	2782.2	89.0
[ia] (1st)	309.7	19.1	1757.4	19.6	3050.8	31.7
[ia] (2nd)	861.9	54.4	1387.4	34.4	2189.0	115.7
[ui] (1st)	444.0	6.0	1092.4	36.7	2311.5	32.7
[ui] (2nd)	371.1	10.2	1851.7	15.3	2469.0	187.0
[iu] (1st)	325.9	10.1	1800.3	18.4	2843.9	32.4
[iu] (2nd)	427.4	17.1	1016.6	47.8	2184.0	86.3
[oi] (1st)	633.9	38.6	1036.1	73.7	2460.4	135.0
[oi] (2nd)	440.2	13.5	1676.7	61.6	2351.7	112.1
[io] (1st)	299.7	4.2	1754.2	118.5	2950.5	143.3
[io] (2nd)	563.6	9.7	940.3	53.6	2452.5	137.0
[au] (1st)	790.3	18.1	1214.9	33.9	2380.6	185.7
[au] (2nd)	570.3	19.2	904.8	27.3	2422.3	134.6
[ua] (1st)	543.9	38.1	945.0	50.1	2140.5	44.3
[ua] (2nd)	887.3	16.4	1306.5	20.2	2166.0	44.4
[eu] (1st)	646.8	20.2	1620.3	48.4	2401.6	75.0
[eu] (2nd)	470.7	8.5	1141.2	21.5	2112.1	78.9
[ie] (1st)	299.2	7.9	1887.4	80.7	3027.3	45.0
[ie] (2nd)	519.6	54.7	1837.8	41.7	2522.5	57.8
[uo] (1st)	440.9	4.9	893.1	50.7	2045.6	30.2
[uo] (2nd)	578.2	12.3	871.8	23.6	2387.7	34.4

Triphthong	Male speaker 10					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	501.5	43.1	1626.0	34.2	2805.1	61.7
[iai] (2nd)	764.7	52.7	1457.3	58.0	2176.0	197.9
[iai] (3rd)	530.1	60.8	1765.1	30.8	2470.0	41.5
[iui] (1st)	301.0	7.5	1823.9	50.0	2948.4	34.7
[iui] (2nd)	435.6	19.0	1123.1	34.1	2248.8	100.4
[iui] (3rd)	349.1	12.3	1825.2	27.7	2523.0	129.2
[iau] (1st)	301.7	8.4	1763.0	63.4	3048.6	18.9
[iau] (2nd)	853.6	8.4	1225.7	25.4	2272.1	98.5
[iau] (3rd)	565.8	14.5	884.1	6.6	2410.8	116.0
[uai] (1st)	523.0	23.6	964.2	41.2	2048.5	61.3
[uai] (2nd)	866.9	7.0	1324.0	15.2	2161.0	37.9
[uai] (3rd)	450.2	39.1	1822.6	17.6	2596.8	139.0

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Male speaker 10			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	195.9 (47.0)	144.8 (24.5)	126.4 (16.2)	467.2 (67.2)
[ia]	104.4 (15.2)	118.6 (19.8)	206.9 (13.3)	430.0 (17.4)
[ui]	95.5 (7.5)	93.5 (31.9)	134.0 (8.1)	323.1 (34.2)
[iu]	237.7 (20.9)	132.5 (22.2)	112.2 (11.9)	482.5 (51.6)
[oi]	141.0 (28.3)	119.8 (21.9)	88.7 (26.5)	349.6 (67.0)
[io]	85.3 (19.0)	108.0 (40.8)	169.5 (26.5)	362.8 (62.9)
[au]	92.6 (26.6)	101.1 (8.2)	87.4 (5.0)	281.1 (24.4)
[ua]	23.5 (9.1)	95.8 (14.1)	210.2 (45.2)	329.7 (44.6)
[eu]	140.0 (38.3)	97.2 (18.7)	91.3 (16.9)	328.6 (50.2)
[ie]	146.3 (37.0)	113.8 (25.6)	357.1 (57.1)	617.3 (103.6)
[uo]	30.6 (6.9)	54.9 (13.7)	257.9 (61.4)	343.5 (74.5)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Male speaker 10					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	60.7 (3.3)	93.3 (5.3)	73.5 (21.1)	118.9 (31.0)	99.6 (18.0)	446.3 (44.4)
[uai]	24.8 (0.4) ③	92.1 (21.0)	80.4 (7.6)	109.3 (38.0)	108.0 (24.0)	404.9 (61.3)
[iui]	70.1 (15.4)	65.2 (5.9)	59.5 (8.7)	85.6 (9.8)	130.2 (17.6)	410.7 (27.9)
[iai]	0.0 (0.0)	54.1 (5.1)	67.2 (10.1)	94.2 (15.7)	78.0 (15.8)	293.7 (21.3)

Vowel	Female speaker 1							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	274.6	31.3	2740.1	24.1	3477.2	71.6	449.2	53.4
[ɛ]	626.7	8.7	2330.8	81.9	2926.3	93.1	404.9	7.2
[a]	1249.6	56.6	1702.0	27.3	3059.7	72.9	358.0	35.8
[o]	638.8	19.3	972.3	81.8	3276.6	53.8	385.8	22.4
[u]	450.3	13.0	885.3	63.6	2888.6	251.4	299.0	15.8
[ʅ]	433.5	23.0	1328.2	52.4	3142.0	115.4	317.9	16.1

Diphthong	Female speaker 1					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1259.9	82.8	1717.8	46.1	3072.4	137.8
[ai] (2nd)	463.3	27.0	2587.1	62.5	3200.0	45.7
[ia] (1st)	283.1	14.4	2863.8	43.6	3759.5	106.6
[ia] (2nd)	1227.6	73.9	1807.6	11.6	2786.0	109.7
[ui] (1st)	568.3	34.0	1200.1	145.1	3102.8	31.9
[ui] (2nd)	377.2	18.9	2557.2	24.6	3105.9	33.4
[iu] (1st)	288.8	41.6	2731.9	51.7	3574.4	183.4
[iu] (2nd)	466.3	48.4	981.5	85.1	2804.1	59.9
[oi] (1st)	821.4	35.4	1251.8	179.3	3222.6	129.3
[oi] (2nd)	522.3	35.8	2367.0	87.4	3108.9	85.9
[io] (1st)	288.0	38.0	2702.3	15.3	3563.1	138.7
[io] (2nd)	773.0	32.8	1097.6	74.0	3198.5	63.5
[au] (1st)	1164.5	27.9	1614.8	51.3	2918.4	167.8
[au] (2nd)	695.6	54.3	973.2	82.4	3116.5	134.0
[ua] (1st)	454.4	57.2	926.1	141.1	2836.1	52.2
[ua] (2nd)	1246.8	35.1	1689.2	28.1	2984.0	126.3
[eu] (1st)	795.6	54.5	2166.0	57.0	3004.5	143.1
[eu] (2nd)	534.5	54.9	986.6	56.0	2825.2	159.1
[ie] (1st)	246.6	14.8	2936.6	38.1	3981.3	100.1
[ie] (2nd)	590.2	38.1	2468.5	39.5	3172.4	60.9
[uo] (1st)	394.5	48.8	853.7	77.7	2763.4	134.8
[uo] (2nd)	686.3	45.1	1239.7	101.1	3113.4	83.6

Triphthong	Female speaker 1					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1^{st})	545.8	45.1	2200.2	80.2	3068.8	86.3
[iai] (2^{nd})	1086.5	58.7	1831.3	98.9	2825.7	49.1
[iai] (3^{rd})	566.7	72.0	2249.5	35.4	3069.4	80.2
[iui] (1^{st})	276.9	18.8	2638.5	82.7	3634.5	163.1
[iui] (2^{nd})	544.7	51.4	1132.4	69.6	2805.5	77.1
[iui] (3^{rd})	383.4	47.3	2462.9	160.3	3045.0	97.8
[iau] (1^{st})	257.8	10.5	2784.3	28.0	3777.4	168.5
[iau] (2^{nd})	1101.1	17.2	1650.9	81.4	2743.3	86.7
[iau] (3^{rd})	631.3	29.9	961.1	41.9	3155.3	76.9
[uai] (1^{st})	435.9	31.4	870.6	81.4	2788.2	202.1
[uai] (2^{nd})	1032.2	37.9	1590.4	82.7	2683.6	155.5
[uai] (3^{rd})	411.0	18.0	2672.3	80.0	3266.8	80.1

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 1			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	122.8 (6.6)	143.2 (13.3)	185.4 (34.9)	451.5 (31.3)
[ia]	110.0 (15.4)	104.8 (27.0)	274.8 (32.2)	489.7 (42.9)
[ui]	101.8 (6.1)	129.1 (9.9)	124.0 (21.3)	355.0 (15.0)
[iu]	224.2 (20.1)	150.4 (27.3)	154.2 (4.9)	529.0 (33.6)
[oi]	141.1 (8.0)	120.9 (27.3)	111.1 (10.2)	373.2 (13.3)
[io]	138.9 (20.1)	94.1 (15.6)	268.0 (6.4)	501.1 (17.4)
[au]	153.9 (14.5)	73.4 (29.9)	162.2 (37.3)	389.5 (34.0)
[ua]	58.5 (13.3)	119.1 (33.9)	280.2 (9.9)	457.9 (32.6)
[eu]	171.9 (11.0)	116.0 (7.1)	122.4 (9.7)	410.4 (14.4)
[ie]	120.8 (5.4)	65.6 (9.6)	310.7 (7.5)	497.3 (12.2)
[uo]	100.1 (29.9)	57.2 (25.4)	323.0 (18.8)	480.4 (27.2)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 1					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	97.6 (5.8)	87.2 (5.4)	106.9 (8.2)	77.0 (27.3)	144.3 (27.7)	513.3 (22.1)
[uai]	38.2 (2.7)	100.5 (14.3)	98.2 (15.2)	120.6 (8.0)	126.4 (25.4)	484.1 (30.5)
[iui]	110.6 (8.0)	93.0 (11.7)	45.2 (12.1)	112.4 (16.8)	102.1 (16.4)	463.4 (30.8)
[iai]	0.0 (0.0)	57.5 (14.2)	79.8 (26.9)	114.4 (20.4)	106.0 (17.1)	357.8 (8.8)

Vowel	Female speaker 2							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	305.5	26.7	2808.9	66.7	3381.0	190.3	556.1	52.3
[ɛ]	569.0	39.7	2379.1	64.7	2899.3	34.3	489.6	39.5
[a]	1094.9	56.5	1539.6	59.9	2740.3	58.8	460.7	28.6
[o]	609.0	19.8	989.4	102.5	3233.4	92.8	521.9	16.2
[u]	436.5	22.3	933.8	63.8	2876.6	55.5	512.7	64.4
[ʅ]	415.2	20.9	1372.9	20.2	2980.1	55.3	469.3	24.9

Diphthong	Female speaker 2					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1107.1	35.7	1601.3	31.4	2712.6	36.3
[ai] (2nd)	320.5	48.9	2666.8	78.1	3234.5	54.3
[ia] (1st)	325.9	19.4	2542.3	145.8	3384.7	127.8
[ia] (2nd)	1071.9	24.8	1635.2	34.8	2779.1	115.3
[ui] (1st)	528.2	4.5	1076.4	56.6	2981.6	68.0
[ui] (2nd)	370.1	33.4	2542.6	44.8	3033.4	35.2
[iu] (1st)	229.2	29.7	2767.8	40.3	3931.0	64.9
[iu] (2nd)	452.3	31.5	910.6	25.8	2858.8	161.0
[oi] (1st)	740.0	33.7	1125.4	70.9	3290.7	27.9
[oi] (2nd)	448.5	32.0	2376.8	79.3	3016.1	108.6
[io] (1st)	253.8	18.5	2663.2	94.9	3413.1	128.5
[io] (2nd)	587.7	109.0	916.9	80.4	3214.8	41.1
[au] (1st)	1033.2	44.2	1465.2	50.6	3003.8	98.9
[au] (2nd)	601.6	42.0	926.1	74.5	3186.6	44.1
[ua] (1st)	486.9	61.1	922.6	63.2	2670.5	64.3
[ua] (2nd)	1101.2	26.3	1591.6	45.4	2730.9	120.0
[eu] (1st)	755.5	7.7	2146.4	21.7	2825.5	54.4
[eu] (2nd)	479.9	22.5	917.1	53.2	2879.8	180.7
[ie] (1st)	265.9	20.4	2799.5	78.8	3708.7	134.8
[ie] (2nd)	647.8	39.2	2455.8	38.7	3126.0	48.7
[uo] (1st)	380.6	42.9	816.2	63.1	2625.6	32.6
[uo] (2nd)	614.0	31.2	1034.4	73.4	3040.7	108.5

Triphthong	Female speaker 2					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	565.6	19.8	2212.6	121.1	2670.1	99.6
[iai] (2nd)	1035.7	50.2	1682.0	57.0	2619.7	75.8
[iai] (3rd)	417.6	25.1	2679.1	120.8	3432.5	200.4
[iui] (1st)	238.4	5.3	2629.8	52.6	3260.5	94.1
[iui] (2nd)	498.4	45.6	1383.7	123.9	2851.9	63.1
[iui] (3rd)	335.0	9.8	2556.3	44.3	3083.0	36.1
[iau] (1st)	269.0	32.4	2612.1	66.9	3394.9	45.2
[iau] (2nd)	1083.2	45.5	1494.0	48.5	2661.5	52.6
[iau] (3rd)	759.0	75.8	1114.6	29.9	3111.8	120.2
[uai] (1st)	485.7	28.8	906.2	58.3	2644.4	40.9
[uai] (2nd)	1050.7	55.4	1512.9	24.5	2545.9	39.3
[uai] (3rd)	366.1	36.4	2593.9	28.7	3120.4	131.7

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 2			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	118.0 (11.5)	205.5 (44.1)	261.7 (32.9)	585.4 (50.2)
[ia]	102.0 (4.9)	103.1 (7.5)	364.3 (26.6)	569.4 (20.5)
[ui]	112.8 (14.0)	121.8 (19.7)	269.2 (35.2)	503.9 (59.6)
[iu]	209.3 (38.2)	248.1 (22.1)	207.4 (42.0)	664.9 (48.3)
[oi]	168.5 (19.5)	211.5 (38.4)	191.6 (30.8)	571.7 (13.7)
[io]	121.1 (29.2)	91.2 (16.8)	344.4 (41.9)	556.9 (70.8)
[au]	160.8 (25.8)	131.4 (27.6)	264.7 (21.0)	557.0 (22.0)
[ua]	18.6 (undefined) ①	110.5 (28.4)	370.0 (39.0)	485.3 (50.1)
[eu]	143.3 (20.7)	121.2 (13.0)	146.5 (31.3)	411.1 (51.1)
[ie]	128.3 (30.4)	77.1 (25.9)	397.0 (31.5)	602.5 (43.0)
[uo]	31.3 (1.7) ③	69.3 (25.5)	398.2 (24.3)	486.3 (27.4)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 2					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	65.2 (15.9)	121.1 (13.5)	99.3 (25.4)	124.6 (6.7)	211.5 (20.0)	621.9 (27.1)
[uai]	0.0 (0.0)	119.4 (9.5)	69.5 (2.7)	223.9 (34.9)	186.4 (29.7)	599.4 (32.5)
[iui]	97.8 (21.0)	97.0 (9.9)	46.1 (8.5)	142.3 (34.1)	181.8 (27.4)	565.1 (26.3)
[iai]	0.0 (0.0)	65.4 (8.6)	78.2 (3.8)	184.5 (34.4)	98.7 (31.6)	427.0 (60.0)

Vowel	Female speaker 3							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	335.5	26.9	2730.1	57.6	3500.9	41.6	593.6	44.6
[ɛ]	752.9	58.2	2431.2	105.0	3435.5	102.9	541.2	27.0
[a]	1240.5	37.8	1686.4	41.8	2730.3	48.9	531.2	32.9
[o]	664.1	14.8	976.6	61.5	3092.9	32.8	588.0	40.5
[u]	428.7	14.5	910.8	19.1	2992.5	51.5	568.2	21.0
[ɿ]	406.1	8.0	1305.7	13.9	2978.4	108.5	514.0	34.8

Diphthong	Female speaker 3					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1286.4	26.1	1722.6	35.5	2757.2	51.6
[ai] (2nd)	508.2	28.0	2640.0	84.6	3152.5	238.5
[ia] (1st)	332.5	18.4	2551.5	19.2	3383.9	26.1
[ia] (2nd)	1122.1	20.3	1642.9	34.5	2554.2	23.2
[ui] (1st)	545.9	37.7	1005.0	41.3	3025.3	45.6
[ui] (2nd)	415.3	8.4	2685.1	25.6	3080.9	28.0
[iu] (1st)	406.4	12.8	2576.5	5.8	3219.2	11.9
[iu] (2nd)	465.5	52.3	988.9	83.1	2943.3	46.1
[oi] (1st)	777.7	35.7	1144.9	39.5	3091.0	120.9
[oi] (2nd)	432.4	14.8	2625.5	84.4	2998.4	121.4
[io] (1st)	293.4	10.8	2506.9	57.8	3221.0	101.9
[io] (2nd)	727.5	37.7	1185.8	67.7	2883.0	67.4
[au] (1st)	1082.0	39.4	1548.9	48.4	2693.6	31.6
[au] (2nd)	603.7	59.6	946.8	109.4	3012.6	100.3
[ua] (1st)	503.8	22.5	881.1	36.0	2689.7	45.6
[ua] (2nd)	1163.7	34.8	1663.0	57.7	2639.3	43.7
[eu] (1st)	749.2	31.7	2142.5	50.4	3009.3	286.0
[eu] (2nd)	519.3	10.8	1044.6	73.0	3050.3	288.9
[ie] (1st)	287.2	9.2	2580.6	38.2	3426.7	28.8
[ie] (2nd)	645.6	50.3	2336.5	18.3	2897.9	88.5
[uo] (1st)	448.6	32.3	816.2	34.8	2740.0	106.1
[uo] (2nd)	755.0	26.0	1122.8	52.8	2943.3	86.3

Triphthong	Female speaker 3					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	499.2	20.0	2171.2	70.9	2732.6	156.3
[iai] (2nd)	1123.3	60.1	1726.8	27.4	2683.8	117.2
[iai] (3rd)	468.3	27.2	2608.2	105.7	3134.8	74.8
[iui] (1st)	276.8	39.4	2478.2	46.5	3256.4	110.3
[iui] (2nd)	515.0	24.4	1075.2	50.6	3015.7	39.2
[iui] (3rd)	411.3	21.8	2657.4	72.0	3108.8	82.9
[iau] (1st)	313.5	11.7	2523.5	32.1	3399.9	76.3
[iau] (2nd)	1086.8	34.3	1552.9	98.1	2697.2	127.8
[iau] (3rd)	599.1	46.3	1015.3	43.1	2856.2	52.6
[uai] (1st)	503.0	50.9	899.1	84.6	2699.0	6.9
[uai] (2nd)	1098.5	19.4	1580.2	65.6	2623.5	84.3
[uai] (3rd)	488.1	22.7	2651.0	37.9	3098.6	113.7

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 3			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	224.1 (14.8)	228.8 (28.3)	163.8 (33.6)	616.8 (37.0)
[ia]	75.5 (17.3)	124.2 (17.4)	321.5 (35.3)	521.3 (58.7)
[ui]	240.0 (56.8)	197.0 (22.7)	148.3 (43.0)	585.4 (37.4)
[iu]	287.3 (23.7)	196.4 (34.3)	185.9 (18.6)	669.7 (25.8)
[oi]	255.3 (19.6)	205.9 (12.0)	150.2 (24.5)	611.4 (22.0)
[io]	57.8 (19.6)	101.9 (19.5)	380.7 (32.1)	540.6 (44.5)
[au]	226.4 (40.2)	211.8 (44.8)	142.1 (61.7)	580.4 (64.6)
[ua]	40.6 (12.4) ③	163.3 (21.6)	384.7 (46.6)	572.4 (45.7)
[eu]	248.5 (20.5)	129.6 (24.5)	88.7 (11.2)	466.9 (30.2)
[ie]	86.7 (8.5)	113.2 (24.3)	446.3 (18.1)	646.3 (41.7)
[uo]	27.4 (8.2)	127.5 (26.4)	425.4 (53.1)	580.3 (49.9)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 3					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	63.6 (5.8)	107.1 (18.4)	166.8 (9.7)	159.6 (12.5)	166.7 (13.4)	664.0 (25.3)
[uai]	25.0 (undefined) ①	161.6 (25.9)	138.3 (43.5)	175.4 (21.5)	151.0 (25.4)	632.7 (16.4)
[iui]	60.3 (15.2)	87.3 (7.9)	147.0 (40.6)	160.8 (17.8)	182.7 (36.2)	638.3 (17.5)
[iai]	0.0 (0.0)	65.7 (26.6)	172.9 (33.3)	164.0 (30.5)	124.9 (21.6)	527.7 (16.4)

Vowel	Female speaker 4							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	303.8	17.8	2749.1	63.7	3674.7	86.5	387.7	19.8
[ɛ]	625.9	23.8	2510.8	55.0	3171.4	110.9	375.1	59.1
[a]	1153.0	16.9	1731.4	44.1	2936.5	136.3	405.4	50.0
[o]	680.4	34.9	1048.4	42.0	2717.1	135.6	415.1	52.2
[u]	438.0	19.1	860.9	82.8	2708.4	100.4	308.6	34.2
[ʅ]	374.5	34.7	1408.5	95.6	2763.3	62.4	349.0	43.4

Diphthong	Female speaker 4					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1086.5	25.0	1700.6	21.8	2880.7	191.4
[ai] (2nd)	458.1	27.5	2601.3	91.8	3435.7	112.9
[ia] (1st)	363.9	40.2	2523.6	64.1	3861.0	139.9
[ia] (2nd)	1185.7	45.3	1779.8	48.6	2636.6	89.1
[ui] (1st)	574.3	29.9	1440.3	120.9	2948.5	92.6
[ui] (2nd)	397.6	33.6	2683.1	23.4	3297.7	71.1
[iu] (1st)	309.1	9.0	2599.7	44.0	3447.2	171.7
[iu] (2nd)	428.4	9.9	1011.0	59.2	2772.5	94.4
[oi] (1st)	848.5	68.6	1288.2	41.0	2686.1	94.2
[oi] (2nd)	601.2	37.3	2476.7	89.6	3174.1	125.4
[io] (1st)	370.0	25.8	2553.7	29.8	3675.4	129.5
[io] (2nd)	844.9	49.7	1290.0	28.8	2821.3	173.6
[au] (1st)	1035.4	35.6	1572.6	28.0	2574.2	73.1
[au] (2nd)	764.2	57.7	1152.3	58.3	2653.3	91.9
[ua] (1st)	487.4	38.1	965.5	42.4	2532.1	62.5
[ua] (2nd)	1123.1	20.6	1687.8	48.0	2618.0	69.0
[eu] (1st)	852.0	20.1	2191.9	100.0	2999.0	123.1
[eu] (2nd)	632.8	17.4	1155.5	85.9	2529.4	135.9
[ie] (1st)	303.5	25.7	2693.4	74.5	3748.7	104.3
[ie] (2nd)	573.7	25.0	2484.5	67.2	3191.9	81.0
[uo] (1st)	466.0	15.6	913.3	28.8	2567.9	148.7
[uo] (2nd)	654.6	18.8	1035.1	55.6	2702.0	110.7

Triphthong	Female speaker 4					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1^{st})	555.2	31.3	2223.6	53.3	3291.6	94.9
[iai] (2^{nd})	1106.5	28.4	1821.4	48.6	2487.3	155.8
[iai] (3^{rd})	775.7	40.8	2287.3	88.6	3153.8	97.9
[iui] (1^{st})	329.4	9.0	2557.2	49.2	3420.8	110.5
[iui] (2^{nd})	560.3	56.0	1427.0	168.4	2819.6	41.7
[iui] (3^{rd})	431.0	27.9	2529.4	96.7	3038.7	126.5
[iau] (1^{st})	366.1	9.0	2482.7	87.4	3606.3	137.5
[iau] (2^{nd})	1013.5	9.4	1500.7	34.4	2698.9	134.3
[iau] (3^{rd})	668.1	35.5	1072.4	45.4	2619.6	99.0
[uai] (1^{st})	560.5	38.4	960.4	24.8	2711.6	47.7
[uai] (2^{nd})	1065.3	51.8	1656.8	29.2	2605.5	40.9
[uai] (3^{rd})	573.1	74.4	2413.1	123.3	3362.0	82.8

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 4			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	53.1 (14.2)	141.3 (26.6)	156.8 (53.7)	351.3 (34.1)
[ia]	54.4 (15.6)	96.3 (22.7)	249.3 (48.0)	400.1 (42.9)
[ui]	39.6 (12.4)	67.0 (10.9)	101.6 (29.3)	208.3 (43.0)
[iu]	152.3 (32.0)	129.4 (20.7)	93.3 (17.4)	375.2 (36.6)
[oi]	68.8 (18.8)	89.8 (8.5)	50.5 (19.8)	209.2 (18.7)
[io]	44.8 (6.9)	79.1 (17.4)	142.5 (68.4)	266.5 (71.9)
[au]	82.3 (22.8)	93.7 (17.6)	103.4 (35.4)	279.5 (26.2)
[ua]	30.1 (13.8) ②	89.2 (11.9)	303.6 (52.4)	407.9 (72.4)
[eu]	99.1 (18.9)	91.3 (19.9)	64.6 (20.9)	255.1 (40.2)
[ie]	59.4 (12.6)	87.8 (12.7)	267.0 (55.3)	414.4 (53.4)
[uo]	32.5 (11.6)	40.2 (9.3)	303.0 (21.2)	375.8 (40.4)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 4					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	68.6 (41.1)	89.7 (5.8)	146.4 (45.3)	122.3 (26.5)	63.3 (23.1)	490.5 (71.3)
[uai]	0.0 (0.0)	86.8 (12.3)	99.4 (19.1)	126.0 (12.2)	84.8 (34.8)	397.2 (51.6)
[iui]	60.6 (13.9)	43.5 (12.5)	41.2 (12.9)	66.7 (16.9)	51.4 (13.2)	263.6 (48.1)
[iai]	0.0 (0.0)	46.8 (16.7)	78.9 (44.3)	69.8 (21.7)	48.5 (10.8)	244.1 (41.2)

Vowel	Female speaker 5							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	312.0	25.8	2819.2	88.5	3681.9	135.0	486.1	19.6
[ɛ]	682.8	14.1	2470.4	7.8	3049.8	78.0	464.0	28.4
[a]	1177.9	61.1	1663.9	64.7	2820.9	153.6	474.7	77.0
[o]	707.9	30.3	1037.1	42.1	3368.2	63.7	535.1	43.5
[u]	483.4	14.9	943.1	96.1	3108.4	138.3	515.8	44.9
[ʅ]	491.0	22.5	1435.0	82.2	3423.4	74.5	440.4	19.3

Diphthong	Female speaker 5					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1159.1	105.6	1707.9	45.0	2908.2	84.7
[ai] (2nd)	467.8	25.0	2764.9	42.3	3303.4	53.3
[ia] (1st)	316.6	57.9	2526.6	61.0	3665.8	112.7
[ia] (2nd)	1180.4	3.4	1757.8	41.5	2919.1	83.9
[ui] (1st)	576.6	33.9	1250.4	61.6	3338.6	86.1
[ui] (2nd)	417.8	32.4	2748.3	39.4	3311.5	91.8
[iu] (1st)	312.0	19.2	2730.2	81.0	3452.5	57.7
[iu] (2nd)	463.1	12.9	1111.0	128.5	3073.2	126.0
[oi] (1st)	850.4	22.6	1180.2	23.7	2216.7	156.7
[oi] (2nd)	470.8	56.9	2700.7	84.6	3294.7	85.0
[io] (1st)	299.8	22.0	2654.0	93.8	3663.5	151.8
[io] (2nd)	807.3	46.2	1180.5	54.8	3363.6	53.8
[au] (1st)	1198.4	28.0	1622.6	28.7	3280.8	231.4
[au] (2nd)	610.3	65.8	1019.4	42.5	3565.6	55.1
[ua] (1st)	511.6	11.8	942.4	35.9	3000.9	90.1
[ua] (2nd)	1179.1	34.9	1706.1	43.1	2916.8	155.3
[eu] (1st)	859.1	24.7	2300.9	32.3	3088.7	64.3
[eu] (2nd)	493.6	60.8	1025.2	75.8	3100.9	145.9
[ie] (1st)	273.4	20.4	2788.5	37.3	3872.0	96.9
[ie] (2nd)	599.0	14.5	2438.4	86.8	3089.8	129.3
[uo] (1st)	509.7	9.9	865.8	52.3	3199.5	79.7
[uo] (2nd)	716.1	23.6	1195.7	37.3	3363.8	67.2

Triphthong	Female speaker 5					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	565.5	24.0	2336.7	70.4	3224.8	73.6
[iai] (2nd)	1193.9	52.0	1837.4	58.7	3068.1	64.7
[iai] (3rd)	473.9	92.7	2730.4	85.4	3432.5	83.2
[iui] (1st)	304.7	11.7	2663.8	57.2	3423.5	82.8
[iui] (2nd)	598.0	46.1	1478.8	163.2	3108.7	122.2
[iui] (3rd)	418.6	24.0	2706.6	52.5	3255.2	55.6
[iau] (1st)	293.5	12.1	2582.8	35.0	3797.7	57.0
[iau] (2nd)	1163.7	34.0	1618.8	32.8	3059.7	151.1
[iau] (3rd)	591.2	60.0	1064.0	69.8	3474.0	81.8
[uai] (1st)	504.1	6.0	955.6	41.9	2901.9	122.0
[uai] (2nd)	1154.4	44.9	1710.5	39.1	2875.7	210.6
[uai] (3rd)	450.4	22.8	2752.6	53.5	3280.2	43.3

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 5			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	166.7 (25.1)	147.8 (37.2)	238.5 (47.5)	553.1 (54.3)
[ia]	81.4 (16.7)	109.1 (17.0)	262.4 (19.7)	453.0 (29.4)
[ui]	139.0 (29.1)	177.6 (15.8)	157.6 (14.4)	474.3 (22.9)
[iu]	214.0 (38.4)	200.4 (57.5)	193.2 (33.1)	607.6 (31.1)
[oi]	193.3 (18.8)	161.8 (27.8)	123.5 (14.7)	478.6 (30.0)
[io]	84.5 (12.7)	115.1 (14.5)	293.4 (30.4)	493.2 (31.9)
[au]	172.7 (37.7)	150.9 (27.6)	100.4 (24.0)	424.1 (29.2)
[ua]	34.1 (3.1)	126.8 (22.6)	272.8 (14.4)	433.8 (23.1)
[eu]	193.1 (17.9)	138.2 (20.7)	129.1 (18.4)	460.5 (40.5)
[ie]	102.6 (27.5)	143.6 (27.7)	418.7 (57.5)	665.0 (80.3)
[uo]	87.0 (26.3)	146.1 (45.3)	340.9 (51.2)	552.3 (43.1)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 5					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	104.1 (30.3)	102.7 (12.3)	127.6 (39.7)	162.8 (55.9)	160.0 (31.5)	657.5 (73.5)
[uai]	39.2 (3.1)	108.7 (9.1)	113.6 (15.8)	129.0 (15.7)	127.5 (22.4)	518.2 (32.8)
[iui]	93.0 (17.2)	98.3 (14.0)	82.1 (27.0)	126.5 (19.3)	176.1 (35.1)	576.2 (34.9)
[iai]	0.0 (0.0)	84.4 (19.3)	96.9 (21.8)	145.1 (20.1)	113.2 (34.1)	439.7 (65.2)

Vowel	Female speaker 6							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	278.6	25.8	2730.3	37.6	3638.7	127.8	505.6	54.6
[ɛ]	580.6	18.3	2377.0	20.8	3031.5	33.0	523.7	9.8
[a]	1047.1	36.0	1549.1	44.8	3140.2	66.0	531.0	36.0
[o]	672.0	42.6	920.0	54.1	2937.2	122.0	546.0	38.5
[u]	484.0	11.1	979.8	52.7	3001.6	63.3	413.3	28.4
[ʅ]	437.8	7.7	1455.8	24.7	3208.8	52.3	442.1	34.2

Diphthong	Female speaker 6					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1095.7	50.6	1605.9	13.1	2745.0	122.1
[ai] (2nd)	499.4	19.5	2465.1	56.6	3091.3	131.8
[ia] (1st)	324.6	21.5	2586.4	17.2	3513.6	140.6
[ia] (2nd)	1021.8	33.8	1554.8	51.1	2640.1	79.1
[ui] (1st)	535.4	21.6	1247.3	29.2	2917.4	28.6
[ui] (2nd)	392.3	21.6	2461.4	30.7	3403.7	39.4
[iu] (1st)	413.3	7.6	2517.9	79.7	3187.6	149.5
[iu] (2nd)	445.3	9.5	900.6	13.8	3019.8	124.9
[oi] (1st)	824.8	10.8	1173.3	73.2	3113.2	167.0
[oi] (2nd)	539.0	15.1	2218.6	28.7	3033.3	95.5
[io] (1st)	288.3	29.0	2584.2	45.0	2958.2	55.9
[io] (2nd)	777.7	9.3	1080.5	19.5	3097.8	45.2
[au] (1st)	1066.8	42.5	1410.6	34.9	2472.7	100.6
[au] (2nd)	658.3	44.7	1012.9	82.0	3029.7	101.9
[ua] (1st)	467.8	52.3	935.3	20.0	2752.5	71.0
[ua] (2nd)	1004.3	4.5	1490.4	32.2	2960.2	142.7
[eu] (1st)	794.9	28.8	2073.1	55.3	2736.9	83.1
[eu] (2nd)	544.0	15.7	1000.5	68.0	2897.0	76.4
[ie] (1st)	301.0	23.4	2772.0	89.6	3685.1	54.9
[ie] (2nd)	678.5	24.3	2316.3	57.3	2964.2	87.8
[uo] (1st)	434.2	4.9	888.8	10.5	2775.8	44.2
[uo] (2nd)	757.6	24.9	992.8	14.2	2965.5	86.8

Triphthong	Female speaker 6					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	550.2	11.8	2327.0	45.2	2999.4	152.4
[iai] (2nd)	1067.2	34.3	1803.6	47.4	2817.3	25.3
[iai] (3rd)	572.3	44.4	2381.4	113.6	3012.9	83.9
[iui] (1st)	278.0	14.5	2677.3	24.2	3108.1	78.5
[iui] (2nd)	516.6	33.9	1243.5	31.4	2918.2	29.5
[iui] (3rd)	351.9	13.2	2437.1	33.2	2905.1	51.8
[iau] (1st)	314.0	9.1	2576.8	62.1	3840.5	71.1
[iau] (2nd)	958.1	28.6	1342.3	39.9	3032.3	99.8
[iau] (3rd)	533.7	15.6	886.1	42.6	3099.3	107.6
[uai] (1st)	514.8	6.9	852.7	14.0	2774.5	106.9
[uai] (2nd)	992.3	28.5	1531.1	52.3	2603.8	99.7
[uai] (3rd)	469.2	28.3	2489.6	76.1	3249.3	99.4

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 6			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	130.0 (7.4)	248.4 (10.1)	179.2 (7.6)	557.7 (10.2)
[ia]	76.4 (12.5)	122.3 (16.2)	300.3 (24.4)	499.1 (26.8)
[ui]	90.3 (14.0)	97.9 (22.1)	144.8 (38.3)	333.1 (34.5)
[iu]	207.7 (43.0)	177.5 (22.3)	231.5 (33.5)	616.8 (24.8)
[oi]	235.8 (37.6)	192.8 (22.5)	132.0 (23.9)	560.6 (37.4)
[io]	72.1 (22.6)	110.4 (21.1)	297.4 (33.4)	480.0 (43.5)
[au]	163.4 (17.3)	162.2 (35.6)	121.7 (21.5)	447.3 (29.1)
[ua]	39.0 (13.8)	132.0 (24.3)	378.0 (23.2)	549.1 (41.7)
[eu]	168.1 (17.4)	112.4 (11.3)	134.5 (15.9)	415.1 (34.3)
[ie]	70.8 (3.6)	82.8 (5.6)	478.0 (15.3)	631.6 (13.5)
[uo]	68.3 (22.9)	77.2 (28.2)	377.2 (27.1)	522.7 (57.1)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 6					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	66.6 (23.9)	118.3 (23.4)	76.6 (18.2)	149.4 (18.6)	151.7 (24.1)	562.7 (57.4)
[uai]	58.6 (12.3)	146.9 (31.3)	90.7 (8.2)	161.9 (19.4)	177.1 (29.7)	635.5 (67.8)
[iui]	81.8 (13.0)	93.6 (7.0)	64.4 (18.7)	135.7 (17.8)	204.0 (36.7)	579.7 (14.5)
[iai]	0.0 (0.0)	68.3 (19.5)	88.1 (18.1)	140.2 (18.2)	121.8 (20.6)	418.5 (27.8)

Vowel	Female speaker 7							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	252.9	12.6	2917.1	63.2	4143.2	88.2	477.2	41.0
[ɛ]	575.2	3.6	2738.6	40.6	3583.3	12.9	462.0	34.5
[a]	1189.7	24.5	1700.7	22.7	3503.4	94.2	483.3	31.4
[o]	695.0	40.0	967.7	22.5	3153.1	69.0	493.5	8.0
[u]	437.8	15.6	877.7	27.7	2913.3	34.0	435.9	16.5
[ɿ]	442.8	14.3	1286.9	28.5	3264.1	97.1	379.4	13.3

Diphthong	Female speaker 7					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1203.9	14.5	1699.8	53.2	3331.1	124.3
[ai] (2nd)	446.5	5.9	2856.3	80.3	3625.9	93.7
[ia] (1st)	287.3	26.9	2910.6	61.9	4203.2	101.5
[ia] (2nd)	1268.5	32.3	1824.7	39.1	3413.3	36.3
[ui] (1st)	580.4	30.8	1187.8	24.9	3090.5	75.2
[ui] (2nd)	352.5	13.5	2960.4	38.0	3506.3	79.9
[iu] (1st)	313.9	31.9	2964.0	79.0	3733.6	162.2
[iu] (2nd)	450.7	4.0	1055.4	74.5	3185.0	43.5
[oi] (1st)	831.8	9.9	1255.4	55.6	3262.0	40.1
[oi] (2nd)	504.7	46.1	2676.9	110.7	3510.0	51.8
[io] (1st)	262.3	7.8	2867.6	55.6	3862.7	68.6
[io] (2nd)	746.5	39.8	1158.0	36.9	3174.3	118.4
[au] (1st)	1061.3	13.2	1537.9	39.7	3610.6	130.0
[au] (2nd)	530.0	17.1	839.5	46.9	3164.8	53.0
[ua] (1st)	482.3	22.9	894.3	29.3	2790.8	122.7
[ua] (2nd)	1139.1	87.3	1720.6	53.0	3140.1	42.7
[eu] (1st)	729.9	30.2	2496.5	70.5	3371.2	185.4
[eu] (2nd)	553.5	48.9	1041.0	45.3	3272.5	67.6
[ie] (1st)	247.0	9.9	2988.1	61.3	4249.4	50.6
[ie] (2nd)	585.7	22.3	2761.6	30.6	3635.3	76.3
[uo] (1st)	398.0	49.4	835.9	60.4	3000.5	169.1
[uo] (2nd)	729.8	22.8	1062.5	55.7	3004.2	61.3

Triphthong	Female speaker 7					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	577.2	23.8	2395.1	53.9	3234.7	83.1
[iai] (2nd)	1099.3	13.5	1792.6	76.8	3153.0	120.0
[iai] (3rd)	479.3	28.5	2784.0	74.9	3696.2	123.0
[iui] (1st)	235.2	2.6	2765.3	127.3	3312.6	169.5
[iui] (2nd)	615.1	4.0	1218.2	67.0	3119.2	44.9
[iui] (3rd)	336.1	9.4	2934.5	15.6	3676.1	81.1
[iau] (1st)	302.4	46.2	2855.7	63.7	4044.3	105.0
[iau] (2nd)	991.3	16.1	1482.3	100.5	3158.6	74.3
[iau] (3rd)	581.8	47.0	934.2	27.8	3178.6	52.7
[uai] (1st)	482.2	37.1	888.4	48.3	2839.7	55.7
[uai] (2nd)	1101.0	88.0	1661.2	75.5	3086.0	80.9
[uai] (3rd)	478.3	59.3	2875.1	33.1	3734.6	68.5

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 7			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	201.8 (22.0)	222.8 (36.9)	176.2 (14.8)	601.0 (33.9)
[ia]	121.3 (15.9)	145.2 (13.1)	216.4 (18.5)	483.0 (28.0)
[ui]	126.0 (12.9)	129.8 (15.8)	191.0 (33.6)	446.9 (48.5)
[iu]	274.2 (39.5)	193.3 (28.3)	168.8 (21.9)	636.5 (41.7)
[oi]	209.6 (23.7)	149.9 (28.1)	119.1 (36.7)	478.7 (42.6)
[io]	112.2 (17.6)	137.5 (22.1)	267.0 (31.1)	516.8 (32.4)
[au]	151.9 (35.7)	181.0 (18.6)	136.2 (24.4)	469.2 (61.0)
[ua]	0.0 (0.0)	175.8 (17.4)	291.2 (26.1)	467.1 (24.5)
[eu]	184.7 (7.9)	140.0 (9.9)	134.6 (5.8)	459.4 (7.1)
[ie]	141.0 (31.5)	114.1 (24.4)	364.7 (18.3)	619.9 (47.1)
[uo]	51.7 (19.7)	89.7 (34.2)	392.3 (39.4)	533.8 (24.1)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 7					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	115.6 (20.9)	140.0 (16.0)	95.5 (20.1)	139.1 (20.9)	130.4 (17.6)	620.9 (7.0)
[uai]	30.8 (6.0)	137.0 (3.7)	79.4 (6.0)	173.5 (10.7)	106.6 (7.7)	527.5 (11.4)
[iui]	117.2 (11.1)	101.3 (17.6)	87.8 (16.8)	154.4 (13.0)	126.9 (21.3)	587.9 (36.5)
[iai]	0.0 (0.0)	99.5 (20.7)	89.2 (10.3)	160.6 (3.0)	108.6 (22.0)	458.2 (13.5)

Vowel	Female speaker 8							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	263.2	6.1	3048.3	43.2	3800.9	77.8	516.3	9.7
[ɛ]	593.8	5.9	2626.8	32.1	3236.3	39.6	436.4	5.7
[a]	1082.7	21.8	1610.4	27.8	2645.7	67.4	436.4	31.0
[o]	614.7	31.5	995.6	28.1	3283.8	119.7	463.4	15.2
[u]	388.1	16.9	994.6	40.7	2922.0	51.6	394.7	41.7
[ʅ]	424.0	30.8	1596.9	52.3	3111.7	47.0	397.3	28.3

Diphthong	Female speaker 8					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1209.2	70.8	1704.6	33.7	3289.2	128.0
[ai] (2nd)	543.8	10.5	3090.3	57.2	3493.6	52.4
[ia] (1st)	273.4	14.0	2867.2	30.2	3731.4	115.6
[ia] (2nd)	1111.3	31.0	1682.5	25.7	2646.1	40.6
[ui] (1st)	516.8	29.0	1178.3	54.3	3038.8	50.3
[ui] (2nd)	336.8	8.6	3021.1	83.2	3366.3	91.1
[iu] (1st)	265.2	5.2	3002.1	67.1	3692.3	89.7
[iu] (2nd)	372.9	35.0	991.9	20.3	3040.5	131.5
[oi] (1st)	784.7	50.3	1180.3	22.0	3330.8	47.3
[oi] (2nd)	353.9	11.4	2985.0	58.5	3333.0	103.6
[io] (1st)	218.0	7.4	3116.3	47.5	3784.4	65.3
[io] (2nd)	654.5	8.7	1012.2	16.3	3559.0	55.0
[au] (1st)	1134.4	71.8	1492.2	38.1	2994.4	120.4
[au] (2nd)	519.8	40.5	867.5	42.7	2783.9	139.3
[ua] (1st)	562.3	37.2	1016.9	57.9	2757.7	57.5
[ua] (2nd)	1105.7	51.5	1536.5	43.4	2616.3	127.8
[eu] (1st)	717.2	22.7	2456.1	34.0	3037.5	108.8
[eu] (2nd)	387.6	46.2	995.3	62.0	2969.1	96.0
[ie] (1st)	232.9	30.0	3237.7	27.4	3992.5	61.8
[ie] (2nd)	547.0	21.2	2740.8	127.2	3338.5	52.0
[uo] (1st)	466.9	29.1	815.7	56.1	2941.1	81.9
[uo] (2nd)	623.7	13.3	1011.5	15.1	3127.0	109.6

Triphthong	Female speaker 8					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	628.3	18.0	2235.5	54.5	3238.5	63.2
[iai] (2nd)	1109.6	13.8	1731.0	44.6	2537.4	21.2
[iai] (3rd)	372.9	16.1	3125.6	79.4	3511.8	94.5
[iui] (1st)	--	--	--	--	--	--
[iui] (2nd)	--	--	--	--	--	--
[iui] (3rd)	--	--	--	--	--	--
[iau] (1st)	263.6	10.0	2973.1	24.1	3890.0	70.2
[iau] (2nd)	1108.6	20.9	1563.3	38.7	2669.8	99.7
[iau] (3rd)	466.8	58.1	950.0	66.3	3072.3	69.0
[uai] (1st)	567.0	33.4	1035.7	39.1	2719.3	34.5
[uai] (2nd)	1101.9	45.5	1511.1	14.6	2674.2	102.6
[uai] (3rd)	550.4	13.5	2909.6	67.4	3315.6	78.3

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 8			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	136.8 (21.4)	191.4 (14.4)	187.8 (13.2)	516.2 (24.9)
[ia]	108.2 (28.1)	117.5 (8.2)	295.3 (21.8)	521.1 (49.0)
[ui]	85.1 (13.2)	153.3 (17.3)	131.9 (28.1)	370.4 (29.3)
[iu]	240.6 (11.9)	177.9 (17.5)	210.9 (17.4)	629.5 (13.3)
[oi]	158.2 (8.3)	157.7 (2.8)	147.5 (28.8)	463.4 (29.4)
[io]	128.0 (10.6)	136.2 (6.4)	306.5 (20.4)	570.8 (33.2)
[au]	111.7 (24.7)	141.3 (13.9)	157.8 (12.0)	410.9 (15.3)
[ua]	33.3 (14.8)	131.6 (12.7)	266.0 (15.9)	430.9 (22.6)
[eu]	144.5 (17.3)	121.7 (10.0)	136.5 (24.4)	402.8 (28.8)
[ie]	157.2 (6.9)	132.4 (37.9)	279.9 (27.7)	569.6 (35.6)
[uo]	40.9 (8.3)	71.0 (13.4)	350.2 (18.0)	462.2 (22.3)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 8					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	117.4 (19.7)	126.6 (7.3)	89.0 (17.4)	143.4 (26.1)	142.2 (24.2)	618.8 (62.4)
[uai]	39.7 (11.5)	97.2 (11.7)	65.8 (4.3)	136.0 (10.3)	154.4 (21.0)	493.3 (31.6)
[iui]	--	--	--	--	--	--
[iai]	0.0 (0.0)	71.2 (11.1)	66.6 (7.3)	139.2 (14.1)	104.7 (11.1)	381.8 (17.8)

Vowel	Female speaker 9							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	396.8	14.1	2909.7	37.2	3840.1	51.4	585.0	15.4
[ɛ]	614.3	27.5	2434.9	26.9	3050.6	97.6	535.8	13.8
[a]	909.4	26.2	1736.6	45.8	2494.4	120.5	515.0	19.8
[o]	685.6	7.6	1186.7	33.7	3132.9	90.3	524.3	25.8
[u]	519.9	34.7	968.5	48.5	3091.0	40.8	510.3	24.6
[ʅ]	534.1	28.9	1690.9	37.8	3250.6	40.2	518.6	32.6

Diphthong	Female speaker 9					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	911.8	16.7	1710.2	41.2	2456.9	100.1
[ai] (2nd)	605.4	28.6	2491.8	93.7	3218.6	70.3
[ia] (1st)	327.1	4.2	2524.3	67.7	3736.5	87.5
[ia] (2nd)	981.4	10.2	1709.6	53.6	2391.3	57.6
[ui] (1st)	622.4	18.4	1213.3	62.5	3106.7	120.6
[ui] (2nd)	341.2	16.1	2655.2	100.7	3349.5	66.0
[iu] (1st)	384.2	26.2	2702.7	38.7	3518.7	76.6
[iu] (2nd)	542.1	54.4	1113.5	93.1	2858.1	258.7
[oi] (1st)	834.2	21.0	1392.2	52.9	2900.1	137.4
[oi] (2nd)	549.4	28.8	2150.3	50.0	3127.0	57.1
[io] (1st)	316.6	5.2	2646.4	66.9	3603.4	75.7
[io] (2nd)	750.6	94.4	1279.7	96.6	3073.4	95.4
[au] (1st)	908.6	21.9	1517.2	31.3	2589.0	140.5
[au] (2nd)	622.7	17.2	1145.8	49.5	2958.1	162.8
[ua] (1st)	490.0	17.2	997.3	27.9	2210.5	119.7
[ua] (2nd)	949.0	23.5	1649.8	30.4	2389.4	57.9
[eu] (1st)	729.6	40.6	2051.2	77.0	2893.3	128.6
[eu] (2nd)	537.6	10.7	1359.6	73.8	3008.7	90.9
[ie] (1st)	332.8	8.7	2753.4	30.1	3842.3	113.5
[ie] (2nd)	576.3	41.0	2500.7	47.8	3176.8	120.2
[uo] (1st)	507.0	25.3	998.1	86.3	2127.5	49.6
[uo] (2nd)	713.3	38.6	1197.4	29.5	3005.3	78.7

Triphthong	Female speaker 9					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1^{st})	617.5	26.9	2083.0	60.4	3200.8	79.4
[iai] (2^{nd})	1037.6	30.8	1704.1	33.2	2381.5	20.4
[iai] (3^{rd})	636.4	26.2	2361.6	68.7	3117.9	153.9
[iui] (1^{st})	319.8	11.0	2663.9	32.5	3368.6	155.0
[iui] (2^{nd})	610.0	23.7	1321.6	63.6	3056.4	44.6
[iui] (3^{rd})	348.2	14.4	2595.8	60.5	3392.8	85.9
[iau] (1^{st})	338.7	6.0	2573.1	15.7	3788.3	107.4
[iau] (2^{nd})	968.2	30.6	1578.3	26.6	2557.3	129.9
[iau] (3^{rd})	602.8	43.2	1065.3	34.2	3395.3	129.4
[uai] (1^{st})	541.3	51.0	1040.0	99.1	3062.0	61.0
[uai] (2^{nd})	966.7	45.2	1639.5	33.3	2323.1	84.2
[uai] (3^{rd})	602.1	31.1	2435.1	134.3	3192.7	46.6

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 9			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	264.6 (27.0)	154.5 (22.8)	117.4 (30.6)	536.6 (24.9)
[ia]	47.2 (8.8)	92.6 (11.1)	345.3 (11.5)	485.2 (11.1)
[ui]	169.3 (22.6)	171.7 (23.5)	92.5 (10.8)	433.7 (12.1)
[iu]	338.9 (55.9)	180.1 (33.3)	113.3 (24.7)	632.4 (25.8)
[oi]	255.4 (16.5)	143.5 (11.5)	88.5 (16.5)	487.4 (21.9)
[io]	71.5 (16.2)	106.3 (17.6)	362.7 (31.9)	540.5 (45.1)
[au]	202.6 (27.2)	128.7 (10.5)	76.1 (10.1)	407.4 (20.9)
[ua]	0.0 (0.0)	125.8 (15.5)	384.0 (52.6)	509.8 (41.2)
[eu]	280.1 (16.2)	95.3 (7.7)	79.0 (17.4)	454.6 (4.4)
[ie]	65.6 (26.6)	63.6 (10.6)	410.7 (39.0)	540.0 (65.7)
[uo]	42.4 (3.3)	94.5 (30.7)	428.3 (39.2)	565.4 (19.8)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 9					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	67.6 (18.2)	131.9 (17.9)	218.6 (18.5)	91.1 (12.7)	88.7 (12.2)	598.1 (34.4)
[uai]	0.0 (0.0)	112.3 (17.8)	179.3 (20.9)	154.6 (19.5)	95.1 (26.3)	541.4 (15.7)
[iui]	63.4 (16.9)	75.3 (13.0)	98.3 (17.7)	182.8 (23.3)	112.5 (3.5)	532.5 (23.0)
[iai]	0.0 (0.0)	70.7 (15.4)	177.5 (23.5)	128.0 (7.8)	61.0 (7.5)	437.3 (12.6)

Vowel	Female speaker 10							
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (Hz)	s.d.	duration (ms)	s.d.
[i]	265.9	3.8	2611.1	85.0	3702.8	46.9	436.7	27.6
[ɛ]	538.2	20.8	2476.0	28.6	3236.0	42.6	346.2	28.0
[a]	967.8	3.9	1691.2	28.2	2539.1	129.9	393.5	22.9
[o]	561.8	14.1	909.2	37.5	3135.3	114.6	434.3	40.6
[u]	344.1	9.7	798.5	22.0	2887.1	187.7	357.0	20.4
[ɿ]	347.7	17.9	1459.8	32.7	2994.3	88.0	331.4	32.6

Diphthong	Female speaker 10					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[ai] (1st)	1034.1	15.5	1677.0	28.8	2544.4	74.0
[ai] (2nd)	274.3	31.9	2700.8	76.3	3363.5	105.0
[ia] (1st)	271.5	13.5	2498.2	64.0	3531.3	120.2
[ia] (2nd)	1008.9	7.5	1707.3	17.6	2573.6	41.3
[ui] (1st)	422.2	17.2	1490.9	133.1	2933.3	140.6
[ui] (2nd)	314.7	6.9	2766.5	23.9	3294.0	71.6
[iu] (1st)	292.7	19.5	2676.5	77.0	3404.4	62.3
[iu] (2nd)	362.2	15.1	949.9	18.2	2949.2	135.3
[oi] (1st)	756.2	23.8	1114.5	43.2	2792.0	132.5
[oi] (2nd)	424.4	11.0	2531.8	63.7	3232.6	74.8
[io] (1st)	288.8	12.9	2654.1	70.2	3482.1	110.9
[io] (2nd)	647.9	12.0	1021.5	49.7	2977.2	111.6
[au] (1st)	961.9	16.1	1461.6	47.2	2534.4	105.3
[au] (2nd)	574.8	49.8	1007.4	15.9	2679.3	115.2
[ua] (1st)	533.2	25.4	1028.7	45.4	2454.1	137.8
[ua] (2nd)	1012.4	30.9	1651.4	19.8	2418.6	128.1
[eu] (1st)	653.8	9.0	2231.3	45.4	2939.4	152.9
[eu] (2nd)	437.6	34.9	1002.7	53.6	2710.5	41.6
[ie] (1st)	260.8	6.2	2654.4	78.6	3578.4	150.0
[ie] (2nd)	522.1	11.9	2505.5	47.9	3299.9	125.0
[uo] (1st)	412.9	13.8	839.2	41.1	2800.6	84.5
[uo] (2nd)	608.8	28.0	939.1	39.5	2998.1	75.1

Triphthong	Female speaker 10					
	F_1 (Hz)	s.d.	F_2 (Hz)	s.d.	F_3 (s.d.)	s.d.
[iai] (1st)	562.0	47.7	2066.9	48.7	2989.7	72.6
[iai] (2nd)	939.9	15.7	1742.5	50.2	2428.8	45.3
[iai] (3rd)	677.1	69.1	2455.4	110.6	3392.3	172.9
[iui] (1st)	299.1	8.8	2617.5	42.4	3523.1	93.1
[iui] (2nd)	454.9	43.7	1434.5	13.6	2919.2	104.0
[iui] (3rd)	321.1	11.8	2714.4	79.5	3356.9	149.7
[iau] (1st)	266.6	6.5	2631.9	76.9	3396.2	113.2
[iau] (2nd)	991.5	25.5	1556.9	27.8	2547.9	50.1
[iau] (3rd)	387.1	20.1	986.5	82.7	2738.7	37.2
[uai] (1st)	539.1	10.4	1010.1	17.8	2471.5	117.0
[uai] (2nd)	950.1	23.2	1626.1	42.2	2378.9	71.7
[uai] (3rd)	278.0	20.2	2753.1	93.5	3566.5	72.1

Diphthong	Temporal organization (in ms) of Hakka diphthongs for Female speaker 10			
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Total duration (s.d.)
[ai]	101.1 (24.8)	180.6 (20.1)	163.1 (43.1)	444.9 (60.6)
[ia]	85.4 (22.9)	95.3 (7.3)	208.1 (22.1)	388.9 (23.9)
[ui]	54.5 (13.5)	119.6 (15.6)	115.4 (49.3)	275.9 (47.1)
[iu]	164.6 (32.6)	162.7 (33.2)	226.7 (25.8)	554.1 (39.9)
[oi]	126.3 (8.7)	123.4 (20.1)	104.2 (9.4)	354.0 (19.6)
[io]	91.8 (26.3)	96.0 (8.7)	238.2 (19.9)	426.0 (43.3)
[au]	113.1 (16.8)	82.3 (8.5)	137.8 (17.8)	333.2 (22.6)
[ua]	0.0 (0.0)	119.6 (20.0)	257.7 (14.9)	377.4 (33.2)
[eu]	131.3 (28.3)	102.8 (15.0)	117.0 (16.4)	351.3 (28.3)
[ie]	95.5 (36.0)	68.5 (4.5)	307.2 (23.8)	471.4 (35.0)
[uo]	36.6 (7.4)	63.1 (11.2)	304.1 (9.5)	403.9 (12.4)

Triphthong	Temporal organization (in ms) of Hakka triphthongs for Female speaker 10					
	First element (s.d.)	Transition (s.d.)	Second element (s.d.)	Transition (s.d.)	Third element (s.d.)	Total duration (s.d.)
[iau]	92.4 (28.1)	104.9 (18.5)	102.8 (18.6)	115.3 (28.8)	177.8 (20.4)	593.4 (41.6)
[uai]	29.2 (3.0)	73.0 (12.4)	78.4 (10.2)	173.1 (27.4)	115.1 (14.5)	469.0 (19.7)
[iui]	65.9 (19.9)	60.7 (4.9)	49.2 (10.4)	85.9 (7.9)	158.7 (21.5)	420.5 (21.6)
[iai]	0.0 (0.0)	53.1 (5.5)	67.4 (8.4)	92.8 (10.1)	83.5 (9.5)	296.9 (3.8)