A STUDY OF NEUTRAL-TONE SYLLABLES IN TAIWAN MANDARIN

A DISSERTATION SUBMITTED TO THE GRADUATE DIVISION OF THE UNIVERSITY OF HAWAI‘I AT MĀNOA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

IN

LINGUISTICS

AUGUST 2012

By

Karen Huang

Dissertation Committee:

Patricia Donegan, Chairperson
Victoria Anderson
Ying-che Li
Kenneth Rehg
David Yang

For my Mother

ACKNOWLEDGMENTS

I am truly grateful to all the help and support I received from my teachers, friends, and families. First of all, I owe my deepest gratitude to my advisor Professor Patricia Donegan. This dissertation would not have been possible without her guidance, support, and patience. I truly appreciate her expertise, assistance in so many ways, and warm kindness. Words just simply cannot describe how deep my appreciation is.

I am deeply grateful to Professor Victoria Anderson, who led me into the wonderful world of phonetics. Aside from all the valuable comments and advice, her enthusiasm really kept reminding me how fun this study is. I am also truly indebted to Professors Ken Rehg, Ying-che Li, David Yang, Amy Schafer, Katie Drager, Haidan Wang, and Song Jiang, for their generous time, energy, suggestions and help.

My sincere thanks also go to my friends in the department of Linguistics and EALL. I would like to thank Akiemi Glenn, Apay Tang, Nian Liu, Carl Polley, Jawee Perla, Elaine Lau, Kaori Ueki, Kanjana Thepboriruk, Diana Stojankovic, Jake Terrell, Chih-kai Lin, Cheng-chun Kuo, Tsai-Hsiu Liu, Hiroko Sato, Laura Robinson, and many others for various suggestions, discussions, and encouragement. I deeply appreciate all the 加油, がんばって, and "You can do it!" from all my friends. Furthermore, I wish to thank our guardian angels Jennifer Kanda and Nora Lum, they are the best!

I would like to show my gratitude to all the people who helped me with my experiments: Rachel Chuang, Mei Xu, Tanny Tang, Shu-ling Wu, Daniel Tom, Nancy Arakawa, many students and staff from National Cheng Kung University, and many others for their generous help and kindness.

I am truly thankful to my family in Taiwan, California, and Ohio. They have always believed in me and been there for me. I cannot thank them enough for their assistance and encouragement. Most of all, I would like to thank my Dad. Even though he always wanted me to finish earlier, he still patiently encouraged me and shared his wisdom with me. I deeply appreciate his trust and love.

Last, but not least, I would like to thank my husband and best friend, John Kupchik, for his love, support, and encouragement. He made me believe in myself and brought smiles to my face.

ABSTRACT

This dissertation studies the realization of the rhythm of Taiwan Mandarin and focuses on the quality of its unstressed (neutral-tone) syllables. Taiwan Mandarin (TM) is often described as more syllable-timed than Standard Mandarin (SM). In TM, the unstressed syllables occur less frequently. The quality of the unstressed (neutral-tone) syllables in TM also seems to differ from those of SM.

The results show that most of the neutral-tone syllables have a mid-low pitch target except in some frequent vocabulary items, or when the original lexical tone is obvious. Unlike SM, the neutral-tone syllables of TM do not undergo tone loss. The neutral tone in TM has a mid-low pitch target in disyllabic words, neutral-tone sequences and novel formations. The pitch target of the neutral-tone syllables is generally not as low as the low tone, and the carry-over effect of the preceding tone is stronger in the neutral tone than in the low tone. Also, there is no general syllable reduction in the neutral-tone syllables in TM. The durations are shorter than their full-tone counterparts only when their rimes are [ə] or [ɨ], and the duration differences are much smaller than in SM. The results suggest that these syllables are better characterized as having a lexical tone, which can be considered the fifth tone in the language. SM, in contrast, has four lexical tones.

The perception tests further show that TM listeners rely on duration, pitch, and vowel/voice quality to distinguish the low tone from the innovative fifth tone. Specifically, for the pitch cues, they rely on low end pitch and fast-falling pitch contour to identify the low tone. The results also show that the perceptual distinctions made by the TM speakers between the low tone and the fifth tone were very weak. The TM listeners could only distinguish 20% of the unaltered pairs. Along with the evidence found in the production tests, this study finds that the innovative fifth tone is on the verge of merging with the low tone.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

1	First person
2	Second person
3	Third person
ASP	Aspect
CL	Classifier
DIM	Diminutive
DUR	Durative
F	Full tone
H	High tone; Mandarin Tone 1
HL	Falling tone; Mandarin Tone 4
L	Low tone; Mandarin Tone 3
LH	Rising tone; Mandarin Tone 2
N	Neutral tone
NML	Nominalizer
PERF	Perfective
PL	Plural
POSS	Possessive
Q	Question particle
REDUP	Reduplicant
S	Singular
Ø	Neutral tone

CHAPTER 1. INTRODUCTION

Characterizing speech rhythm has been a goal of both phonologists and phoneticians for many decades. Most of the earlier works focused on evaluating the isochrony hypothesis (Pike 1945; Abercrombie 1967): stress-timed languages are based on the isochrony of the interstress intervals; syllable-timed languages are based on the isochrony of the syllables. However, efforts to find isochrony were not successful (Lehiste 1977; Roach 1982; Betran 1999). In the last thirty years, researchers have focused on the phonetic and phonological properties that characterize speech rhythm instead of searching for isochrony (Dasher & Bolinger 1982; Dauer 1983; Ramus, Nespor & Mehler 1999; Grabe & Low 2002). As a result, the continuum between stress-timed and syllable-timed languages is evaluated by syllable complexity and the tendency for lengthening accented syllables and reducing unstressed syllables. Much of the subsequent research has also applied the concepts and compared the speech rhythms across different dialects, for example British and Singapore English (Low, Grabe & Nolan 2000; Deterding 2001), Eastern and Western varieties of Arabic (Ghazali, Hamdi & Barkat 2002), European and Brazilian Portuguese (Frota & Vigário 2001), different varieties of Italian (Russo & Barry 2004), two Peruvian Spanish dialects (O'Rourke 2008), and so on. However, these dialectal rhythmic differences have all been related to segmental differences such as vowel deletion, vowel reduction, vowel epenthesis, and diphthongization.

On the other hand, the rhythmic differences reflected in tone languages have never been carefully examined. The rhythmic manifestation on tones has been studied much less than the segmental features. This is likely because pitch is a different dimension from tempo, and tone languages typically have simple syllable structures. Also, tone languages that assign tones to syllables, such as East-Asian tone languages, are often believed to be syllable-timed, while those that assign tones lexically to words tend to be stress-timed (Auer 1993). Nevertheless, tone languages that assign tones to syllables can have stress, and unstressed syllables can undergo vowel reduction and tone loss. Mandarin, a tone language, is an example of this.

Despite the general belief that Mandarin is a syllable-timed language (Grabe & Low 2002; H. Lin & Wang 2007) due to its simple syllable canon (C)(G)V(C), Mandarin in casual speech style actually maps closely with stress-timed languages such as English (Benton, Dockendorf, Jin, Liu & Edmondson 2007). The stress-timed features in casual speech style are related to syllable de-stressing in colloquial Mandarin and the behavior of unstressed syllables. In Standard Mandarin, about one-third of all syllables are unstressed and toneless in connected speech (Duanmu 2000, 257-258). When an unstressed syllable reduces, the coda nasal is lost and the vowel reduces towards a schwa (Duanmu 2000, 256). For example, the second syllable of /ʈʂ^{h}wən-t^{h}jan/ ‘spring’ is unstressed, therefore the word is pronounced [ʈʂ^{h}wən-t^{h}jə̃]. An unstressed syllable is also shorter in duration (Chao 1968). The mean duration of syllables with neutral tones is about 50%-60% of the mean duration of those with lexical tones (M. Lin & Yan 1980; Lee & Zee 2008).

Furthermore, unstressed syllables in Mandarin undergo tone loss. The pitch of the unstressed syllable is decided by the lexical tone of the preceding syllable. For example, /fei-lə/ [H-Ø] ‘fly-PERF’ is [55-2], /lai-lə/ [LH-Ø] ‘come-PERF’ is [35-3], /mai-lə/ [L-Ø] ‘buy-PERF’ is [21-4], and /mai-lə/ [HL-Ø] ‘sell-PERF’ is [51-1]. The actual pitch of the /lə/ is determined by the lexical tone of the preceding syllable. These unstressed syllables are often referred to as having a “neutral tone” (Chao 1968, 36). A neutral-tone syllable is toneless because it has lost its underlying tone and its pitch is determined by the preceding syllable (Duanmu 2007, 242-243).

On the other hand, the rhythm of Mandarin varies across dialects (Benton et al. 2007). Taiwan Mandarin (TM), a Mandarin dialect spoken in Taiwan, is considered more “syllable-timed” than Standard Mandarin (Kubler 1985). Since the syllable canons are similar across Mandarin dialects, the differences in rhythm most likely lie in the treatment of unstressed syllables. Unstressed (neutral-tone) syllables and syllable de-stressing occur less frequently in TM (Kubler 1985, 161; Tsao 2000; Swihart 2003, 110; Duanmu 2007, 308). This difference is even noticed by the government. According to the dictionary compiled by Ministry of Education in Taiwan, there are fewer unstressed syllables in TM. Furthermore, many of the prescriptively unstressed[1] syllables are obviously produced with lexical tones. For example, /caŋ-jiŋ/ ‘a housefly’ is marked as /H-Ø/ in dictionaries in Taiwan, but it is pronounced as /H-LH/ by TM speakers. The rime of the “unstressed” syllable is not reduced at all either. The syllable is clearly stressed.

The lower frequency of the unstressed syllables in Taiwan Mandarin, prescriptively and descriptively, contributes to the fact that TM is more syllable-timed than Standard Mandarin. Moreover, it seems that the rhythmic difference even affects the remaining unstressed syllables. These unstressed syllables prescriptively have a neutral tone, and they either do not behave like the lexical tone before destressing, or they cannot be traced back to a lexical tone at all. The pitch of these neutral-tone syllables does not seem to be determined by the preceding syllable as in Standard Mandarin either. The pitch of these neutral tone syllables seems to have a target (J. Li 2005). For example, the perfective suffix /lə/ in TM sounds low regardless of the tones of the preceding syllables. If /lə/ had a neutral tone as in TM, I would expect it to have different pitches depending on the preceding tone.

The fact that the neutral-tone syllables in Taiwan Mandarin have a certain pitch target suggests that these syllables are not really toneless. However, if they are not toneless, what is the tonal identity of these neutral-tone syllables? Are they assigned a new tone or do they keep one of the four lexical tones in Mandarin? Are these neutral-tone syllables really unstressed in Taiwan Mandarin? How has the rhythmic difference in Taiwan Mandarin affected these supposedly unstressed syllables?

This dialect variation within Mandarin provides an opportunity to study tonal behavior in different rhythms. As Taiwan Mandarin is perceived as more syllable-timed than Standard Mandarin, the treatment of the so-called “unstressed” syllables in Taiwan Mandarin is a perfect chance to help us understand how rhythmic differences are manifested in a tone language. Specifically, what happens to the features of a stress-timed tone language when this language is no longer stress-timed? There has been no complete study that examined the features of the neutral tone in Taiwan Mandarin. It would be

[1] A syllable is prescriptively unstressed when it is spelled with a neutral tone in dictionaries. A Neutral-tone syllable is marked by the dot in *zhuyin fuhou*, which is the phonetic spelling system taught and used in Taiwan.

important to characterize the neutral-tone syllables in Taiwan Mandarin in order to answer the questions above.

This dissertation studies the realization of the rhythm of TM and focuses on the quality of its “unstressed” (neutral-tone) syllables. It examines the acoustic properties of the TM neutral-tone syllables, specifically the pitch behavior in different contexts; it compares the stressed (full-tone) with the unstressed (neutral-tone) syllables in TM in terms of duration and pitch contours; and it includes perception experiments to see whether TM listeners can distinguish TM stressed and unstressed syllables and it examines the perceptual cues they may use.

Chapter Two summarizes the phonology of Standard Mandarin, and provides a background of the linguistic area. It also outlines the features of Taiwan Mandarin and gives a more detailed description of the neutral-tone syllables in Taiwan Mandarin.

Chapter Three explores the pitch behaviors of the TM neutral tone in different contexts in order to characterize the tonal identity of these neutral-tone syllables. I examine the pitch of the TM neutral tone in disyllabic words. I also examine the pitch patterns of the neutral-tone syllables when they are attached to novel words in order to exclude possible lexicalized pitch contours. Lastly, I investigate the consecutive “neutral-tone” syllables to see if they each have their own phonemic pitch specification instead of gradually moving toward a mid-target as has been observed in Standard Mandarin.

One important feature of syllable-timed languages is that the differences between stressed and unstressed syllables are not as obvious as those of stress-timed languages. Chapter Four compares the stressed (lexical-tone) syllables and the unstressed (neutral-tone) syllables in TM. Each neutral-tone syllable is compared to its lexical-tone counterparts that have similar pitch contours. Both pitch contours and durations are examined to show the similarities and the differences between the lexical-tone and the neutral-tone syllables.

Chapter Five contains two perception tests. The first perception test examines whether the difference between the lexical-tone and the neutral-tone spoken by four TM speakers is perceptible by the TM listeners. The test also tries to search for possible perceptual cues that were used by the TM listeners. The second perception test uses speech with manipulated pitch contours to find out the pitch cues that were used by the TM listeners to distinguish the neutral tone from the lexical tone.

Chapter Six summarizes the results and discusses the limitations of this dissertation. It also presents the conclusion which evaluates the current status of the neutral-tone syllables in Taiwan Mandarin.

The purpose of this study is to examine the neutral-tone syllables in Taiwan Mandarin and further understand how rhythmic difference affects the unstressed syllables in a tone language. This study is the first to give a proper portrait of the neutral-tone syllables in TM. It investigates the acoustic phonetics of the neutral-tone syllables in detail, along with perception studies on how these syllables are perceived by the TM listeners.

This study is not limited to exploring the rhythmic influence on the “unstressed” syllables in TM. These features of “unstressed” syllables may also imply that Taiwan Mandarin has a different phonology compared to Standard Mandarin. For example, if the neutral tone were a fifth tone, a five tone system would have to be proposed for Taiwan Mandarin. If the neutral tone were in fact a low tone (Tone 3), certain phonological rules such as Tone 3 sandhi (underlying /L-L/ changing to [LH-L]) might need to be modified for Taiwan Mandarin, since [L-L] would actually be possible in Taiwan Mandarin.

CHAPTER 2. Mandarin and Taiwan Mandarin

This chapter includes three parts. Section 2.1 describes the main features of Standard Mandarin phonology. Section 2.2 first gives a short introduction on the linguistic situation in Taiwan, and then describes the features of Taiwan Mandarin as a Mandarin dialect.

2.1. Standard Mandarin phonology

Mandarin Chinese, one of the Chinese languages, is the lingua franca of China. Mandarin is spoken natively in China roughly in the north of the Yangtze River and in most of Southwest China. It is also spoken natively in Taiwan, Singapore, and by immigrants in various countries. Standard Mandarin, also known as *Putonghua*, is based on the Beijing Mandarin dialect. Although Beijing dialect is different from SM, the differences are mostly lexical. Most linguists use Beijing Mandarin speakers to represent Standard Mandarin.

2.1.1. Segments and tones

There are 25 consonants in Standard Mandarin. Table 2.1 organizes the consonants in the way that Chinese linguists often do: labials, alveolars, alveolar sibilants, retroflexes, alveolo-palatals, and velars are presented in rows, while manners of articulation are presented in columns. The consonants /ŋ, w, j, ɥ/ cannot be in the initial onset position, and the voiceless velar fricative /x/ can be a voiceless glottal or uvular fricative depending on the dialect.

Table 2.1 Mandarin consonant inventory

	unaspirated	aspirated	nasals	fricatives	voiced continuants
labials	p	p^h	m	f	w*
alveolars	t	t^h	n		l
alveolar sibilants	ts	ts^h		s	
retroflexes	ʈʂ	$ʈʂ^h$		ʂ	ɻ
alveolo-palatals	tɕ	$tɕ^h$		ɕ	j* ɥ*
velars	k	k^h	ŋ*	x (χ, h)	

There are five vowel phonemes in Mandarin: /i, y, u, ə, a/ (Duanmu 2007, 35) as shown in Table 2.2. The high vowels contrast in triplets such as /li/ 'strength', 'lu' 'road', and /ly/ 'green'. The high vowel /i/ also has an allophone [ɨ] when it is after the alveolar sibilants /ts, ts^h, s/ or the retroflexes /ʈʂ, ʈʂh, ʂ, ɻ/.[2]

Table 2.2 Mandarin vowel phonemes

	front	central	back
high	i y		u
mid		ə	
low		a	

The mid vowel and the low vowel have several variants. In an open syllable, the mid vowel /ə/ has variants [e] after [j] or [ɥ], [o] after [w], and [ɤ] elsewhere. When the mid vowel is before the coda /u̯/ or /i̯/, /əi̯/ becomes [ei̯] and /əu̯/ becomes [ou̯]; when the mid vowel is unstressed, it is [ə]. As for the low vowel /a/, it becomes [ɑ] in an open syllable, or when it is before a velar nasal /ŋ/ or the coda /u̯/. The Mandarin vowel allophones are shown in Table 2.3.

[2] [ɨ] is sometimes described as apical vowels /ɿ, ʅ/. Some linguists analyze [ɨ] as syllabic consonants [z, ʐ]. [ɿ] or [z] are found after /ts, ts^h, s/, and [ʅ] or [ʐ] are found after /ʈʂ, ʈʂh, ʂ, ɻ/.

Table 2.3 Mandarin vowel allophones

	front		Central	Back	
	unround	round	unround	unround	round
High	i	y	ɨ		u
Mid	e		ə	ɤ	o
Low			a	ɑ	

The syllable canon of Mandarin is (C)(G)V(X). The medial glide can be one of the three glides /w, j, ɥ/, and the coda can only be one of the consonants /ɻ, n, ŋ/ or one of the non-syllabic vowels /u̯, i̯/. There are four VV̯ rimes in Mandarin: /ai̯, əi̯, au̯, əu̯/. In VN rimes, only /i, ə, a/ can combine with final nasals. There are six VN rimes in Mandarin: /in, iŋ, an, aŋ, ən, əŋ/.

Table 2.4 Mandarin tones

Tone	Pitch description	Tone mark in *pinyin*	Chao's scale	Notation in this study
Tone 1	High-level	ā	55	H
Tone 2	High-rising	á	35	LH
Tone 3	Dipping (low)	ǎ	214 (21)	L
Tone 4	High-falling	à	51	HL
Neutral tone	-	a	-	Ø

As shown in Table 2.4, there are four lexical tones in Mandarin: Tone 1, 2, 3, and 4. They are high-level, rising, dipping, and high-falling respectively. There is also a neutral tone, and its pitch contour is decided by the preceding lexical tone. In Standard Mandarin, a syllable with any of the four lexical tones can have a neutral tone when it is unstressed. These unstressed syllables can be found in the second syllables of some disyllabic compounds. However, some syllables, mostly suffixes and particles, always have a neutral tone. The most commonly adopted description of tones is Chao's five point scale (1956; 1968), in which a pitch range is divided into five levels, [5] being the highest and [1] being the lowest. The four tones of Standard Mandarin are marked as [55, 35, 214, 51] respectively. Tone 3 has two variants. It is a dipping tone [214] when it is in isolation or final position, and it is a low tone [21] when it is in non-final position.

In this study, instead of using Chao's five point scale or listing the tones numerically from 1 to 4, I adopt Duanmu's model (2007, 225-235) and represent Tone 1 (high level) as /H/, Tone 2 (rising) as /LH/, Tone 3 (low) as /L/, Tone 4 (falling) as /HL/, and the neutral tone as /Ø/, because this notation is easier to follow for non-Mandarin-speaking readers. Furthermore, as I will discuss in 2.2.2.3, Chao's description of Standard Mandarin tones does not apply to Taiwan Mandarin. Therefore it is better to describe the tones in terms of categories instead of phonetic/perceptual values.

2.1.2. Rhythm

According to Duanmu (2007, 137-139), SM has trochaic disyllabic feet, and stress is sensitive to syllable weight. As a result, licit foot structures include heavy, heavy-heavy, or heavy-light. A light syllable does not occur word initially. A light syllable is unstressed and toneless. Unlike most languages with stress where most of the stresses can be predicted by the stress assignment rule, unstressed syllables in Standard Mandarin are not predictable. The unstressed syllables are lexically determined in Standard Mandarin. They would be marked by zero tone mark in *hanyu pinyin*, the phonetic spelling system used in China.

In SM, about 15-20% of the syllables in written texts are considered unstressed (W. Li 1981, 35). They are found in suffixes, clitics, final particles, postpositions, directional complements, the second syllable of reduplications, and the second syllable of some disyllabic words. X. Chen (2004) analyzed the words that contain the neutral tone in the 1996 edition of *Xiandai Hanyu Cidian* 'Modern Chinese Dictionary'. She found that of the 2882 entries that contain the neutral tone, 1551 of them (53.8%) are syntactically based bound morphemes. The other 1371 (47.6%) are lexically-based, which are the second syllables of disyllabic words with unstressed second syllables. These words include common nouns like *luo-bo* /LH-Ø/ 'radish' and *bo-li* /H-Ø/ 'glass'; body parts such as *mei-mao* /LH-Ø/ 'eyebrow', *yan-jing* /L-Ø/ 'eye'; and verbs like *xi-huan* /L-Ø/ 'to like' and *ji-du* /HL-Ø/ 'to envy'.

Whether a syllable is unstressed is not always consistent. X. Chen (2004) also found that in the 1996 edition of Modern Chinese Dictionary, 375 words are listed with an optional neutral tone. For example, *ti-liang* 體諒 ‘to understand’ can be read with /L-HL/ or /L-Ø/. There is also inconsistency across different dictionaries. A syllable might be marked with a neutral tone in one dictionary, but marked with a lexical tone in the others.

Aside from these prescriptively unstressed syllables, destressing also occurs in colloquial SM. Destressing refers to the process in which an underlying full syllable becomes weak and toneless. As mentioned earlier, about one-third of all syllables are unstressed and toneless in SM in connected speech (Duanmu 2000, 257-258). This is higher than the percentage of unstressed syllables in written texts (15-20%). Jing (2002) and Deng et al. (2008) also notice that unstressed syllables are unstable in SM. SM speakers tend to stress the unstressed syllables if the word is infrequent. On the other hand, more frequent disyllabic words change into a heavy-light rhythm, and the second syllable becomes unstressed.

An unstressed syllable reduces in SM: the coda nasal is lost and the vowel centralizes towards a schwa. Also, in connected speech, an unstressed vowel can be devoiced or deleted when the vowel is high and after a fricative or an aspirated stop or affricate (Duanmu 2000:257). For example, /i-tɕ^{h}i/ ‘together’ becomes [i-tɕ^{h}] and /ʈʂ^{h}u-tɕ^{h}y/ ‘get out’ becomes [ʈʂ^{h}u-tɕ^{h}] in connected speech because the second syllables are unstressed. Furthermore, in connected speech, the initial of an unstressed syllable often undergoes lenition (Chao 1968, 38; S. Xu 1980, 159; Duanmu 2000, 255-256). Duanmu (2000:256) lists several examples: /lipa/ becomes [liba] ‘fence’, /pau-ʂaŋ/ becomes [pau-ɹə̃] ‘in the newspaper’, and /kaŋ-kaŋ-tɕ^{h}y/ becomes [kã-ɣã-tɕ^{h}y] ‘just went’. Unstressed syllables are also shorter in duration, and their tones are neutralized (Chao 1968). The phonetic features of the neutral-tone syllables are discussed further in Section 2.1.3.

Typologically Mandarin has been called a syllable-timed. Lin and Wang (2007) followed Ramus et al. (1999) and Grabe and Low (2002) and measured four rhythmic parameters: vowel percentage, consonant standard deviation, rPVI, and nPVI of Mandarin; and they found that Mandarin grouped closely with syllable-timed languages.

However, Mandarin seems to have a lot of variation in rhythm depending on speech style and dialect (Benton et al. 2007). In their research, Benton et al. found that Mandarin in casual speech style actually mapped closely with stress-timed languages such as English. This is probably due to the unstressed syllables and destressing in colloquial Mandarin.

2.1.3. Neutral-tone syllables in Standard Mandarin

As mentioned above, the pitches of the neutral tone in Standard Mandarin depend on the preceding syllables. Chao (1968) described the neutral-tone syllables as "completely unstressed" (Chao 1968: 27) and "the tone range is flattened to practically zero and the duration is relatively short" (Chao 1968:35). He also characterized the pitch of the neutral tone as *half-low* after /H/ [55], *middle* after /LH/ [35], *half-high* after /L/ [21], and *low* after /HL/ [51] (Chao 1948:27). However, no quantitative data were presented to support Chao's description.

Lin and Yan (1980) did an acoustic study of the neutral tone in Beijing Mandarin. They compared neutral-tone syllables with lexical-tone syllables in minimal pairs, where disyllabic words of the shape heavy-light were compared with minimal pairs of the shape heavy-heavy. They found that light syllables are about 50% shorter than their counterparts in the pairs, and their intensity is also reduced. They also measured the pitch contours of the neutral tone and confirmed that the pitch varies depending on the preceding lexical tone: the pitch contours of the neutral tone were [41], [51], [44]/[33], and [21] after the lexical tones /H/ [55], /LH/ [35], /L/ [21], and /HL/ [51] respectively.

Lee and Zee (2008) revisited the neutral tone and found that the F0 contours of unstressed syllables are *mid falling* after /H/, *high falling* after /LH/, *mid level* after /L/, and *low falling* after /HL/, which are similar to the previous studies. They also found that the F0 contours of the unstressed syllables are also influenced by both the preceding and following citation tones. For example, the neutral tone is *high level* between /H, LH/ and /HL/; and it is *high-mid level* between /L/ and /L/. Aside from the pitch contours, they also looked at the intensity and duration of the neutral-tone syllables. The result shows that the intensity of the neutral-tone syllables is not necessarily lower than their preceding lexical-tone syllables. The intensity curves co-vary with the pitch contours. Their durations are mostly shorter than their neighboring lexical-tone syllables. Their results of the four speakers showed that in disyllabic words, the duration of the initial lexical-tone syllables is 1.87 to 2.01 times as long as the duration of the following neutral-tone syllables.

Chen and Xu (2006) found that when there are consecutive unstressed syllables, the consecutive neutral-tones syllables have a mid-level target. Figure 2.1 shows how the consecutive neutral tones are influenced by the preceding lexical tones and slowly merge to a mid-target. For a full lexical tone, the influence of the preceding lexical tone is overcome quickly and the underlying tonal target is implemented immediately. In contrast, the mid target in a neutral tone is slowly approached with weak articulatory strength, and the influence of the preceding lexical tone, although it decreases over time, is still substantial at the end of the consecutive neutral tone syllables.

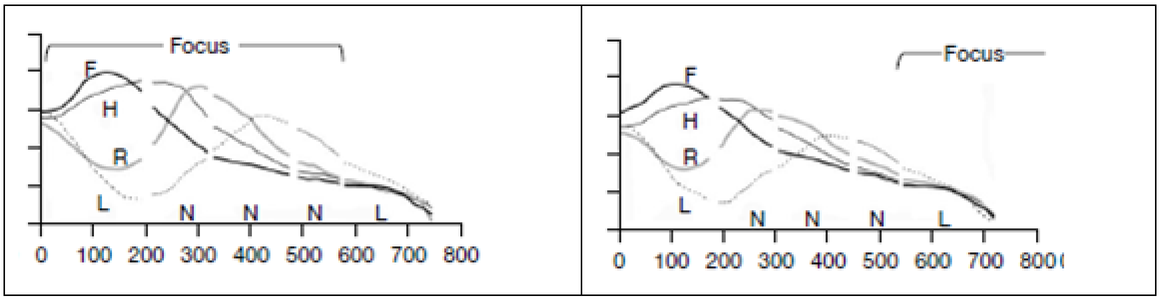


Figure 2.1 Mean F0 contours of 3 consecutive neutral-tone syllables in different tonal contexts under different focus conditions before a low tone in SM (Y. Chen & Xu 2006, 54)

Another feature of neutral-tone (unstressed) syllables in SM is vowel centralization. Lin and Yan (1990) compare the rimes of the unstressed syllables with the rimes of these syllables when they are stressed. For example, they compare the rimes of *chou* in *bao-chou* /HL-LH/ 'revenge' and *bao-chou* /HL-Ø / 'reward'. They plot the vowels using the measurements of F1 and F2. The result shows that a monophthong is centralized when it is unstressed. As for diphthongs, in unstressed syllables, the vowel is also centralized and the off-glide is usually deleted or centralized. Similarly, final nasals in unstressed syllables are usually deleted while the vowels are nasalized.

Aside from the production research on neutral tone in Standard Mandarin, several perception studies have also been done. Lin (1985) chose seven minimal pairs of heavy-heavy and heavy-light disyllabic words such as *ya-tou* /H-LH/ 'duck head' vs.

ya-tou /H-Ø/ ‘girl’. He then manipulated the pitch contour, the intensity, and the duration of the second syllables and had the Standard Mandarin listeners distinguish between the minimal pairs. The subjects could choose between the pairs or the option ‘cannot tell’. The result shows that the duration is the main cue by which subjects distinguish the pairs, and the beginning pitch of the neutral tone also takes an important role. On the other hand, the intensity is not an important cue, and the pitch contour plays only a minor role in distinguishing the pairs because the neutral-tone syllables are usually short.

Lin and Wang’s perception test (1984) also shows that the beginning pitch of the neutral-tone syllable is an important perceptual cue. The synthesized *bai-de* syllables are manipulated so the first syllables are always on the level of 115 Hz, and the second syllables are falling contours of (110-70 Hz, 120-80 Hz, 130-90 Hz, and 140-100 Hz). The listeners identified *bai-de* as /H-Ø/ 掰的 ‘to fabricate-NML’ ‘the made-up one’ when the second syllables are 110-70 Hz and 120-80 Hz, and they identified *bai-de* as /LH-Ø/白的 ‘white-NML’ ‘the white one’ when the second syllables are 130-90 Hz and 140-100 Hz. The results show that the listeners used the relative pitch differences between the first and the second syllables to identify the tonal patterns. /LH-Ø/ was identified because of the high falling pitch contour of the second syllable, and /H-Ø/ was identified because of the low falling pitch contour of the second syllable. The results also show that the pitch contour of the neutral-tone syllable is also important because no listeners identified any of the stimuli as *bai-de* /L- Ø/ 擺的 ‘to place-NML’ ‘the placed one’, in which the neutral tone has a level pitch contour.

2.2. Taiwan Mandarin

Taiwan Mandarin is a dialect of Mandarin, but rather than developing simply as a regional dialect, it has a rather special historical background. In 2.2.1, I will first introduce the linguistic history and the linguistic situation in Taiwan. In 2.2.2, I will briefly discuss the formation of the Taiwan Mandarin dialect, and then I will summarize the phonological and phonetic features of Taiwan Mandarin.

2.2.1. Taiwan as a linguistic area

The earliest inhabitants of Taiwan were Austronesian people. Around four hundred years ago, influxes of Han Chinese from coastal area of Southern China started to build settlements in Taiwan. The majority of them came from the Fujian province of China and spoke Southern Min[3], and some of them were Hakka speakers who came from the Guangdong province. In 1896, Taiwan was ceded to Japan after the First Sino-Japanese War, and the Japanese colonization lasted almost 50 years. To this day, many Taiwanese over 70 years old can speak some Japanese. In 1945, Taiwan was returned to the Republic of China (ROC) government after the end of World War II. Shortly after, the ROC government lost control over mainland China after losing the Chinese civil war to the Communist Party (which later established the People's Republic of China) in 1949. As a result, the ROC government withdrew to Taiwan, bringing about 1.2 million of the mainland population to Taiwan. The mainlanders are commonly referred to as *Wai-sheng-ren*外省人 (literally "outside province people"), and a majority of them spoke Mandarin as a native language. This history of Taiwan resulted in four main identity groups in Taiwan: as of 2002, 76.9% of Taiwan's population are Southern Min, 10.9% are Hakka, 1.4% are Formosan, and 10% are Mainlanders and their descendants[4].

---

[3] Southern Min is another Chinese variety spoken in Taiwan and Fujian area. It belongs to Min, a different Chinese language.

[4] *Survey on Living Conditions of Citizens in Taiwan-Fuchien Area for the year of 2002 九十一年台閩地區國民生活狀況調查*, published by Department of Statistics, Ministry of the Interior, Republic of China.

However, the linguistic situation does not match up with the ethnic composition. With the establishment of the ROC government, Mandarin, commonly known as *guoyu* 國語 (literally "national language"), was highly promoted and the other local languages such as Formosan languages, Hakka, and Southern Min were prohibited in the public domain. The strict Mandarin Language Policy was enforced in schools from 1945 to 1987, and the school-based policies had an impact on family-based speaking practices (Sandel 2003). With the painful Mandarin learning experience in school, the non-Mandarin speakers started to speak Mandarin to their children. As a result, Mandarin became more dominant in formal domains and non-Mandarin local languages are more dominant in informal domains (Tse 1983; Van den Berg 1986). The generation that started school in the 1970s and 1980s in general are not as fluent in Southern Min as they are in Mandarin (Young 1988; Huang 1993; Sandel 2003).

After 1987, Martial law was lifted, and so was the Mandarin Language Policy. The non-Mandarin local languages are no longer banned in the schools. With democratization in Taiwan in the 1990s, the language situation began to turn around. The local languages extended their use to the public domain, and speaking a local language in public or speaking Mandarin with an accent was no longer shameful, but was deemed as a symbol of one's willingness to identify with Taiwan (Mo 2000).

Even with the easing of the Mandarin only policy and the rise of the local languages, Mandarin had already become the dominant language. According to the 2010 census[5], among all the Taiwanese above 6 years old, 83.6% of the population use Mandarin at home, 81.9% of the population also use Southern Min at home, and only 6.6% and 1.4% of the population use Hakka and Formosan languages at home. The percentages across different age groups show a strong trend of language shift. As shown in Table 2.5, Mandarin is much more dominant among younger generations. We can see that the percentages of using Mandarin rise from 45.3% to 96% as age decreases. On the other hand, the use of Southern Min or Hakka decreases as age decreases. Van den Berg (1986) compared the language ability of three generations and also found that Mandarin became more prominent among the younger generations.

[5] 99 年人口及住宅普查, conducted by Directorate General of Budget, Accounting and Statistics, Executive Yuan, R.O.C. 行政院主計總處

Table 2.5 Percentage of Language use at home of Taiwanese above 6 years old from 2010 census*

Age	Population	Percentage of Language use (%)				
		Mandarin	Southern Min	Hakka	Formosan Languages	Others
6 - 14	2,418,610	96.0	69.7	3.8	1.0	0.8
15 - 24	3,146,521	94.9	78.6	4.8	1.3	1.0
25 - 34	3,799,930	91.9	83.2	5.6	1.3	1.8
35 - 44	3,531,622	90.4	84.1	6.4	1.5	2.3
45 - 64	6,068,715	78.9	86.3	8.1	1.5	2.6
65-	2,441,837	45.3	81.7	10.1	1.3	3.1
Total/Average	21,407,235	83.6	81.9	6.6	1.4	2.0

*Each person may use more than one language at home.

Also, many researchers observed the diglossic situation in the society, with Mandarin as the high variety (Tse 1983; Liao & Lii-Shih 1993; Tsao 1999). For bilingual speakers of both Southern Min and Mandarin, Southern Min is mainly used at home or only between interlocutors of close relationships, while Mandarin is dominant in all public domains. Southern Min also tends to be associated with in-group solidarity. In contrast, Mandarin is the standard language, and its authority is still maintained in many social domains (M.-y. Chang 1996). Code-switching between Mandarin and Southern Min is also common as a tool to negotiate interpersonal relationship (Huang 2000; Liao 2000; Su 2001; Wei 2003).

There is also geographic difference in terms of language use. Van den Berg (1986) reported that there is an urban vs. rural difference where Mandarin is more dominant in the urban areas while Southern Min is more dominant in the rural areas. He also found that generally Mandarin is used more in the north and Southern Min is used more in the south. Liao (2000) also observed that the code-mixing between Mandarin and Southern Min is much more common in the south than in the north. The frequency increases as one goes south.

2.2.2. Taiwan Mandarin dialect

In this study, I focus on the dialect variation between Standard Mandarin and Taiwan Mandarin. Mandarin can be classified into four subgroups: northern, northwestern, southwestern, and eastern (*Jiang-Huai*) (Norman 1988, 190-191). Standard Mandarin

(SM) is based on the dialect of Beijing Mandarin, a northern dialect. Taiwan Mandarin (TM), a Mandarin dialect spoken in Taiwan, and has been influenced by Southern Min. Kuo(2005) used the first Taiwan population census in 1956 and described the origins of the Mainland immigrant. She further analyzed the linguistic features of these regional varieties and found that several phonological features of TM such as the absence of retroflexion were in fact the outcome of a koineization process among the mainlanders. Hsu (2005) further showed that aside from the koineization among the mainlanders, there was also a phonological leveling between the Mandarin varieties of mainlanders and the Southern Min speakers since 1950. Thirty years later the leveling of Taiwan Mandarin had almost been completed, as the speakers born between 1981 and 1990 shared similar phonological features regardless of their ancestries.

Over the span of sixty years, special features in terms of phonology, lexicon and syntax developed in Taiwan Mandarin (Kubler 1979; Cheng 1985; Kubler 1985; Swihart 2003; H.-I. Tseng 2003). Although many Taiwanese speak Mandarin as a second language, most of the younger generations (born after 1980s) speak Mandarin as a first language. This Mandarin variety spoken in Taiwan has distinct features, therefore it should be treated as a separate dialect, Taiwan Mandarin (Huang 1993; Her 2010). In the following sections, I will review the phonological differences between SM and TM, including consonants, vowels, and pitch patterns.

2.2.2.1. **Consonants**

One of the most salient features of TM is the articulation of retroflexes. Many scholars have pointed out that the retroflexed initials /ʈʂ, $ʈʂ^h$, ʂ, ɻ / tend to merge into alveolar initials /ts, ts^h, s, z or l/ in TM (Kubler 1985; C.-y. Chen 1991; Duanmu 2000). Aside from that, /f/ is merged with /h/ or has a phonetic variant [ɸ]. Some examples are listed in Table 2.6.

Table 2.6 Initial consonants correspondences

	Correspondence		Example		
pinyin	SM	TM	SM	TM	Gloss
zh	[ʈʂ]	[ts]	[ʈʂɤ]	[tsɤ]	‘this’
ch	[ʈʂ^{h}]	[ts^{h}]	[ʈʂ^{h}a]	[ts^{h}a]	‘tea’
sh	[ʂ]	[s]	[ʂɨ]	[sɨ]	‘ten’
r	[ɻ]	[z] [l]	[ɻɨ] [ɻu]	[zɨ] [lɨ]	‘day’
f	[f]	[ɸ][h]	[fu]	[ɸu][hu]	‘luck’

Not all TM speakers merge their initials. Some TM speakers have completely merged these initials; some speakers merge these initials in fast speech; and some TM speakers still maintain the phonemic distinctions. Tse (1998) compared the pronunciations of /ʈʂ, ʈʂ^{h}, ʂ/ vs. /ts, ts^{h}, s/ of college students and reported that females generally maintain the distinction better than males. Also, it was more common for /ʈʂ/ to merge with /ts/ than for /ʈʂ^{h}, ʂ/ to merge with /ts^{h}, s/.

Furthermore, the phonetic realizations of the retroflexes are quite different in TM. The degree of tongue retraction of the retroflexes in TM is not as extreme as in SM. The retroflexes /ʈʂ, ʈʂ^{h}, ʂ/ in TM are only slightly retracted to the post-alveolar area [tʃ, tʃ^{h}, ʃ][6] (Chung 2006). The difference between /su/ ‘crunchy’ and /ʂu/ ‘book’ is often realized as [su] and [ʃu] in TM for those speakers who still maintain the phonemic distinction.

The mergers in the syllable final position are the other salient features of TM because these final consonant mergers or deletions are found in most TM speakers. The examples are illustrated in Table 2.7.

Table 2.7 Final consonants correspondences

Correspondence		Example		
SM	TM	SM	TM	Gloss
[əɻ]	[ɤ]	[əɻ]	[ɤ]	‘two’
[in]	[iŋ]	[jin]	[jiŋ]	‘music’
[əŋ]	[ən]	[ʂəŋ]	[ʂən]	‘raw’
[ən]	[əŋ]	[ɻən]	[ɻəŋ]	‘merciful’

[6] The post-alveolar [tʃ, tʃ^{h}, ʃ] are similar to those in English, except for they are not rounded.

In TM, most speakers have completely lost final [ɻ]. For example in words like *er-qie* [əɻ-tɕʰjɛ] 'and', most TM speakers drop the final [ɻ] and pronounce it as [ɤ-tɕʰjɛ]. As a result, the *-er* diminutive suffix, which is a main characteristic of Beijing Mandarin and many Northern dialects, is not productive in TM. Only certain vocabulary items prescriptively keep the *-er* suffix such as [i-tjəɻ] 'a little bit' and [ʈʂəɻ] 'here'. However, most TM speakers either omit the suffix or pronounce the suffix as an independent, non-rhotic syllable. For example [i-hwəɻ] 'a while' is pronounced [i-hwə] or [i-hwə- ʔɤ] 'a while'.

A final nasal merger is found in TM (C.-y. Chen 1991; Tse 1992). Furthermore, the ethnic gap of the final nasal mergers has been leveled. This feature is found among TM speakers regardless of their ancestry (H.-j. Hsu & Tse 2007). Many TM speakers confuse /n/ with /ŋ/ especially when it is after the vowel /i/. As for the direction of the merger, Chen (1991:146-47) found that /-in/ is merging into /-iŋ/, and /əŋ/ is merging into /ən/. Tse (1992) did a study on the production and the perception of the final nasals in TM, and concluded that TM is in the stage of merging /-ŋ/ into /-n/. Although Chen and Tse disagree on the direction of the merger when the preceding vowel is /i/, they both claimed that /əŋ/ is merging into /ən/, and that the /aŋ/ and /an/ distinction is the most stable due to the allophonic differences of /a/ (a→ɛ/_n; a→ɑ /_ŋ).

Phonetically, in SM, /əŋ/ is realized as [ʌŋ] regardless of the initial consonants. In TM, /əŋ/ becomes [ɔŋ] after labials, and /əŋ/ remains [əŋ] after non-labials. For example, /fəŋ/ 'wind' is pronounced [fɔŋ] in TM and [fʌŋ] in SM; while /kəŋ/ 'more' is pronounced [kəŋ] in TM and [kʌŋ] in SM. TM speakers seem to use the vowel distinction to be the acoustic cue. Therefore, /ən/ and /əŋ/ are not confused after labials in TM.

2.2.2.2. **Vowels**

In TM, certain vowels or glide-vowel combinations have been replaced by other vowels as shown in Table 2.8. These features have been reported by Kubler (1985) and Tsao (2000). These vowel mergers are less common compared to the initial or final consonant mergers. TM speakers with such phonological features are considered to have a strong Southern Min accent.

Table 2.8 Vowels correspondences

Correspondence		Example		
SM	TM	SM	TM	Gloss
[y]	[i]	[y]	[i]	‘rain’
[wo]	[ɔ]	[kwo]	[kɔ]	‘country’
[jɛ]	[ɛ]	[tʰjɛ]	[tʰɛ]	‘to stick’
[ɨ]	[u] ~ [ɯ]	[ʈʂɨ]	[tsu]~ [tsɯ]	‘to know’
[ɤ]	[ɔ]	[kɤ]	[kɔ]	‘older brother’

The first two examples in Table 2.8, the delabilialization of /y/ and the glide deletion of /wo/, were used as variables in Tseng and Huang’s research (2010) of Taiwan Mandarin. They found that these phonological variants are highly correlated with the degree of Southern Mandarin exposure, which they quantified by whether Southern Min is used for communicating with the parents and the siblings. On the other hand, they found that socio-economic factors did not correlate with the use of the TM variants.

Another feature of TM vowels is that the rimes /əu/ [ou] and /əi/ [ei] are usually monophthongized into [o] and [e]. For example [tou] ‘all’ is pronounced [to] and [fei] ‘to fly’ is pronounced [fe] in TM. These monophthongizations are also found in words with prenuculear glides, [ljou] ‘six’ is pronounced [ljo] and [twei] ‘correct’ is pronounced [twe] in TM. These features can be found in the lyrics in Taiwanese pop music. /-ei/ can rhyme with /-e/ and /-ou/ can rhyme with /o/. For example in the refrain of Jaye Chou’s song *qilixiang* 七里香 ‘Orange Jasmine’, /je/夜 rhymes with /ʂwei/水, /je/葉, /tje/疊, /fei/非, /tɕye/卻, /je/頁. /e/ can rhyme with /ei/ and /e/.

Also, the vowels in /ai/ and /au/ in TM are not raised as much as in SM. /ai/ and /au/ are pronounced as [æɛ] and [ɑɔ] in TM. For example, /mai/ ‘to sell’ is pronounced [mæɛ] and /mau/ ‘hat’ is pronounced [mɑɔ].

Chang(1998) investigated the acoustic properties of the monophthongs in Taiwan Mandarin, and she found that the vowel space in Taiwan Mandarin is smaller than Standard Mandarin. Also, the vowel targets in Standard Mandarin are implemented much faster.

2.2.2.3. **Pitch differences**

In TM, the pitch height is lower than SM and the pitch range is smaller than SM. Fon and Chiang (1999) revisited the five tone values based on Chao's scale and found that opposed to Chao's description, where the five point scale of /H, LH, L, HL/ in SM are [55], [35], [214] and [51], the five point scale of /H, LH, L, HL/ in TM are [44], [323], [312], and [42]. Torgerson (2005) compared 6 speakers each from SM and TM, and the result also showed that SM has higher and wider tone register than TM.

TM and SM also differ in terms of the distribution of the /L/ variants. In SM, the pitch contour of /L/ (the dipping tone [214]) becomes low falling only at non-final position. In contrast, /L/ in TM is reported to have the low-falling variant in both isolated or final position (Kubler 1985; Lo 1991; Sanders 2005; Deng et al. 2008). In the production test done by Fon et al. (2004, 273), only 43.24% of the Tone 3 in isolated syllables recorded were dipping, more than half of /L/ was low-falling.

Also, the TM /LH/ tone does not rise as much as the SM /LH/ tone. The five point scale of TM tones investigated by Fon and Chiang (1999) shows that both /LH/ and /L/ in TM have a dipping contour in the mid-low pitch range. Fon, Chiang & Cheung (2004) followed up and compared /LH/ and /L/. They found that in TM, /L/ tends to have a creaky voice quality toward the end, while /LH/ has modal voice throughout; /L/ has steeper falling and rising portion, but the rising portion is longer in /LH/. Their perception tests show that the TM listeners rely on the steeper initial pitch fall to recognize /L/, while they rely on the final rising pitch to recognize /LH/. The rising portion in /L/ does not contribute in the identification of /L/, even though it has a steeper rise. Fon et al. (2004) suggests that it is because most of the /L/ of TM are low-falling rather than dipping even at final position.

2.2.2.4. **Neutral-tone syllables**

As mentioned earlier, unstressed syllables are less frequent in TM (Kubler 1985, 161; Tsao 2000; Swihart 2003, 110; Duanmu 2007, 308). The second syllables of reduplications or of many disyllabic words are rarely unstressed, while they are unstressed in Standard Mandarin. Some phonemic distinctions are therefore lost. For example, the distinction between *dong-xi* /H-H/ 'east-west' and *dong-xi* /H-Ø/ 'things' are lost, both of them are *dong-xi* /H-H/ in Taiwan Mandarin (Lo 1991). The second syllables do not reduce their vowels or lose their original tones in TM. Some examples are shown in Table 2.9.

Table 2.9 Reduplication and disyllabic words in SM and TM

	pinyin	Gloss	SM	SM tone	TM	TM tone
1	*xing-xing*	'star'	[ɕiŋ-ɕĩ]	/H-Ø/ [55-2]	[ɕiŋ-ɕiŋ]	/H-H/
2	*kan-kan*	'take a look'	[kan-kə̃]	/HL-Ø/ [51-1]	[kan-kan]	/HL-HL/
3	*da-suan*	'plan'	[ta-suə̃]	/L-Ø/ [21-4]	[ta-suan]	/L-HL/
4	*ming-bai*	'understand'	[miŋ-pɛ]	/LH-Ø/ [35-3]	[miŋ-pai]	/LH-LH/

Table 2.9 shows that TM speakers do not reduce the second syllables in these reduplications and compounds. These syllables maintain their original tones. The second syllables are not unstressed. These variations are even noticed by the government, and these syllables are spelled with the original lexical tones in the dictionary that is compiled by the Ministry of Education of Taiwan.

Table 2.10 Unstressed grammatical morphemes in SM and TM

pinyin	Gloss	SM	SM tone	TM	TM tone
mai-guo	buy-EXPE 'had bought'	[mai-kwə]	L-Ø [21-4]	[mai-kwo]	L-HL [21-51]
qu-guo	go-EXPE 'had gone'	[tɕʰy-kwə]	HL-Ø [51-1]	[tɕʰy-kwo]	HL-HL [51-51]
kan-qi-lai	look-up-come 'looks like'	[kan-tɕʰi-lə]	HL-Ø-Ø [51-1-1]	[kan-tɕʰi-lai]	HL-L-LH [51-21-35]

Also, as shown in Table 2.10, some grammatical morphemes are unstressed in SM but not in TM. For example the experiential aspect marker *guo* [kwo] and many directional complements such as *shang* /ʂaŋ/ 'up' and *lai* /lai/ 'come' are not unstressed in TM; their vowels are not reduced and their lexical tones remain. They are also spelled with a full tone by the Ministry of Education of Taiwan.

In order to understand the scale of the differences of the neutral tone between TM and SM, I compared the words in the List of neutral-tone words for the Standard

Mandarin Proficiency Test[7] with the online Revised Mandarin Chinese Dictionary published by the Ministry of Education in Taiwan. Out of 545 words in the list of the SM neutral-tone words, 402 of them (73.8%) were marked as having a neutral tone in the dictionary provided by the Ministry of Education in Taiwan, 134 of them (24.6%) were marked as having a lexical tone on those neutral-tone syllables. Six of them (1.1%) were marked as free variations; both lexical and neutral tone are acceptable. The detailed result is shown in Table 2.11.

Table 2.11 The neutral-tone words tonal marking in TM

Word types		Marked tone in Taiwan	
		Neutral tone	Full tone
X-bound morphemes (244)	*-zi* 子 'DIM'	206	0
	-men 們 'PL'	7	0
	-tou 頭 'DIM'	21	0
	-me 麼 'such'	5	0
	-ge 個 'CL'	1	0
	-le 了 'PERF;	2	0
	-de 得 'Complement marker'	1	0
	-de 的 'NML'	1	0
Reduplicants (20)		20	0
Compound words(272)		138	134
- 6 compounds were marked as free variations. - 3 compounds were not available in the dictionary.			

As shown in Table 2.11, all the suffixes and the reduplicants in the list are marked as having the neutral tone in Taiwan Mandarin. However, only about half of the compound words (138/272) were marked as having the neutral tone; the other half of the compound words (134/272) were marked with their original lexical tones. For example, one of the words on the list is *xia-ba* /HL-Ø/ 'chin', and it is also marked as having /HL-Ø/ in TM. However, *yue-liang* /HL-Ø/ 'moon' on the list is marked as /HL-HL/ in TM. This

[7] *Putonghua shuiping ceshi yong qingsheng cibiao* 普通话水平测试用轻声词表, which is a study guide of neutral-tone words for non-native speakers to study in order to pass the Standard Mandarin Proficiency Test held by the Chinese government.

comparison shows that only half of the full-neutral disyllabic words are prescriptively marked as having full-neutral tones in Taiwan Mandarin.

However, the prescriptive markings do not seem to reflect TM speakers' phonological representation of these words with neutral tones. It seems that only some of the prescriptively marked neutral-tone syllables are recognized by Taiwan Mandarin users. I examined these 402 prescriptively neutral-tone words by trying to input these words in the computer's system. One of the most common character input methods in Taiwan is the New Phonetic Input[8] on the Microsoft operating system, in which the users have to input both the segment and the tone (4 lexical tones and 1 neutral tone) in order to input the character. After attempting to input the 402 prescriptively neutral-tone words, I found that many prescriptively neutral-tone syllables cannot be found under the neutral-tone entry. All of the suffixes are found under the neutral-tone entry by entering the segments and the neutral tone. For example, the suffix *-zi* /Ø/ 子 can be input by typing in *zi* (ㄗ in *zhuyin*) + neutral tone (˙ in *zhuyin*). However, 4 out of 20 reduplicated items cannot be found under the neutral-tone entry. For example, *xīng-xing* /H-Ø/星星 'star' has a neutral tone in the dictionary. However, when typing *xing* + neutral tone, the character was not available. The character is only available when typing *xing* plus high tone. Also, 116 out of 138 (84%) of the TM prescriptive full-neutral compound words do not have neutral-tone input on their neutral-tone syllables. For example, in the dictionary in TM, the syllable *fu*服 in *shu-fu* /H-Ø/舒服 'comfortable' has a neutral tone. However, the character *fu* /Ø/服 is not available when inputting *fu* + neutral tone; we have to type *fu* + rising tone. The computer input system in this way reflects the users' phonological representations of the syllables. These full-neutral disyllabic words were stored as full-full in the input system in order to let TM users input the words in the most intuitive way. This piece of evidence shows that even though the Ministry of Education in Taiwan has already recognized that there are fewer full-neutral compound words in TM, there are even fewer full-neutral compound words existing in TM speakers' phonology.

---

[8] 新注音輸入法。New Phonetic Input is a phonetic input method operated in the Microsoft operating system in which users can type characters by inputing the phonetic elements using *zhuyin fuhao* (also known as *bopomofo*), which is the phonetic spelling system taught and used in Taiwan.

On the other hand, there are still some unstressed syllables in TM; that is, they are spelled with a neutral tone, and TM speakers presumably know that these syllables are spelled with a neutral tone. Although many reduplications and compound words nowadays obviously lose the neutral tone as shown in Table 2.9, some reduplication and compounds still have a prescriptive neutral tone. Table 2.12 shows some examples.

Table 2.12 Reduplication and disyllabic words with neutral tones

Gloss	Pinyin	IPA	prescriptive tone	TM tone
'younger brother'	*di-di*	/ti-ti/	HL-Ø	HL-?
'uncle'	*shu-shu*	/ʂu-ʂu/	LH-Ø	LH-?
'aunt'	*shen-shen*	/ʂən-ʂən/	L-Ø	L-?
'older brother'	*ge-ge*	/kə-kə/	H-Ø	H-?
'radish'	*luo-bo*	/lwo-pwo/	LH-Ø	LH-?
'ear'	*er-duo*	/əɻ-two/	L-Ø	L-?
'key'	*yao-shi*	/jao-ʂi/	HL-Ø	HL-?
'warm'	*nuan-huo*	/nwan-hwo/	L-Ø	L-?

As illustrated in Table 2.12, these vocabulary items contain a neutral tone. The second syllables do not obviously resemble their original tone. For example, the first word *didi* is not pronounced as /HL-HL/. The pitch of *duo* in *er-duo* /L-Ø/ does not resemble the original tone *duo* /L/ of the syllable. However, the phonetic nature of these neutral-tone syllables is unclear. We do not know whether these syllables have a neutral tone like those in SM. In addition, the rimes in the second syllables are not obviously reduced like those in SM by observation. It is not clear whether the second syllables are unstressed.

Aside from the lexical items, many grammatical morphemes are also marked as having a neutral tone. Table 2.13 is a list of these grammatical morphemes and particles. There are also many sentence final particles that have a neutral tone. They are: *ba* 吧, *na* 哪, *ne* 呢, *la* 啦, *lo* 咯, *ya* 呀, *ye* 耶, *wa* 哇, *a* 啊, *en* 嗯.

Table 2.13 Prescriptively neutral-tone grammatical morphemes in Taiwan Mandarin

character	Pinyin	IPA	Function	syllables before destressing
的	*de*	[tə]	possessive (POSS) nominalizer (NML)	*di* /HL/ ‘target’
地	*de*	[tə]	adverbializer	*di* /HL/ ‘land’
得	*de*	[tə]	complement marker	*de* /LH/ ‘get’
了	*le*	[lə]	perfective (PERF)	*liao* /L/ ‘finish’
著	*zhe*	[ʈʂə]	durative (DUR)	*zhao* /LH/ ‘contact’
們	*men*	[mən]	plural (PL)	*men* /LH/ ‘PL’[9]
子	*zi*	[tsɨ]	diminutive (DIM)	*zi* /L/ ‘son’
頭	*tou*	[tʰou]	diminutive (DIM)	*tou* /LH/ ‘head’
個	*ge*	[kə]	classifier (CL)	*ge* /HL/ ‘individual’
嗎	*ma*	[ma]	question particle (Q)	unknown
呢	*ne*	[nə]	question particle (Q)	unknown

It is unclear whether these syllables in Table 2.13 have a neutral tone in TM. As mentioned above, the pitch of a syllable with a neutral tone is decided by the tone of the preceding syllable. However, in Taiwan Mandarin, by observation, the pitch of a neutral tone syllable does not seem to be decided by the preceding syllable. This can be illustrated by Table 2.14. As shown in Examples 1 and 2, the perfective suffix /lə/ in TM has a low pitch regardless of the preceding lexical tone. If /lə/ has a neutral tone, I would expect /lə/ to have different pitches depending on the preceding tones as shown in SM. Examples 3-6 also show that the possessive suffix /tə/ and the question particle /ma/ are not neutral tones in TM because their pitches do not change according to the tone of the preceding syllables as in SM.

[9] The character *men* 們 originally was /HL/ ‘fat’ when it was first documented in the dictionaries. In early *Qing* dynasty (17th century), the character was used as a phonetic element to represent the plural suffix *men* /LH/.

Table 2.14 The pitches of neutral-tone syllables in SM and TM.

		Gloss	IPA	tone	SM pitch	TM pitch
1	*mai-le*	buy-PERF 'bought'	mai-lə	L-Ø	[21-4]	[21-1]
2	*qu-le*	go-PERF 'went'	$tɕ^h$y-lə	HL-Ø	[51-1]	[51-1]
3	*wo-de*	1.S-POSS 'my'	wo-tə	L-Ø	[21-4]	[21-1]
4	*ta-de*	3.S-POSS 'his/her'	ta-tə	H-Ø	[55-2]	[55-1]
5	*hao-ma*	good-Q 'good?'	hao-ma	L-Ø	[21-4]	[21-5]
6	*qu-ma*	go-Q 'go?'	$tɕ^h$y-ma	HL-Ø	[51-1]	[51-5]

These observations are confirmed by Tseng (2004). She compared five sentences produced by 6 SM and 6 TM radio announcers, and she found that the neutral tone in TM has a low entering tone regardless of the preceding tones. Li (2005) also compared the neutral tone between the speakers of North Mainland China and Taiwan, and concludes that the neutral tone in TM "has a certain target" regardless of the preceding lexical tones, suggesting the syllables have a certain lexical representation. If these syllables have a certain lexical representation, they are not toneless. If that is the case, it implies that these syllables were assigned new lexical tones. We might hypothesize that the perfective suffix /lə/ and the possessive suffix /tə/ are assigned a low tone in TM as shown in 1-4 in Table 2.14; on the other hand, the question particle /ma/ acquired a high tone. This hypothesis still remains to be tested.

It is also unclear whether these syllables are unstressed. Although many of these morphemes such as *de*, *le*, *zhe*, *ne* have a reduced vowel /ə/, they are grammaticalized from stressed lexical items historically. Synchronically the lexical representation of these morphemes contains /ə/, and the original syllables before the diachronic destressing are not accessible because they have been grammaticalized. It is possible to have a stressed /ə/ in TM such as *de* /tə/ /LH/ 'to receive'. However, the quality difference between a stressed /ə/: [ɤ] and an unstressed /ə/: [ə] is not obvious (M. Lin & Yan 1990). It is hard to observe obvious vowel reduction like those in disyllabic words. Other morphemes also do not have obvious vowel reduction. For example, although the neutral-tone *tou* / t^hoʊ/ in TM is pronounced [t^ho], the off-glide is deleted everywhere, as mentioned above in the vowel section. There is no obvious vowel reduction observed. Also, Duanmu (2000, 266-267) noticed that, unlike in SM, the vowel devoicing and deletion are not found in unstressed syllables in TM. Consonant reduction is found in TM, but it is not restricted to unstressed syllables; it can happen in stressed syllables too. Even consonants in word initial position can reduce in Taiwan Mandarin, for example [xən-dwo] ‘many’ becomes [ɦən-dwo], or [zan-xəu] ‘afterwards’ becomes [ã-əu]. This may be another manifestation of the rhythmic differences between SM and TM.

CHAPTER 3. PRODUCTION OF NEUTRAL-TONE SYLLABLES

This chapter discusses the production of the neutral-tone syllables in Taiwan Mandarin. I hypothesize that the pitch contours of the neutral-tone syllables in Taiwan Mandarin do not behave like those in Standard Mandarin. The pitch contours of the neutral-tone syllables seem to have a low target regardless of the preceding lexical tones. In order to test my hypothesis, I design a series of production tests. Section 3.1 investigates the pitch contours of the neutral-tones syllables in disyllabic words. Section 3.2 tests the neutral-tone syllables in novel occurrences. Section 3.3 explores the pitch contours of the consecutive neutral tones in Taiwan Mandarin.

3.1. Pitch Contours of Neutral-tone Syllables in Disyllabic Words

The first production test looked at the pitch contours of the neutral-tone syllables in disyllabic words in Taiwan Mandarin. The subjects were asked to read frequent disyllabic words that contain a lexical tone and a neutral tone.

3.1.1. Research design

Several prescriptively neutral-tone syllables that are fairly frequent were selected. These syllables include grammatical markers such as perfective *le*, possessive/nominalizer *de*, durative *zhe*, plural marker *men*, diminutive suffixes *zi* and *tou*, derivational suffix 'like' -*me* and classifier *ge*. I also included the second syllables of reduplications such as *didi* /HL-Ø/ 'younger brother' because all these syllables are prescriptively neutral-tone.[10]

In order to test whether or not these syllables show neutral-tone features, these syllables were elicited after syllables with the four different lexical tones in Mandarin (H,

[10] Sentence final particles such as *ma, ne, le, a, o* also have neutral tones. They are not included in the test because it is unnatural to have subjects read final particles in a short phrase without any context. There are also some other disyllabic compound words with neutral tones on their second syllables prescriptively. These words are not included in this study because many of them obviously have lexical tones descriptively.

LH, L, HL). For example, in order to observe the features of neutral-tone perfective suffix -*le*, subjects were asked to read *chi-le* /H-Ø/ 'eaten', *lai-le* /LH-Ø/ 'come', *mai-le* /L-Ø/ 'bought', and *qu-le* /HL-Ø/ 'gone'. In Taiwan Mandarin, the pitches of the neutral-tone syllables do not seem to be determined by the lexical tone of the previous syllables. Instead, I expect these neutral-tone syllables to retain a stress, which means that they keep the original lexical tones from being lost, or that they have acquired a new lexical tone that does not depend on the preceding lexical tone.

Table 3.1 shows the list of the pitch patterns that I expected these syllables to show. Some neutral-tone syllables do not have a known original syllable before destressing such as -*me*; some neutral-tone syllables do not have segmentally identical original syllables such as -*le*, -*de*, and -*zhe*. In that case, I expect these syllables to have a low pitch target. On the other hand, some neutral-tone syllables have segmentally identical original syllables before destressing. In that case, I expect the neutral-tone syllables to be pronounced with either the original tone or with a low pitch. For example, the diminutive suffix *tou* /Ø/ was destressed from *tou* /LH/. In that case, I expect *tou* to be pronounced with its original tone /LH/ or with a low pitch.

Table 3.1 The original lexical tones and the expected new tones of the tested neutral-tone syllables

neutral-tone syllables	original syllables before destressing	segmentally identical	expected new pitch pattern
-*le* 'PERF'	*liao* /L/ 'finish'	no	low
-*de* 'POSS'/ 'NML'	*di* /HL/ 'target'	no	low
-*zhe* 'DUR'	*zhao* /LH/ 'contact'	no	low
-*me* 'such'	unknown	no	low
-*ge* 'CL'	*ge* /HL/ 'individual'	yes	low or HL
-*zi* 'DIM'	*zi* /L/ 'son'	yes	low (L)
-*men* 'PL'	*men* /LH/ 'PL'	yes	low or LH
-*tou* 'DIM'	*tou* /LH/ 'head'	yes	low or LH
(reduplicants)	the reduplicated syllables	yes	H / H_ low /[LH, L, HL]_

3.1.2. Materials

As mentioned above, eight suffixes that are prescriptively neutral-tone syllables were selected. These syllables were elicited after syllables with the four different lexical tones in Mandarin (H, LH, L, HL). Aside from that, four reduplication forms with four different lexical-tone syllables being reduplicated were included. There were a total of nine (tested neutral-tone syllables)*4 (preceding tones) =36 words tested. The tested disyllabic words were elicited in a frame sentence: *qing shuo*_____*ba ci*. 'Please say______ eight times.' Table 3.2 is a list of the tested words.

Table 3.2 Tested disyllabic words with the neutral-tone syllables

Tested syllable	H-Ø	LH-Ø	L-Ø	HL-Ø
le 'PERF'	*chi-le* 'eaten'	*lai-le* 'come'	*mai-le* 'bought'	*qu-le* 'gone'
de 'POSS' or 'NML'	*suan-de* 'sour'	*tian-de* 'sweet'	*ku-de* 'bitter'	*la-de* 'spicy'
ge 'CL'	*ba-ge* 'eight'	*yi-ge* 'one'	*wu-ge* 'five'	*si-ge* 'four'
me 'such'	*duo-me* 'such'	*she-me* 'what'	*ze-me* 'how'	*na-me* 'such'
zhe 'DUR'	*ting-zhe* 'listening'	*du-zhe* 'reading'	*xiang-zhe* 'thinking'	*zuo-zhe* 'sitting'
zi 'DIM'	*zhuo-zi* 'table'	*jia-zi* 'clip'	*yi-zi* 'chair'	*nie-zi* 'tweezer'
men 'PL'	*ta-men* 'they'	*ren-men* 'people'	*wo-men* 'our'	*jiao-shou-men*[11] 'professors'
tou 'DIM'	*zhuan-tou* 'brick'	*she-tou* 'tongue'	*zhen-tou* 'pillow'	*nian-tou* 'thought'
(reduplication)	*ma-ma* 'mother'	*shu-shu* 'uncle'	*shen-shen* 'aunt'	*di-di* 'brother'

[11] A disyllabic /HL-Ø/ word with *-men* suffix was not available. As a result, a trisyllabic *jiao-shou-men* /HL-HL-Ø/ 'professors' was included to elicit the /HL-Ø/ *-men* occurrence.

3.1.3. Subjects

The subjects were 6 male and 6 female Taiwan Mandarin speakers who were between 21 and 33 years old without any speaking or hearing impairment. Subjects all grew up in Taiwan and their parents all speak Southern Min as their native language. Among the TM speakers, some speak Southern Min as their first language; some speak Mandarin as their first language. Even though the koineization of Taiwan Mandarin is almost completed (H.-j. Hsu & Tse 2009), this criterion is to exclude possible dialect influences from other Chinese languages/dialects such as Hakka, Wu, or other Mandarin dialects. Also, subjects have not resided in any foreign countries for more than three months in the past year to exclude possible language influences from other Mandarin varieties.

The subjects were first asked to finish the linguistic background form (see Appendix A). All twelve subjects speak Southern Min, Mandarin, and English. Of the 12 subjects, three grew up speaking only Southern Min at home. Two grew up speaking both Southern Min and Mandarin, and seven grew up speaking only Mandarin. As for their Mandarin abilities, three subjects considered their Mandarin “not so standard”, and only one subject considered her Mandarin “very standard”. The rest of the subjects considered their Mandarin “fairly standard”.

The subjects were notified of their rights as human subjects and asked to sign the consent forms. The subjects were also informed that their speech would be recorded, but their personal information and identity would remain confidential.

3.1.4. Equipment

The subjects were recorded with a Zoom H2 handy recorder in quiet offices in National Cheng Kung University. The speech was recorded with the sample rate of 44100 Hz.

3.1.5. Procedures

The subjects were asked to read 36 disyllabic words carried in the frame sentence *qingshuo XX baci* ‘Please say XX eight times’, where XX represents the target word. All the sentences were written in Chinese characters and the subjects were asked to read every sentence just once. The subjects could repeat the sentences if they felt they made a mistake. The subjects were asked to read 118 sentences in this task, including 36 tested sentences in this section, 30 tested sentences in Chapter Four which were also disyllabic words/phrases, and 52 filler sentences. The test words and fillers were randomly ordered, so the subjects would not notice the goal of the task and produce too carefully. The full list of the sentences is attached in Appendix B. The whole procedure took about 15 minutes.

3.1.6. Data Analysis

The recorded speech was analyzed in Praat. In this study, only the rimes of the syllables were segmented. In this way, it is easier to compare the F0 contour across different syllables. Otherwise a syllable with an initial obstruent only has F0 readings in half of the syllable while a syllable with an initial sonorant has F0 readings throughout the syllable. Phonetic studies also show that the pitch contours in the initial consonants are irregular, and the tones are not implemented until the beginning of the rimes (Howie 1976; Y. Xu 1999).

The tested syllables were hand labeled at the beginning and the end of the rime. The beginning of the rime was measured from the abrupt increase of the amplitude of the waveform. The end of the rime was measured to the beginning of the stop closure or the beginning of the fricative noise in the following word. The beginning and the end of F2 energy were used when the waveform was ambiguous. Segmentation was also cross-checked with the audio signals and the spectrogram.

The F0 of the rimes were measured with the aid of a Praat script.[12] The script divided any labeled interval (rime) into tenths, and measured the midpoint of all the 1/10 intervals. Therefore each rime had 10 F0 measurements distributed evenly.

In order to compare the pitches across different speakers, the F0 readings need to be normalized. Zhu (1999) compared six main normalization strategies and concluded that logarithmic z-score is the best strategy, because a z-score weighs in the F0 distribution of each speaker, and a logarithmic scale corresponds to human's perception of pitch. Adopting Zhu's analysis, the F0 readings in this study were normalized into semitone z-scores. Each F0 measurement was first transformed to a semitone[13], which is on a logarithmic scale. The semitones were then normalized to z-scores[14] for cross-speaker comparison. Each speaker's pitch range and pitch height were normalized into an approximate range between -2 ~ 2 semitone z-scores. The data were plotted using MS Excel. The x-axis abstractly represents a normalization of the time, and each number on the x-axis represents one-tenth of each rime. The y-axis represents the pitch in semitone z-scores.

The F0 readings were sometimes abnormal or absent due to the creakiness or breathiness of the production. Obvious pitch halving or pitch doubling were manually repaired. However, when no readings were available, the pitch contours were either excluded from the mean value, or the F0 values were manually inserted. Below are the criteria: because the second syllables of the disyllabic words are the syllables under investigation, three F0 points were selected from the rimes of the second syllables. They are the 2nd, the 5th, and the 8th F0 readings out of the 10 readings. Out of these three F0 readings, if one rime has two or more F0 readings unavailable, the F0 readings of this rime were excluded from the averaging process. At the same time, the data from its preceding syllable was also excluded. When the 2nd or the 5th F0 readings was not available, the neighboring F0 readings were used to apply linear interpolation in order to obtain the missing F0 readings. When the 8th F0 reading was missing, this suggests the

[12] The Praat script is originally written by Mietta Lennes, and later edited by Diana Stojanovic, Victoria Anderson, and Kaori Ueki.

[13] The semitone is calculated with a reference frequency of 50 Hz, which is close to the lowest pitch of all the speakers. Semitone = $12 * \log_2(F0/50)$.

[14] $z = (x-\mu)/\sigma$. X is the calculated semitone, μ is the average semitone of the speaker, and σ is the standard deviation of the speaker's semitone.

voice quality was creaky due to the lowering of the pitch. Therefore 0.5 semitone z-score was subtracted from the preceding available F0 reading. The creakiness of the syllables was verified by examining the spectrogram. If the repair method caused the F0 readings to be different from the bundle of the pitch contours, the whole pitch contour would be excluded from the averaging process.

3.1.7. Results

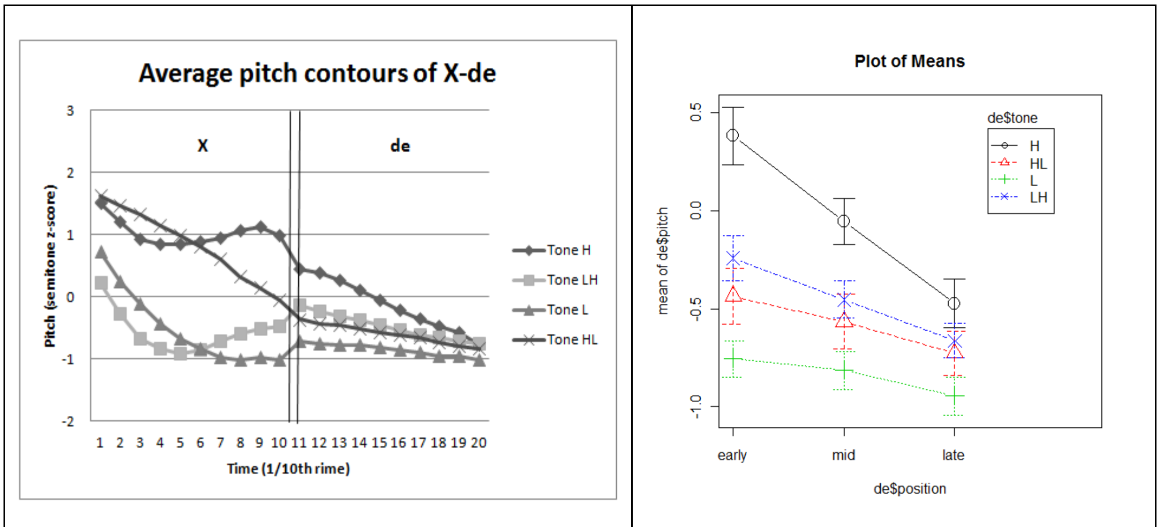
The results of the prescriptively full-neutral disyllabic words are discussed in the following sections. The neutral-tone suffixes with a /ə/ rime are discussed together in 3.1.7.1. The other neutral-tone suffixes are discussed in the following sections.

3.1.7.1. **Neutral-tone suffixes with a /ə/ rime**

Pitch patterns of perfective *le*, possessive/nominalizer *de*, durative *zhe*, derivational suffix ‘like’ *-me* and classifier *ge* are discussed in this section. All these neutral-tone syllables have a /ə/ rime because these unstressed forms were reduced to /ə/. As shown earlier in Table 3.1, out of these five neutral-tone suffixes, only the classifier *ge* has a segmentally identical original syllable before destressing: *ge* /kəHL/ ‘individual’. The rest of the four neutral-tone suffixes were grammaticalized from a segmentally different syllable where their relationships before and after destress are not obvious for most TM speakers.

3.1.7.1.1. X-*de*

Figure 3.1a shows the results of the full-neutral disyllabic words with the neutral-tone possessive *-de*. The chart illustrates the average pitch contours of all the speakers except for the data that was excluded as mentioned in 3.1.6. The chart shows the average pitch contours of two rimes: the first rime (X-axis 1-10) is the varied preceding syllable with lexical tone X, and the second rime (X-axis 11-20) is the tested neutral-tone syllable *-de*. There are double lines separating the two rimes to show that there are initial consonants in between. The four lines in the chart represent H-Ø, LH-Ø, L-Ø, and HL-Ø respectively. As shown in Figure 3.1a, the first rime (X-axis 1-10) shows the four pitch contours of the high, falling, rising, and low tones. The high tone /H/ slightly rose at the end, and the rising tone /LH/ had a dipping pitch shape where the pitch rise was very small. The falling tone /HL/ had a high fall while the low tone /L/ had a low fall. The second rime (Figure 3.1a, X-axis 11-20) shows the pitch contours of the rime of *de* lowering from the offset of the previous syllable to pitch that ranges between -0.5~ -1 semitone z-score. Figure 3.1b shows the mean pitches and the standard error of the four pitch contours at the different positions (Early, Mid and Late) in the second rime. “Early” is the 2nd of the ten measurement points within the rime; “Mid” is the 5th of the ten measurement points within the rime; and “Late” is the 8th of the ten measurement points within the rime.



a. Average pitch contours of *X-de*

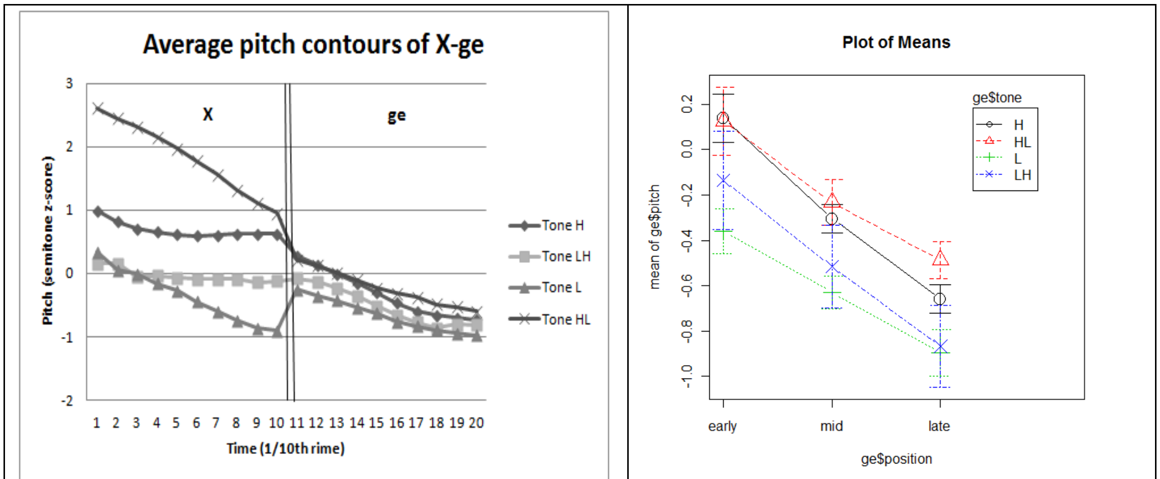
b. Pitch of *de* at early, mid, and late position

Figure 3.1 Result of full-neutral *X-de*

The influences of the preceding tones are obvious at the early and the mid position. A one-way analysis of variance shows that the influence of preceding tones on pitch was significant at the early position [$F(3,31) = 12.64$, $p < .001$] and the mid position [$F(3,31) = 7.39$, $p = < .001$]. However, the influences of the preceding tones were not obvious at the late position. A one-way ANOVA shows that the effect of preceding tone was not significant at the late position [$F(3, 31) = 2.81$, $p = 0.055 > .05$].

3.1.7.1.2. X-*ge*

Figure 3.2 shows the result of *X-ge*. Figure 3.2a shows the four pitch contours with four preceding lexical tones. As we can see, the pitch of the neutral-tone *ge* generally fell from around 0 semitone z-score to around -0.75 semitone z-score. Figure 3.2b further shows the mean pitches and their standard errors at the three positions in the *ge* rime.



a. Average pitch contours of *X-ge*

b. Pitch of *ge* at early, mid, and late position

Figure 3.2 Result of full-neutral *X-ge*

A one-way ANOVA shows that the influence of the preceding tones was significant at the early position [$F(3,31) = 2.91$, $p = 0.0499 < .05$] and the mid position [$F(3,31) = 2.9767$, $p = 0.047 < .05$], but the influence of the preceding tone was not significant at the late position [$F(3,31) = 2.8156$, $p = 0.055 > .05$].

Also, two of the productions of *X-ge* [LH-Ø] (*yi-ge*) were excluded in the data above. Subject M2 and M3 produced the word *yi-ge* [LH-Ø] 'one' with the pitch pattern of [HL-Ø]. The pitch contours of their productions of *yi-ge* are shown in Figure 3.3. This is a very interesting bit of data. In Mandarin, the underlying tone of *yi* 'one' is /H/, and *yi* 'one' undergoes a special tone sandhi rule[15] where /H/ becomes [LH] before /HL/, and /H/ becomes [HL] before /H, LH, L/. In this experiment I included *yi-ge* as a tested word because the underlying lexical tone of *ge* is /HL/. The tone sandhi rule ordinarily applies

[15] This tone sandhi rule only apply on four syllables: *yi* 'one', *qi* 'seven', *ba* 'eight', *bu* 'no'.

to *yi* first, changing *yi-ge* into [LH-HL], and then *-ge* is destressed after the tone sandhi rule. The processes are illustrated as follows:

Phonemic representation of *yi-ge*	/H-HL/
Tone Sandhi (H→ LH/ _HL)	[LH-HL]
Destressing	[LH-Ø]

But these two subjects produced [HL] for *yi* . This shows that the speakers did not treat *ge* as an unstressed syllable. Instead, they treated *-ge* like a /L/ syllable, perhaps because *-ge* has a rather low pitch target in other neutral-tone occurrences. As a result, a different tone sandhi rule applied: *yi* /H/ became [HL] before /L/. The processes are illustrated as follows:

Phonemic representation of *yi-ge* by M2 and M3	/H-L/
Tone Sandhi (H→ HL/__L)	[HL-L]

The application of the tone sandhi rule shows that *ge* was considered as a /L/ phonologically by these two speakers. The application of the tone sandhi rule gave us a valuable opportunity to find out the phonological representation of the neutral-tone *ge* by these two speakers.

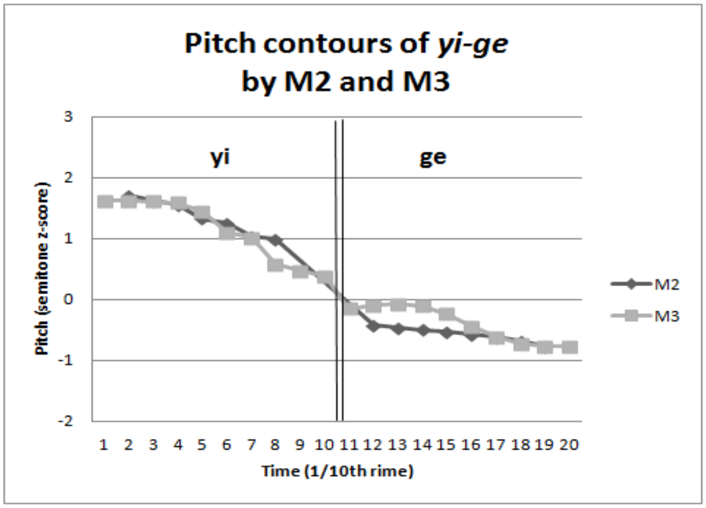
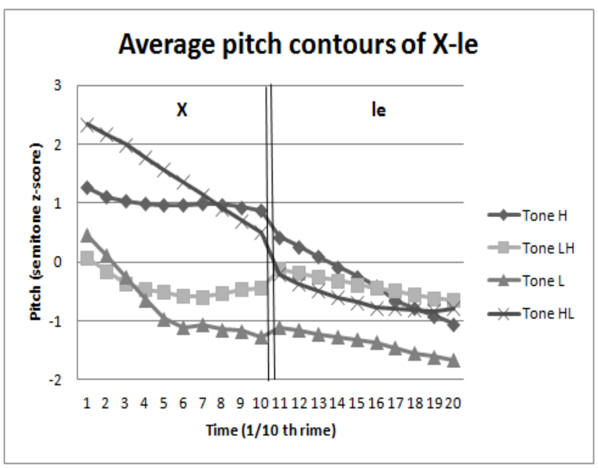


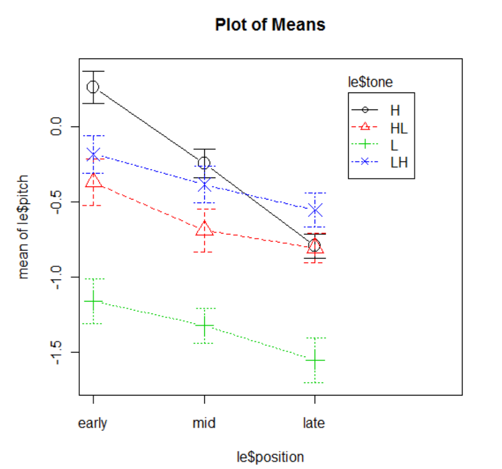
Figure 3.3 Pitch contours of *yi-ge* /HL-Ø/ by subject M2 and M3

3.1.7.1.3. X-*le*

Figure 3.4 shows the result of the full-neutral *X-le* words with four varied lexical tones of *X*. As shown in Figure 3.4a, except for /L-Ø/, the pitch contours reached to -0.5 ~ -1 semitone z-score. In *X-le* /L-Ø/ (*mai-le* 'bought'), the low tone ended at -1.25 semitone z-score and lowered to around -1.65 semitone z-score, which was different from the /L-Ø/ pitch pattern of *X-de* and *X-ge*. Figure 3.4b also shows the mean pitches of the four pitch contours and their standard errors at the early, mid, and late position. The influence of the preceding tones was obvious in the early position. One-way ANOVA shows that the differences of the four pitch contours were statistically significant [$F(3,36) = 20.18$, $p < .001$]. The pitches at the mid and the late position were also influenced by the preceding tones. The differences were also statistically significant [Mid position: $F(3,36) = 16.29$, $p < .001$; Late position: $F(3,36) = 14.75$, $p < .001$]. Post hoc comparison using the Tukey HSD test indicates that the mean pitch of /L-Ø/ was lower than the other three pitches at the .001 level of significance even at the mid and the late position.



a. Average pitch contours of *X-le*

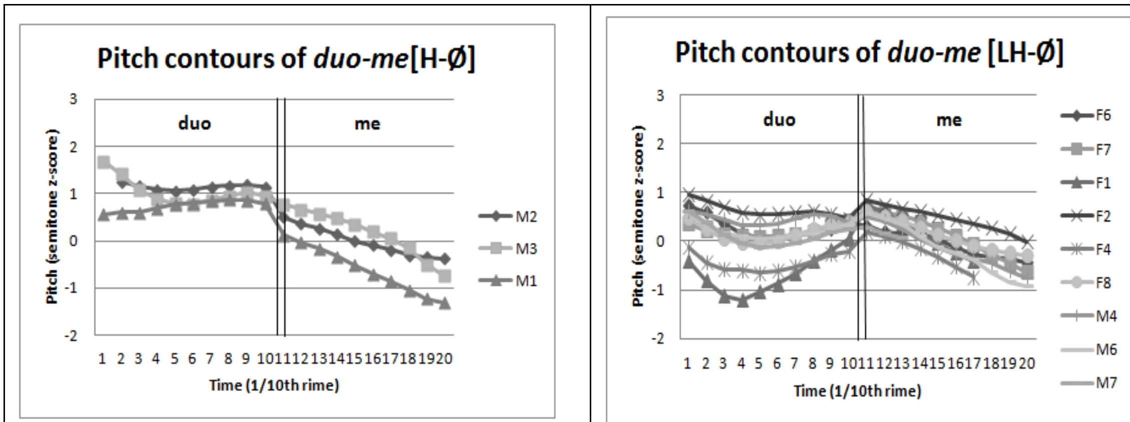


b. Pitch of *le* at early, mid, and late position

Figure 3.4 Result of full-neutral *X-le*

3.1.7.1.4. X-*me*

The tested suffix -*me* is not productive anymore. Only a few vocabulary items use the -*me* suffix. Furthermore, some of these words showed tonal variation on both the preceding syllables and the tested syllables. Figure 3.5 shows the pitch contours of *X-me* /H-Ø/ (*duo-me* ‘such’). Figure 3.5a shows that only three of the subjects produced [H-Ø] as expected. Figure 3.5b shows that nine of the subjects produced the first syllable *duo* with a rising tone /LH/.



a. Pitch contours with [H-Ø] *duo-me*

b. Pitch contours with [LH-Ø] *duo-me*

Figure 3.5 Pitch contours of *X-me* /H-Ø/ (*duo-me*)

There were also tonal variations for the productions of X-*me* /LH-Ø/ (*she-me* ‘what’). The word *she-me* ‘what’ was expected to be [LH-Ø]. However, only five subjects produced the first syllable *she* with a rising tone as shown in Figure 3.6. The pitch contours of *she* [LH] were not very transparent. As shown in the figure, only M7 had a rising pitch contour in *she*; F7 and F8’s *she* had flat pitch contours and F1’s *she* had a falling pitch contour. However, perceptually these *she* syllables all resembled a rising tone /LH/ in TM even though the pitch contours did not necessary rise. It seems that the rising pitch contours were realized in the initial /m/ of the following syllables. The rising pitch was carried over to the next syllable possibly because the rime of *she* is rather short. As a result, the beginning pitches of the following rimes [ə] were all higher than the end pitches of *she*.

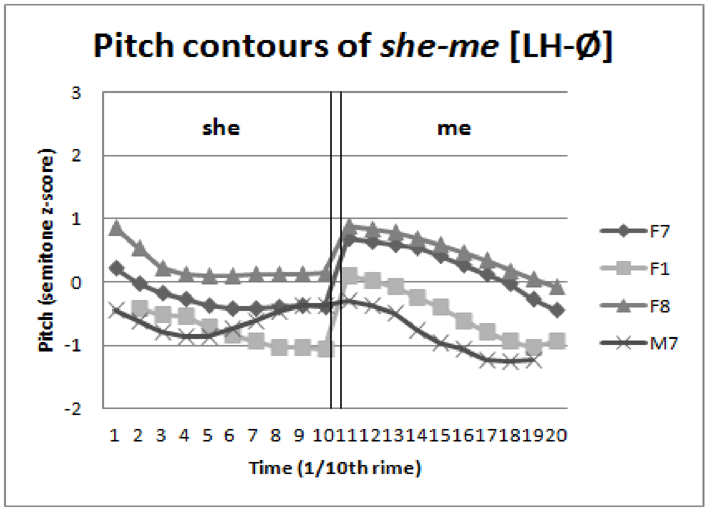
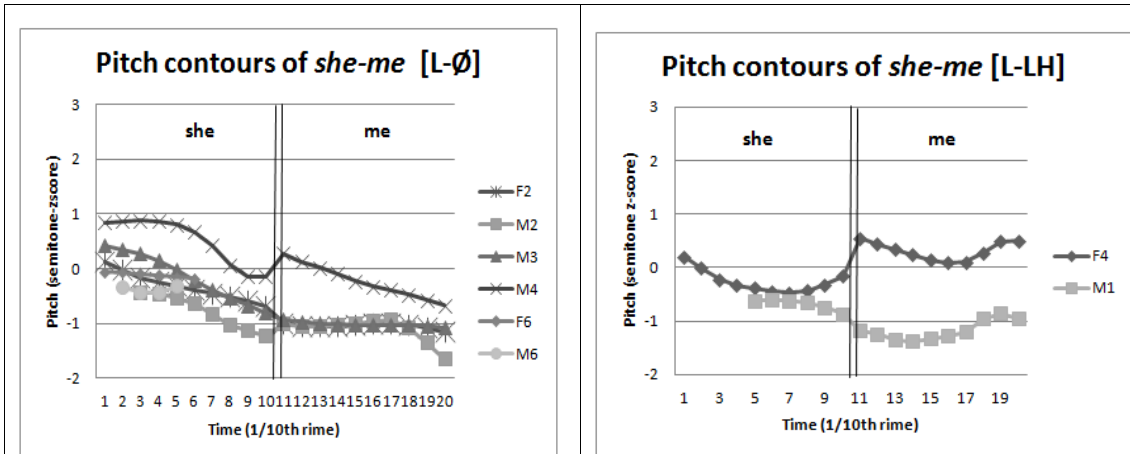


Figure 3.6 Pitch contours of X-*me* /LH-Ø/ (she-me) with a rising-tone first syllable

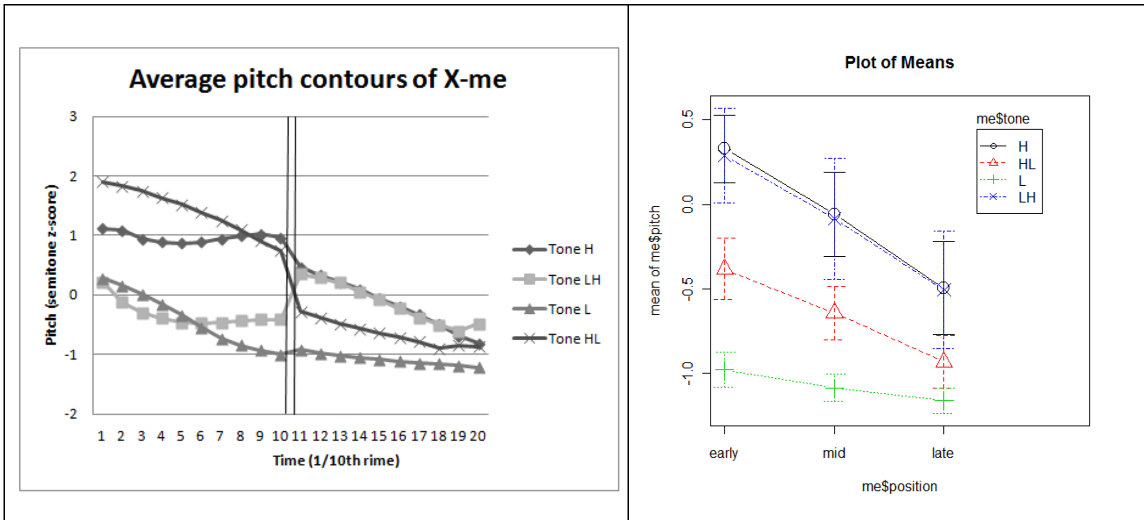
Figure 3.7 shows the other productions of *she-me* /LH-Ø/. Figure 3.7a shows one variation of *she-me* produced by six subjects. These six subjects produced the first syllable *she* with a low tone, and they produced the second syllable *me* with a low pitch which resembled the TM neutral tone. Subject F6 and M6 did not have pitch readings for the second syllables because they were creaky. On the other hand, two subjects produced the first syllable *she* with a low tone, and produced the second syllable *me* with a rising tone as shown in Figure 3.7b.



a. Pitch contours of *she-me* with [L-Ø]	b. Pitch contours of *she-me* with [L-LH]

Figure 3.7 Pitch contours of X-me /LH-Ø/ (*she-me*) with a low-tone first syllable

The tonal variations of *duo-me* /H-Ø/ 'so/such' and *she-me* /LH-Ø/ 'what' are probably because both are high frequency words and high-frequency words are more likely to undergo reductions that eventually change their lexical representations (Bybee 2001, 61-62). The word 'what' was probably produced in various forms because this word is often spoken in various intonations. For example, the word 'what' might have a stronger rising intonation in the end when a speaker wants to express his/her doubt. Therefore it may have different lexical representations with respect to tone, or the lexical representation is neutral with respect to tone. As for *duo-me* 'so/such', the tonal variation might be related to its emphatic meaning. In fact some dictionaries spell the word with /LH-Ø/, but some spell it with /H-Ø/. Ideally these words with tonal variation should not have been included. Unfortunately, only few words in Mandarin contain the fossilized morpheme *-me*.



a. Average pitch contour of *X-me*

b. pitch of *me* at early, mid, and late position

Figure 3.8 Result of full-neutral *X-me*

The result of *X-me* is presented in Figure 3.8. The average pitch contours in Figure 3.8a show a similar pattern to other neutral-tone syllables with /ə/. The pitch contours generally slowly reached to -0.5 ~ -1 semitone z-score. Figure 3.8b also shows the mean pitches and their standard errors at the early, mid, late position of *me*. As we can see, the mean pitches of /H-Ø/ and /LH-Ø/ have a rather high standard error (n = 3, 4 respectively). Therefore statistical analysis was not performed.

Similar to -*le* in Figure 3.4, the pitch contour of *me* after /L/ was lower than the other three. This difference is not observed in the results for -*de* and -*ge*. It is likely related to the fact that both of the initials of the neutral-tone syllables -*me* and -*le* are sonorant while the others are voiceless stops. The pitch of the rime is probably influenced by the consonant perturbation. The voiceless obstruent is known to raise the F0 of the following vowel, and the sonorant is known to lower the F0 of the following vowel (Hombert, Ohala & Ewan 1979). The pitch perturbation of the voiceless consonant raised the pitch of -*de* and -*ge* in /L-Ø/ as shown in Figure 3.1a and Figure 3.2a. On the other hand, the pitch perturbation is not found in -*me* and -*le* in /L-Ø/ as shown in Figure 3.4a and Figure 3.8a. Furthermore, Hombert's data from Yoruba showed that the perturbation caused by a voiceless consonant on a following low tone is greater than the perturbation caused by a voiceless consonant on a mid tone (Hombert 1977). This explains the pitch raising found in -*de* and -*ge* in /L-Ø/, but not in /H-Ø/, /LH-Ø/, or /HL-Ø/.

3.1.7.1.5. X-*zhe*

The result of full-neutral *X-zhe* is shown in Figure 3.9. As shown in Figure 3.9a, the pitch contours overcame the preceding lexical tones and reached to the pitch range of -0.5 ~ -1.25 semitone z-score in the end of the second syllable. At the *zhe* syllable, the /H-Ø/ and /LH-Ø/ pitch contours both fell from 0.5 to -0.5 semitone z-score, and the /L-Ø/ and /HL-Ø/ pitch contours both approximately fell from -0.5 to -1 semitone z-score. The pitch contour of /L-Ø/ drops in the second half of the *zhe* syllable because some syllables have slight creaky voice which decreased the F0 readings.

Figure 3.9b shows the mean pitches of the four pitch contours and their standard errors at the early, mid, and late position. The influence of the preceding lexical tones was statistically significant in all three positions according to the one-way ANOVA tests [Early position: $F(3,50) = 21.19$, $p < .001$; Mid position: $F(3,50) = 13.44$, $p < .001$; Late position: $F(3,50) = 6.64$, $p < .001$]. Post hoc comparison using the Tukey HSD test indicates that at the early and mid position, the pitches of /HL-Ø/ and /L-Ø/ were both lower than the pitches of /LH-Ø/ and /H-Ø/ at the .001 level of significance. At the late position, the pitch of /L-Ø/ was lower than the pitch of /LH-Ø/ ($p < .01$) and /H-Ø/ ($p < .05$), but the pitch of /HL-Ø/ was only lower than the pitch of /LH-Ø/ ($p < .05$).

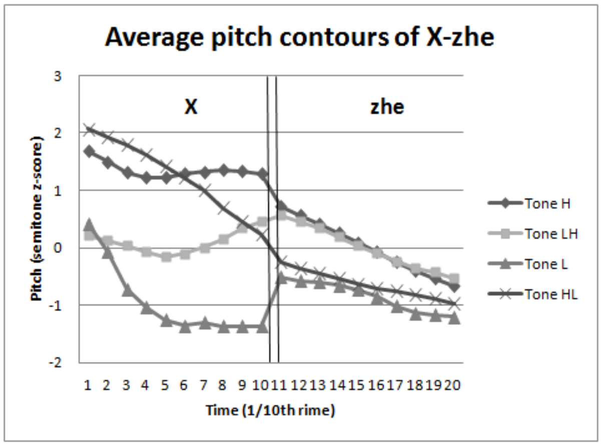
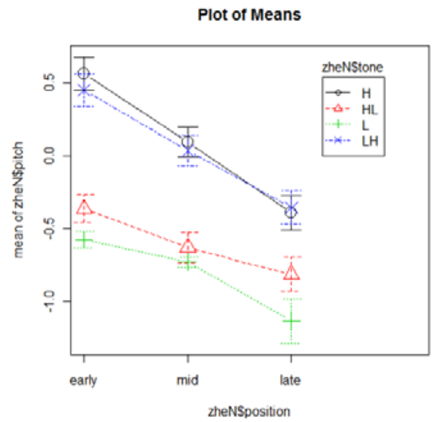
	
a. Average pitch contour of *X-zhe*	b. pitch of *zhe* at early, mid, and late position

Figure 3.9: Result of full-neutral X-*zhe*

3.1.7.1.6. Summary

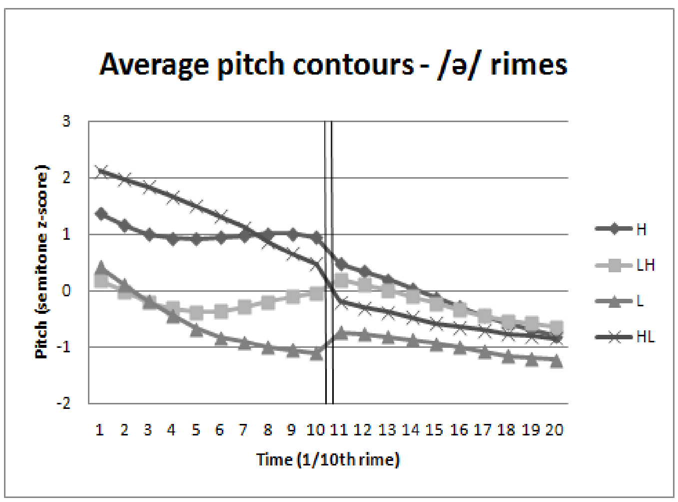


Figure 3.10 Average pitch contours of five neutral-tone syllables *-ge*, *-le*, *-de*, *-zhe*, and *-me*

To summarize, all the neutral-tone syllables with /ə/ rimes generally had a pitch contour that fell from where the preceding tone ended to roughly -0.5 semitone z-score to -1.2 semitone z-score. Figure 3.10 shows the average pitch contours of all the neutral tone syllables with a /ə/ rime discussed above.

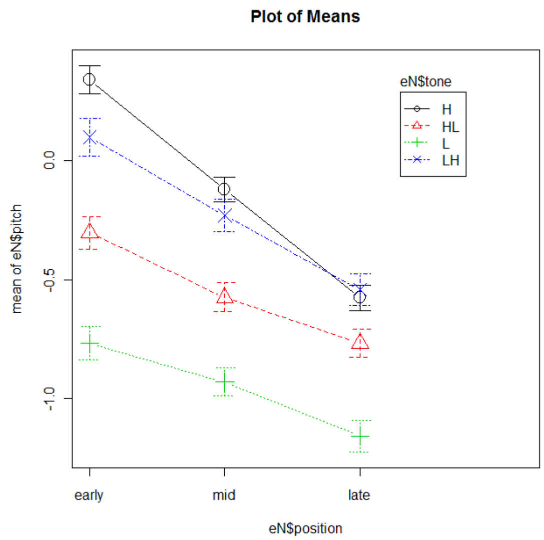


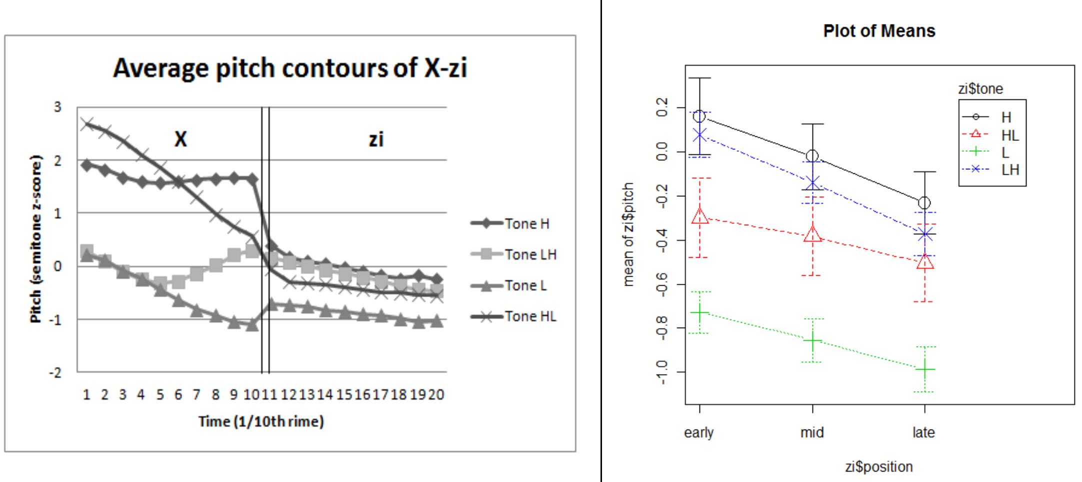
Figure 3.11 Mean pitches and their standard errors of all the neutral-tone /ə/ rimes at early, mid, and late position

The mean pitches and their standard errors of all the neutral-tone syllables with a /ə/ rime at the early, mid, and late position are shown in Figure 3.11. A two-way ANOVA shows that the preceding tones, the tested neutral-tone syllables, and their interactions all influenced the pitches at all three positions. At the early position, there was a significant main effect from the preceding tones [$F(3,169) = 51.67$, $p < .001$], a significant main effect from the tested neutral-tone syllables [$F(4, 169) = 6.17$, $p <.001$], and a significant interaction between the preceding tones and the neutral tone syllables [$F(12, 169) = 3.23$, $p < .001$]. At the mid position, there was a significant main effect from the preceding tones [$F(3, 169) = 35.13$, $p < .001$], a significant main from the tested syllables [$F(4, 169) = 5.07$, $p < .001$], and a significant interaction between the preceding tones and the tested syllables [$F(12, 169) = 2.84$, $p = 0.0014$]. At the late position, there was a marginally significant main effect from the preceding tones [$F(3,169) = 2.95$, $p = 0.0218 < .05$], a significant main effect from the tested syllables [$F(3,169) = 18.56$, $p < .001$], and a significant interaction between the preceding tones and the tested syllables [$F(12,169) = 2.28$, $p = 0.0104 < .05$].

These results show that while the preceding tones had carry-over effects on the following neutral-tone syllables, the pitch of the neutral-tone also varied depending on the tested syllables. While the influence of the preceding tones were overcome at the late position for *-de* and *-ge*, the influence of the preceding tones were still significant at the late position for *-le* and *-zhe*.

3.1.7.2. ***-zi***

The result of the full-neutral *X-zi* is shown in Figure 3.12. The average pitch contours varying the four preceding tones are shown in Figure 3.12a. At the second syllable (neutral-tone *zi*), the four pitch contours ranged between 0 ~ -1 semitone z-score. All four pitch contours at *-zi* were quite flat likely because the rime of *-zi* is very short. Figure 3.12b shows the mean pitches and their standard error at the early, mid, and late position of the *-zi* syllable. The influence of the preceding tones was significant at the early position [$F(3,51) = 9.43$, $p < .001$], the mid position [$F(3,51) = 8.47$, $p < .001$], and even the late position [$F(3,51) = 6.51$, $p < .001$]. The pitches of /L-Ø/ were significantly lower than /H-Ø/ and /LH-Ø/ at all three positions.



a. Average pitch contour of *X-zi*	b. Pitch of *zi* at early, mid, and late position

Figure 3.12 Result of full-neutral *X-zi*

3.1.7.3. *-men*

The result of the neutral-tone *-men* shows speaker and lexical variations. Out of 4 (words with *-men* suffix)*12 (speakers) = 48 production of *-men*, four instances of *-men* had the pitch contours that resemble the rising tone /LH/ as shown in Figure 3.13. These four *-men* instances had a rising pitch contour most likely because the suffix *-men* /Ø/們 is written with a phonetic element of 門, which has a rising tone /LH/. This variation is however not speaker or lexical item specific. M6 produced two instances of rising *-men* but his /HL-Ø/ and /H-Ø/ did not have a rising *-men*.

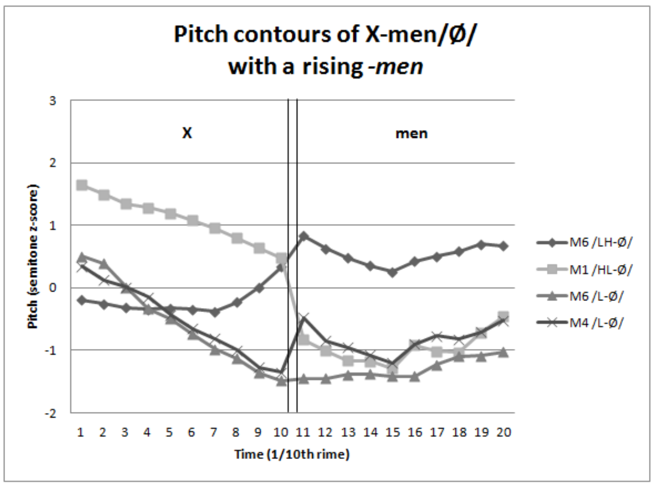
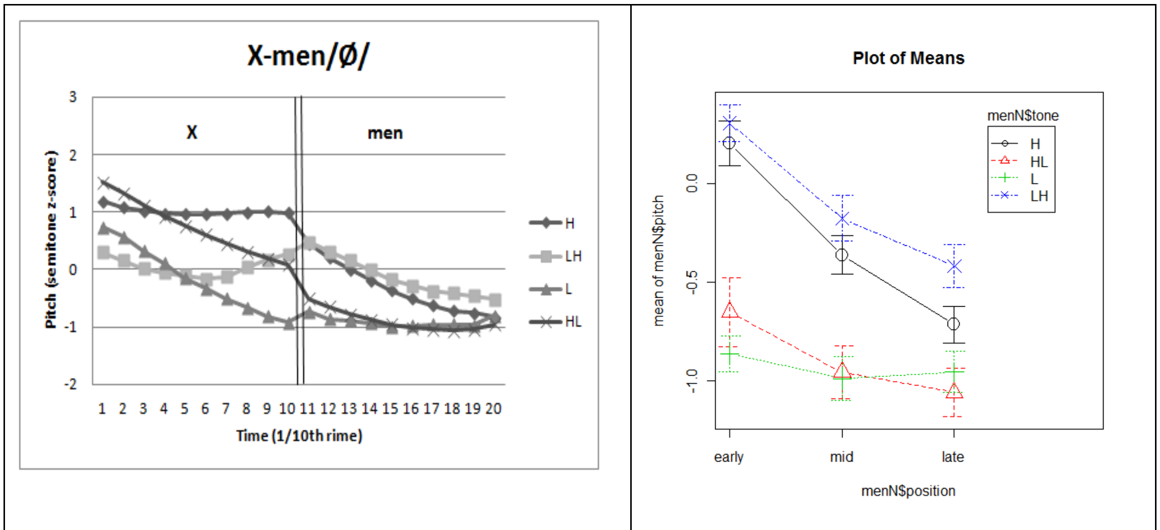


Figure 3.13 Pitch contours of the *-men* tokens that have a rising pitch /LH/

The rest of the *-men* instances had slight falling pitch contours as did the other neutral-tone syllables mentioned above. The average pitch contours are shown in Figure 3.14a. As shown in the figure, the pitch of the second syllable *-men* tried to overcome the previous tones and reached to around -0.5 ~ -1 semitone z-score. However, as shown in Figure 3.14b, the influence of the preceding lexical tone was still significant at the late position [$F(3,33) = 7.11$, $p < .001$]. The pitch of /LH-Ø/ at the late position of *-men* (M = -0.42, SD = 0.36) was significantly higher than the pitch of /HL-Ø/ (M = -1.06, SD = 0.36).



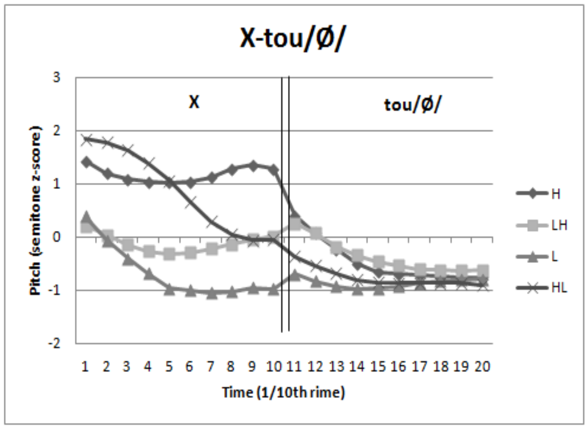
a. Average pitch contours of *-men* that have mid-low pitch target

b. Mean pitches and standard errors at the early, mid, and late position of *-men*

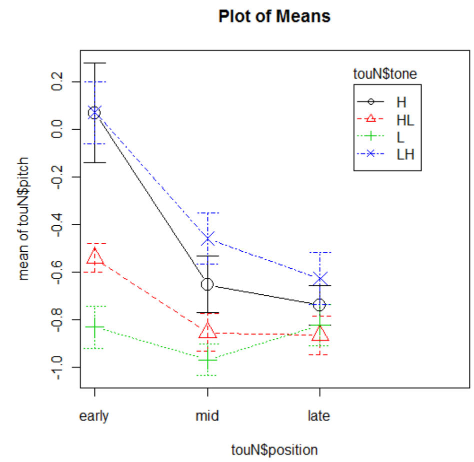
Figure 3.14 Pitches of *-men* that have mid-low pitch target

3.1.7.4. ***-tou***

The neutral-tone -*tou* was different from the seven neutral-tone syllables discussed above. As shown in Figure 3.15a, the pitch of *tou* /Ø/ overcame the previous pitch quickly in the first half of the -*tou* syllable and remained flat around -0.7 semitone z-score. Unlike the neutral-tone syllables discussed above, the pitch falls between the early and the mid position were much faster (steeper) than the pitch falls between the mid and the late position as illustrated in Figure 3.15b. One-way ANOVA also shows that while the preceding tones influence the pitch significantly at the early and mid position [Early position: $F(3,37) = 14.23$, $p < .001$; Mid position: $F(3,37) = 6.28$, $p = 0.0015 < .01$], the influence of the preceding tones was not significant at the late position [$F(3,37) = 1.38$, $p = 0.26$]. These pitch contours show that -*tou* 'diminutive suffix' was pronounced with the original lexical tone /LH/ 'head' as if it were a free morpheme. The TM speakers did not treat the syllable as unstressed (or having a neutral tone), so they simply produced the syllable with their knowledge of the character, which is /LH/. Also, these neutral-tone -*tou* syllables perceptually resemble /LH/ even though the pitch contours did not have an obvious pitch rise as have been noticed by Fon and Chiang (1999).



a. Average pitch contours of *X-tou* /Ø/



b. Pitch of *tou* /Ø/ at early, mid, and late position

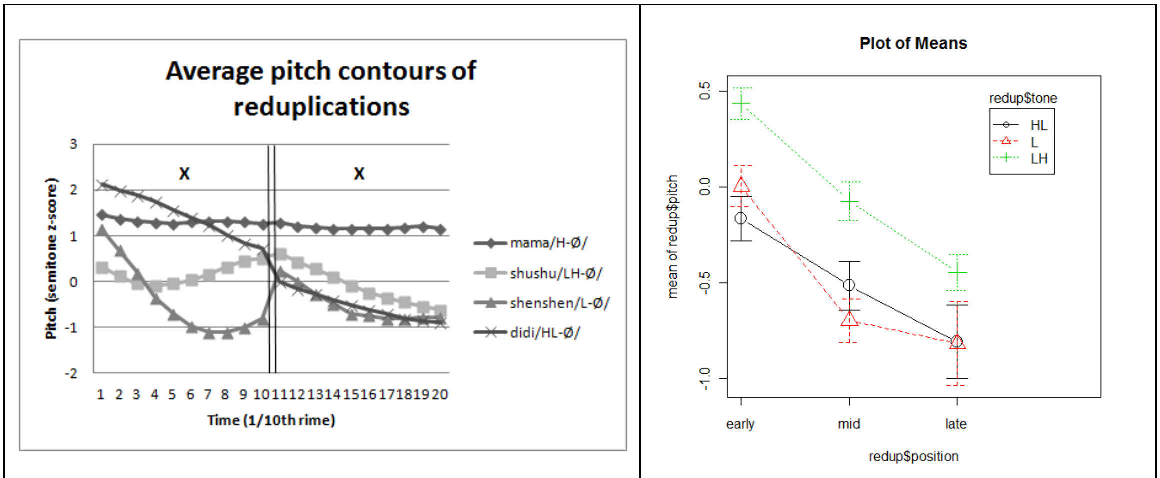
Figure 3.15 Result of full-neutral *X-tou* /Ø/

3.1.7.5. **Reduplication**

In Standard Mandarin, reduplicated syllables are always unstressed, resulting in a neutral tone. In this production test I also included words that were reduplications. However, the reduplications in Taiwan Mandarin are limited to mostly kinship terms. Therefore all four tested words are kinship terms: *ma-ma* /H-Ø/ ‘mother’, *shu-shu* /LH-Ø/ ‘uncle’, *shen-shen* /L-Ø/ ‘aunt’, and *di-di* /HL-Ø/ ‘younger brother’.

Figure 3.16a shows the four average pitch contours of the reduplications. However, some productions were not included because they were obviously different from the others. These different pitch contours will be discussed separately. As shown in Figure 3.16a, the high tone reduplication *ma-ma* /H-Ø/ ‘mom’ had a high pitch reduplicant. The reduplicated high tone did not possess a neutral tone; instead it possessed a high tone. On the other hand, the reduplicants of /LH/, /L/, and /HL/ have similar pitch contours that acted like other neutral-tone syllables which reached to -0.5 ~ -1 semitone z-score. The three pitch contours reached to a similar pitch target at the second syllable. As illustrated in Figure 3.16b, one-way ANOVA shows that at the early and mid positions of the reduplicant, the pitches of the three reduplicants (LH-Ø, L-Ø, HL-Ø) were significantly different [early position: $F(2, 28) = 9.58$, $p < .001$; mid position: $F(2,28) = 7.63$, $p =$

0.0023 < .01], Post hoc comparison using the Tukey HSD test indicates that the /LH/ reduplicant was statistically higher than the other two at the early position [M = 0.44, SD = 0.27] and the mid position [M = -0.07, SD = 0.34] . However, at the late position, the three reduplicants were not statistically different [$F(2,28) = 1.57$, $p = 0.23$]. The influence of the preceding tones was overcome at the late position.



a. Average pitch contours of reduplications in TM

b. Pitch of reduplications of /LH, L, HL/ at early, mid, and late position

Figure 3.16 Result of reduplications

Three pitch contours of the reduplication *shen-shen* /L-Ø/ 'aunt' were excluded. They were produced by M1, M6, and F8. As shown in Figure 3.17, M1 and M6 produced the reduplicated *shen* with a high pitch and F8's *shen-shen* had the pitch contour of [L-LH]. M1 and M6's *shen-shen* /L-Ø/ were similar to Standard Mandarin in which the neutral tone has a post-L rising. It is very likely that this pitch pattern of low-high was lexicalized with the word *shen-shen* 'aunt' because M1 and M6 did not produce post-L rising in any of the neutral tones mentioned above.

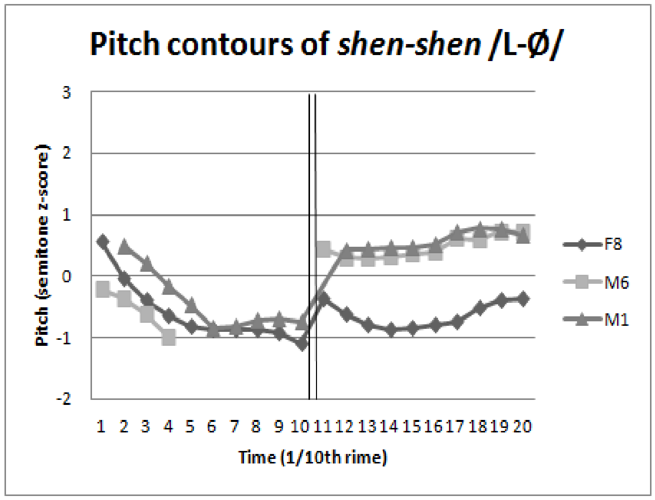


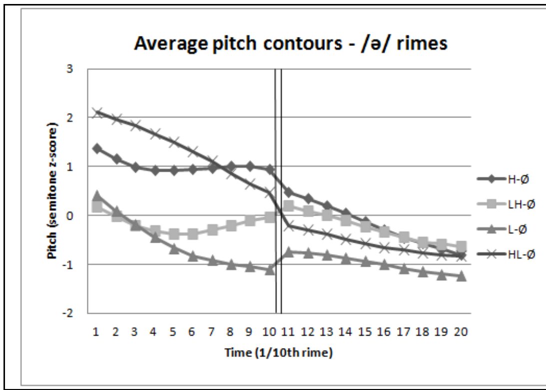
Figure 3.17 Pitch contours of the /L/ reduplication by M1, M6, and F8

F8's *shen-shen* produced *shen-shen* /L-Ø/ with the pitch contour of [L-LH], which is a common way to show endearment. In Taiwan Mandarin, a common way to form nicknames or to show endearment is to put the pitch pattern [L-LH] on reduplicated words (H.-c. Hsu 2006). For example, most of the kinship terms can take the [L-LH] tonal pattern such as *ma-ma* 'mother', *ba-ba* 'father', *ye-ye* 'grandfather', *shu-shu* 'uncle'. Similar patterns can also be found in child speech in referring to animals. For example *gou* /L/ 'dog' becomes *gou-gou* [L-LH], and *xiong* /LH/ 'bear' becomes *xiong-xiong* [L-LH].

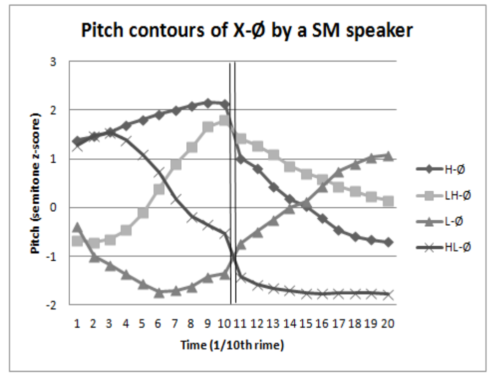
3.1.8. Discussion

Section 3.1.7 showed the results of eight neutral-tone syllables (suffixes) and reduplication in the environment of four preceding lexical tones. Except for the suffix -*tou*, which shows /LH/ tone-like pitch contours, and the reduplicant of /H/, which has a high tone, generally the pitch contours of the neutral-tone syllables went from the end pitch of the preceding syllable and reached to a mid-low range between -0.5 and -1 semitone z-score. In -*de*, -*ge*, and three reduplications, the influences of the preceding lexical tones were even overcome at the late position. The pitch target was implemented in the second half of the syllable, which is a lexical tone feature.

This pitch behavior is very different from what has been documented in Standard Mandarin. Figure 3.18 compares the full-neutral pitch patterns between Taiwan Mandarin and Standard Mandarin. Figure 3.18a is the same as Figure 3.10 and Figure 3.18b shows the reduplications of four lexical tones produced by a SM speaker. The pitch patterns of Figure 3.18b are very similar to what Chao (1968) has described as the neutral tone pitch patterns, where /H-Ø/ is 5-2, /LH-Ø/ is 35-3, /L-Ø/ is 21-4, and /HL-Ø/ is 51-1.



a. Average pitch contours of X-Ø with /ə/ rimes by 12 TM subjects



b. Pitch contours of X-Ø by a SM speaker

Figure 3.18 Comparison of full-neutral pitch contours between TM and SM

Comparison between the neutral tones of TM and SM shows that the four preceding lexical tones are very different between TM and SM as mentioned by Fon and Chiang (1999). Also, the pitch contours of the neutral tone between the two dialects are also very different. The /H-Ø/ contours in TM stays at 1 semitone z-score in /H/ and lowers to -0.75 semitone z-score in /Ø/, but the /H-Ø/ contours in SM reaches to 2 semitone z-score in /H/ and falls to -0.75 semitone z-score in /Ø/. The end pitches of the neutral tone in /H-Ø/ between the two dialects are quite similar. As for /LH-Ø/, in TM it slowly rises to 0 semitone z-score in /LH/ and falls to -0.5 semitone z-score, but in SM it rises to 2 semitone z-score in /LH/ and falls to 0 semitone z-score in /Ø/. The end pitches of /Ø/ in /LH-Ø/ are quite different; this is likely due to the fact that /LH/ in TM does not rise as high as SM. The most apparent difference between TM and SM seems to be in /L-Ø/. In TM the pitch of /L/ lowers to -1 semitone z-score and the pitch of /Ø/ stays low in the same pitch range. In SM, the pitch shows post-L rising; the pitch contour of /L/ lowers to -2 ~ -1.5 semitone z-score and that of / Ø/ rises to 1 semitone z-score. As for /HL-Ø/, in

TM the pitch of /HL/ falls to about 0 semitone z-score and the pitch of /Ø/ lowers to almost -1 semitone z-score. However in SM the pitch of /HL/ falls to almost -1 semitone z-score and the pitch of /Ø/ stays low to about -1.75 semitone z-score. Similar to /LH-Ø/, the differences of /Ø/ in /HL-Ø/ between TM and SM are probably due to the fact that in TM the /HL/ contour does not fall as low as in SM.

Most importantly, the pitch of the neutral tone in SM depends on the preceding lexical tone, which is a sign of tone loss. On the contrary, the pitch of the neutral tone in TM has a mid-low pitch target. The pitch target suggests that the neutral-tone syllables in TM are not toneless, but have an underlying tonal target.

There were also some variations observed in the data of TM. The suffix -*me* has a lot of pitch variation due to the possible high frequency of the tested disyllabic words. As noted in 3.1.7.1.2, two instances of *yi-ge* [LH-Ø] were produced as [HL-L]. In 3.1.7.3, four instances of -*men* /Ø/ were produced with a rising pitch contour. In 3.1.7.5, three instances of reduplication /L-Ø/ were produced with different pitch contours. These sporadic variations may be lexicalized pitch patterns. However, the production of *yi-ge* [HL-L] suggests the -*ge* suffix might be treated as a low tone phonemically by some TM speakers.

3.2. Neutral-tone syllables in novel occurrences

Many of the words that contain a "prescriptive" neutral tone are high-frequency words such as *wo-de* 1.S-POSS 'my', *mai-le* buy-PERF 'have bought', and *wo-men* 1-PL 'our'. Therefore the pitch of these neutral-tone syllables might surface like a neutral tone of SM because of the high frequency of certain phrases or words. For example, as we can see in the previous experiment, two instances of the neutral-tone *shen* in *shen-shen* 'aunt' had a post-L rising which is a SM neutral tone realization. However, it is very likely that the high pitch of *shen* in TM is because this word-form *shen-shen* is lexicalized as /L-H/ due to high frequency input from SM speakers. I propose that the TM speakers do not really acquire the process of tone neutralization of SM. Therefore, if I test these prescriptively neutral-tone syllables in novel words, I may show that the process of tone neutralization that applies in SM does not apply in TM speech.

3.2.1. Method

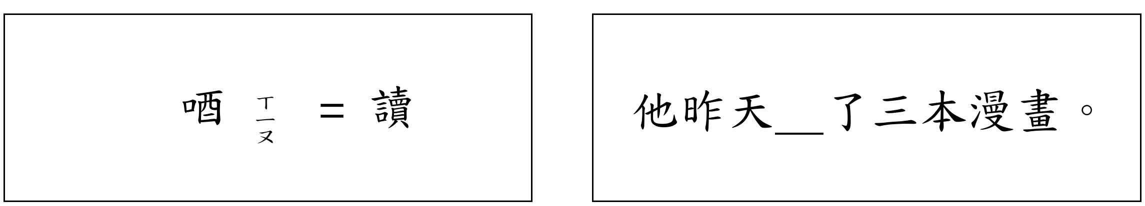
Four of the tested neutral-tone suffixes are still productive. They are perfective *le*, durative *zhe*, possessive *de*, and plural *men*. In this experiment, subjects were given an obsolete character with a made-up spelling and a made-up meaning, and they were asked to read a meaningful sentence with a tested suffix attached to this made-up character. The obsolete characters attached by the first two suffixes were made-up as verbs, and the obsolete characters attached by the last two suffixes were made-up as nouns. Each suffix was combined with four different verbs (or nouns) in four different sentences respectively. Therefore the subjects were asked to read 4 (tested suffixes) * 4 (preceding nouns or verbs in 4 lexical tones) = 16 sentences. Table 3.3 shows the four sentences used to elicit -*le* in novel occurrences.

As illustrated in Table 3.3, for each suffix, the four verbs (or nouns), which were written in four different obsolete characters, have the same segments but different lexical tones in order to control for the consonant perturbation. Also, for each suffix, the four sentences have similar sentence structures with minor meaning differences, so that the intonation will be the same for all four tested parts. The full list of the 16 sentences and their testing elements are shown in Appendix C.

Table 3.3 Materials of creating *-le* in novel occurrences

suffix	obsolete characters	made-up sounds and meanings	given sentences
-le (-PERF)	呦	*xiu* /H/ 'to eat'	*ta zuotian* __***le*** *sanwan chaofan* 'He ____three bowls of fried rice yesterday'
	唀	*xiu* /LH/ 'to drink'	*ta zuotian* __***le*** *sanbei hongcha* 'He ____three cups of black tea yesterday'
	哂	*xiu* /L/ 'to read'	*ta zuotian* __***le*** *sanben manhua* 'He ____three comic books yesterday'
	啁	*xiu* /HL/ 'to buy'	*ta zuotian* __***le*** *sanzhi yashua* 'He ____three toothbrushes yesterday'

Subjects were the same as in 3.1.3. This experiment is one of the tasks they were asked to complete. The equipment and data analysis were the same as in 3.1.4 and 3.1.6. Subjects were given a made-up character and they were asked to read a meaningful sentence with the tested morpheme attached to this made-up character. For example, as shown in Figure 3.19, 哂 (the obsolete character) was written on the front of the slip of paper. There were also *zhuyin fuhao* phonetic symbols showing its pronunciation (*xiu* /ɕjou/ /H/) and the character 讀 *du* 'to read' written next to it showing the meaning of the character. The back of the paper showed a sentence 他昨天__了三本漫畫 *ta zuotian* ____ *le sanben manhua* 'He ___ three comic books yesterday.' The subjects were first shown the front of the paper, and they were asked to put 哂 in the space and read the whole sentence on the back after they flipped to the back of the slip.



a. Front of the paper b. Back of the paper

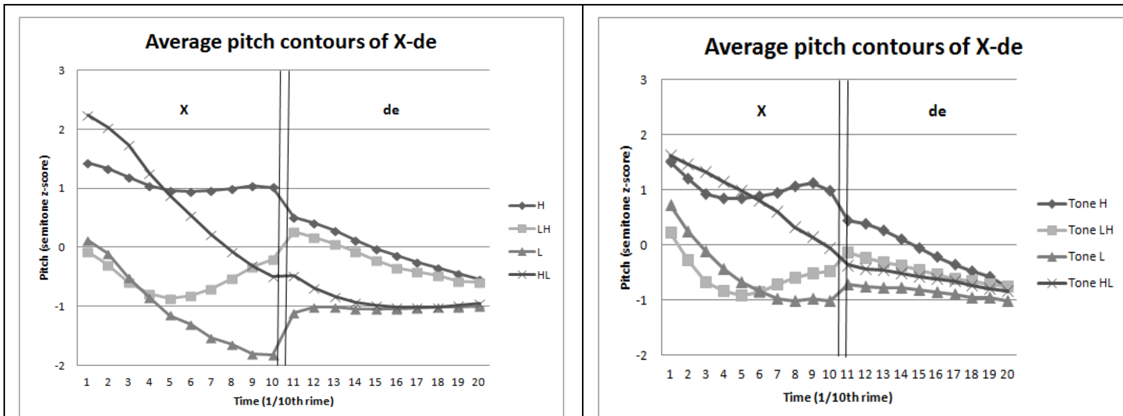
Figure 3.19 The sample of production test on novel occurrence of the neutral tone

There were 16 obsolete characters and 16 matching sentences (See Appendix C). They were written on the both sides of 16 slips of paper. The subjects were given 16 pieces of paper one by one, with the front of the paper on top. The order of the papers was randomized. The slips of paper were shuffled differently for each speaker.

3.2.2. Results

The subjects were asked to produce neutral-tone syllables in novel formations. The results of the four suffixes from the 12 TM speakers are shown at the a. sections from Figure 3.20 to Figure 3.23. These charts show the average pitch contours of the four productive suffixes following four different lexical tones. The rime of the first syllable (varied made-up syllable) is shown on 1-10 on the X-axis, and the rime of the suffix is on 11-20. They are separated by a double line.

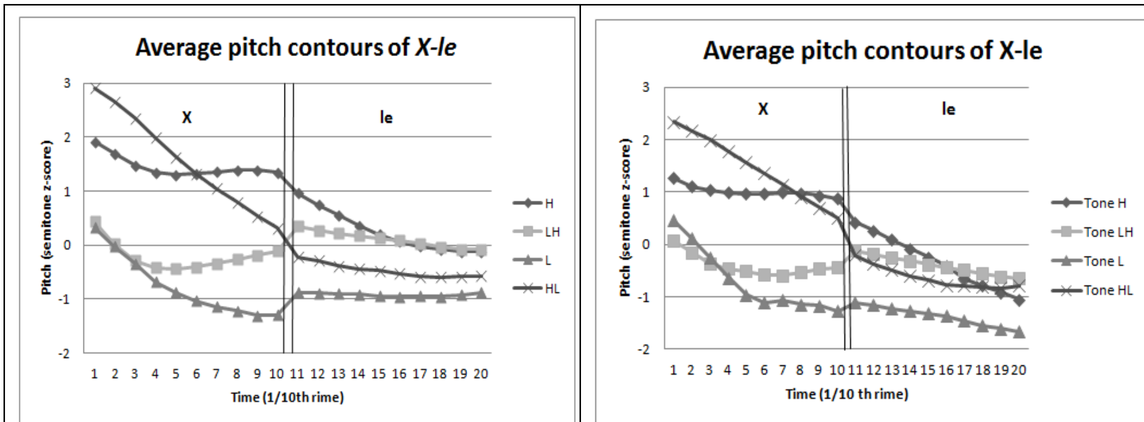
The result shows that generally the pitch of the neutral-tone suffixes gradually moved to the mid-low level (0~-1 semitone z-score). The four suffixes are discussed individually as follows.



a. *X-de* in novel occurrences | b. *X-de* in disyllabic words

Figure 3.20 The results of *X-de* in novel occurrences and in disyllabic words

Figure 3.20a shows the result of *X-de* in novel occurrences, and Figure 3.20b shows the result of *X-de* in disyllabic words as discussed in Section 3.1.7.1. Both of the charts show that the pitch contours of -*de* reach toward -0.5 ~ -1 semitone z-score. However, the pitch ranges of the first syllable X are larger in the novel occurrences. This is probably because in novel occurrences the subjects talked slower and emphasized the test syllables more. The result of *X-de* in the novel occurrences also shows the influence of the preceding lexical tones. At the late position of -*de*, the influence of the preceding tones on the pitch was significant [$F(3,41) = 6.37$, $p = 0.0012 < .01$].

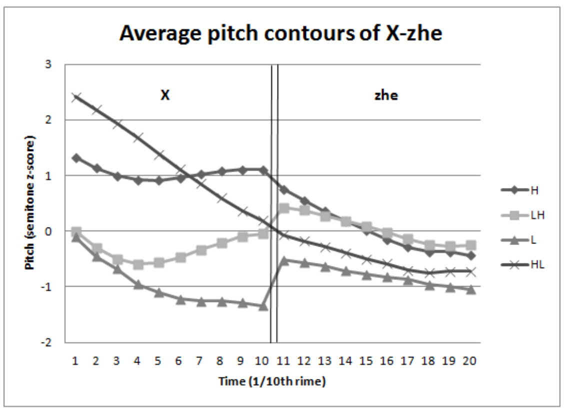


a. *X-le* in novel occurrences	b. *X-le* in disyllabic words

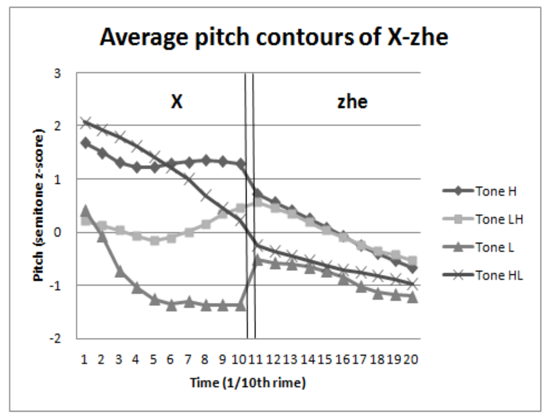
Figure 3.21 The results of X-*le* in novel occurrences and in disyllabic words

Figure 3.21a shows the result of *X-le* in novel occurrences. We can compare it to the result of *X-le* in disyllabic words as shown in Figure 3.21b. As we can see, similar to the result of *X-de*, the pitch of X shows a larger range in novel occurrences. The pitches of -*le* in novel occurrences reached to the range of 0 ~ -1 semitone z-score. The pitches of the neutral-tone -*le* in novel occurrences were influenced by the preceding lexical tones. Even though the pitch contours slowly merged, the influence of the preceding tones was still significant at the late position [$F(3,40) = 12.985$, $p < .001$].

Comparing the two graphs in Figure 3.21, the pitch contours of /H/ and /LH/ in X ended higher in the novel occurrence than in the disyllabic words. The pitches of *le* following /H/ and /LH/ in novel occurrences ended high at 0 semitone z-score. *le* after /L/ in a disyllabic word (shown in b) has a pitch contour lowering to -1.75 while *le* after /L/ in novel occurrences (shown in a) stayed around -1 semitone z-score. Overall, the pitch range at the end of -*le* in novel occurrences is about 0.5 semitone z-score higher than in disyllabic words. This difference is possibly due to the experiment's design. Even though both of the test words were before a high-tone syllable, *X-le* in novel occurrences was in the meaningful sentence *ta zuotian* ***xiu-le*** *sanwan chaofan* ‘He **ate** three bowls of fried rice yesterday’ before an object ‘three bowls of fried rice’. On the other hand, *X-le* in disyllabic words was in a frame sentence *qingshuo* ***chi-le*** *baci* ‘Please say **ate** eight times’. Speakers were much more likely to have a prosodic break between the disyllabic word and the frame “eight times” than between a transitive verb ‘to eat’ and its object. This difference caused the *le* in disyllabic words tend to have slight creakiness which slightly lowered the pitch readings.



a. *X-zhe* in novel occurrences

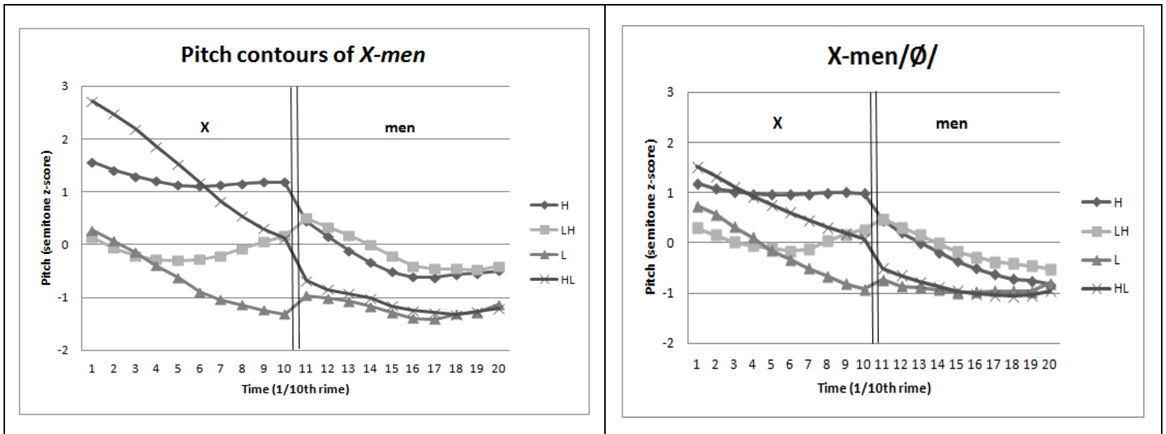


b. *X-zhe* in disyllabic words

Figure 3.22 The results of X-*zhe* in novel occurrences and in disyllabic words

The result of *X-zhe* in novel occurrences is shown in Figure 3.22a. The result is similar to the result of *X-zhe* in disyllabic words as shown in Figure 3.22b. Both figures show that the pitch contours of *zhe* all reached toward a mid-low range. The pitch contours of *zhe* in novel occurrences ended between -0.2~ -1 semitone z-score, but the pitch of *zhe* in disyllabic words ended between -0.5 ~ -1.2 semitone z-score. This difference is probably due to the experiment design as mentioned in the *X-le* section: the disyllabic words were elicited in frame sentences and the novel occurrences were elicited in meaningful sentences. A one-way ANOVA also shows that the pitch of the *zhe* in novel occurrences is influenced by the preceding lexical tone even at the late position [$F(3,36) = 4.20$, $p = 0.012 < .05$].

The result of *X-men* in novel occurrences is shown in Figure 3.23a. The pitch contours of *-men* generally reached to a mid-low range, but they were also influenced by the preceding lexical tone: the pitch contours of *-men* after /LH/ and /H/ were higher than those after /HL/ and /L/. The differences were significant even at the late position of *-men* [$F(3,39) = 12.40$, $p < .001$].



a. *X-men* in novel occurrences	b. *X-men* in words

Figure 3.23 The results of X-men in novel occurrences and in disyllabic words

Figure 3.23 also shows that the pitch of *-men* in novel occurrences slightly raised in the last one-third of the rime, but based on native speaker's perception, they do not have a rising tone. Similar to X-*le* and X-*zhe*, this pitch rise was likely due to the experiment design. X-*men* in this experiment was elicited in meaningful sentences where the anticipation effect of the following *zai* /HL/ was probably stronger than when X-*men* was put in a frame sentence. The comparison between Figure 3.23a. and b. also shows that the pitch range is larger in novel formations than in real words.

3.2.3. Discussion

Overall, the results of the four neutral-tone syllables in novel occurrences are similar to what I have found in Section 3.1. The pitch contours of the neutral-tone syllables slowly overcame the preceding lexical tones and reached to a mid-low target. The pitch patterns of the neutral tone in Taiwan Mandarin are very different from the neutral tone in Standard Mandarin. The results in novel formations also suggest that the post-L risings that resemble Standard Mandarin (as found in Section 3.1) are lexicalized pitch patterns. They only occur sporadically in certain lexical items and produced by certain speakers.

The results in novel formations also show different aspects of the neutral-tone syllables in Taiwan Mandarin. Due to the design of the experiment, the full-neutral utterances were elicited in meaningful sentences. However, the syllables after the neutral-tone in this experiment were not controlled. They varied in terms of their lexical tones, initial consonant, and their grammatical relations with the preceding tested nouns/verbs. As a result, the end pitch of the neutral tone seems to show some variation compared to the neutral tone elicited in a frame sentence. Although the influences of the following syllables on pitch are not as important as the preceding syllables, the influence of the following syllable on pitch has been reported in terms of following tone (Lee & Zee 2008) and following consonant perturbation (Y. Xu 2001). It would be interesting to run the same experiment with various following lexical tones to see how the anticipatory effect applies.

In this experiment, the influences of the preceding lexical tones on the neutral tone syllables seemed to be stronger than the results in 3.1. Statistical analysis shows that the influence of the preceding tones was significant at the late positions of all four tested neutral-tone syllables. This result might be explained by the experiment design. In this experiment, the sentences that were elicited focused on the syllables before the tested neutral-tone syllables. Because the subjects had to insert the made-up syllables into the sentences, the made-up syllables with various lexical tones were pronounced with emphasis. As a result, the pitch ranges of these preceding lexical tones (made-up syllables) were larger. In Mandarin, lexical tone is known to have a carry-over effect on the following lexical tone (Shen 1990; M. Lin & Yan 1991; Shih & Sproat 1992; Y. Xu 1994; Y. Xu 1997). The syllables with emphasis might create a stronger carry-over effect on the following syllables.

In addition, several pitch contours produced by one subject were not included in the average pitch contours above. However, these pitch contours show us the phonological realization of the neutral tone of this subject. Subject M3 consistently produced the preceding low tone (X) with a rising tone in the novel occurrences task. As shown in Figure 3.24, all four /L-Ø/ novel formations produced by M3 have a rising tone on the first /L/ syllable.

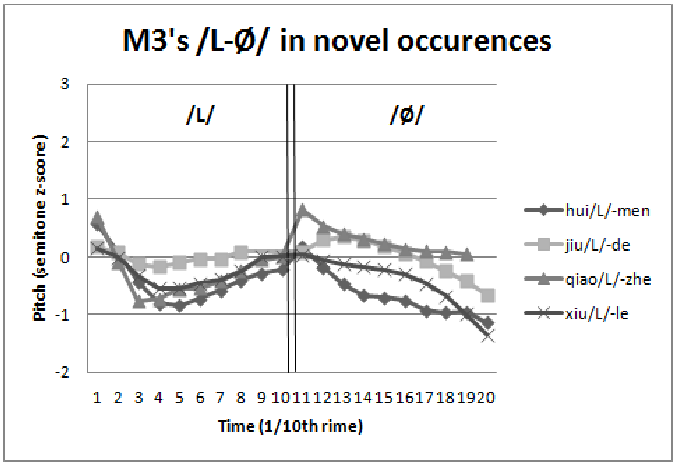


Figure 3.24 Subject M3's /L-Ø/ in novel occurrences

As demonstrated in Figure 3.25, the pitch contour of *hui-men* /L-Ø/ is almost identical to the pitch contour of *hui-men* /LH-Ø/.

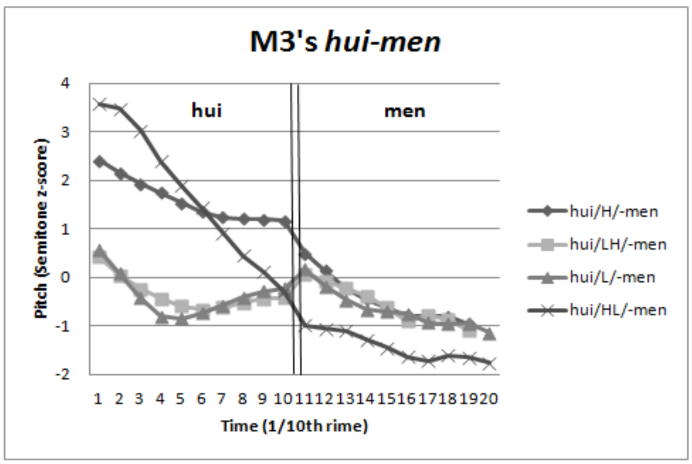


Figure 3.25 X-*men* /L-Ø/ in novel occurrences produced by Subject M3

The production /L-Ø/ in novel occurrences by M3 shows that M3 treated all the suffixes as a low tone phonologically. As a result, the /L-Ø/ combination was realized as /L-L/ by M3, and he applied Tone 3 sandhi to /L-L/ resulting in the production of [LH-L].

xiu-le	/L-Ø/	Prescriptive tone
xiu-le	/L-L/	M3's representation
xiu-le	[LH-L]	M3 applied Tone 3 Sandhi: L-L → LH-L

The fact that M3 treated the neutral-tone syllables like a low tone phonologically shows that for some speakers of Taiwan Mandarin the low pitch of the neutral tone has led them to categorize these neutral-tone syllables as a low-tone syllable. The [L-L] pitch patterns I observed in the disyllabic words are in fact lexicalized for these TM speakers. In a completely novel situation, these neutral-tone syllables were treated like a low tone.

3.3. Consecutive neutral-tone syllables

According to Chen and Xu (2006), the pitch of consecutive neutral-tone syllables in Standard Mandarin moves from the previous full tone toward a mid target gradually. Since the neutral tone syllables in Taiwan Mandarin appear to have a target, I hypothesize that the consecutive neutral-tone syllables in TM will behave like consecutive full tones. Instead of gradually reaching to a mid target as in Standard Mandarin, consecutive neutral-tone syllables in TM will each have their own target.

3.3.1. Method

Several prescriptive neutral-tone syllables (reduplication, nominalizer/ possessive *de*, plural marker *men*, durative marker *zhe*, and perfective marker *le*) were tested. These tested neutral-tone syllables were combined into four sets of consecutive neutral tone syllables: 1) -*X-de* '-reduplication-POSS', 2) -*X-men-de* '-reduplication-PL-POSS', 3) -*zhe-de* '-DUR-NML.', and 4) -*le-de* '-PERF-NML.'). These four sets had either two or three meaningful consecutive neutral-tone syllables. Each set of consecutive syllables was put into a meaningful sentence. Adopting Chen and Xu's (2006) methodology, in order to see the influences of the preceding lexical tone, I varied the lexical tone before the consecutive neutral tones. All four lexical tones were included. Also, in order to control for contextual influence, I varied the lexical tone following the consecutive neutral tones. Only /L/ and /HL/ were included because /L/ starts at the lowest pitch and /HL/ starts at the highest pitch. An example is given as followed.

		X	$N_{2\text{-}3}$			Y		
	H	H	N	N	N	L	H	
ta	*shuo*	*ma*	*ma*	*men*	*de*	*hao*	*duo*	*le*
3.S.	say	mom	-REDUP	-PL	-POSS	good	more	ASP

'He said moms' are better'.'

As illustrated in the example, X (varies in 4 lexical tones) represents the syllables before the consecutive neutral-tone syllables, N represents the neutral-tone syllables (two or three), and Y (varies between 2 lexical tones) represents the syllable after the consecutive neutral-tone syllables. As a result, subjects read 4 (lexical tone of X)* 4 (sets of consecutive neutral-tone syllables)* 2 (Y) =32 sentences.

The subjects were the same as in 3.1.3. This experiment is one of the tasks they were asked to finish. The equipment and data analysis were the same as in 3.1.4 and 3.1.6. The subjects were asked to read the 32 sentences written in Chinese characters. The list of the 32 sentences is shown in Appendix D. There was no practice section but the subjects were encouraged to re-read the sentence if they felt they had made a mistake.

3.3.2. Results

The results are presented in the following section 3.3.2.1 to 3.3.2.3. The pitch contours are shown from Figure 3.26 to Figure 3.29. In these figures, rimes of the tested syllables were measured and graphed. There are double lines between the syllables to separate the syllables and to show that the syllable-initial consonants were not included in the graph. The first and the last syllables were the variables X and Y. The first syllables X can be one of the four lexical tones in Mandarin (H, LH, L, HL), and the last syllable Y is either /HL/ or /L/. Each chart shows the average pitch of the 12 TM speakers. However, if a pitch contour had no pitch readings at the mid position (the 5^{th} out of 10 F0 readings) of any tested neutral-tone syllable, the whole pitch contour was excluded from averaging. Each chart has 4 (lexical tone of the preceding X)* 2 (lexical tone of the following Y) = 8 pitch contours.

3.3.2.1. **-perfective-nominalizer**

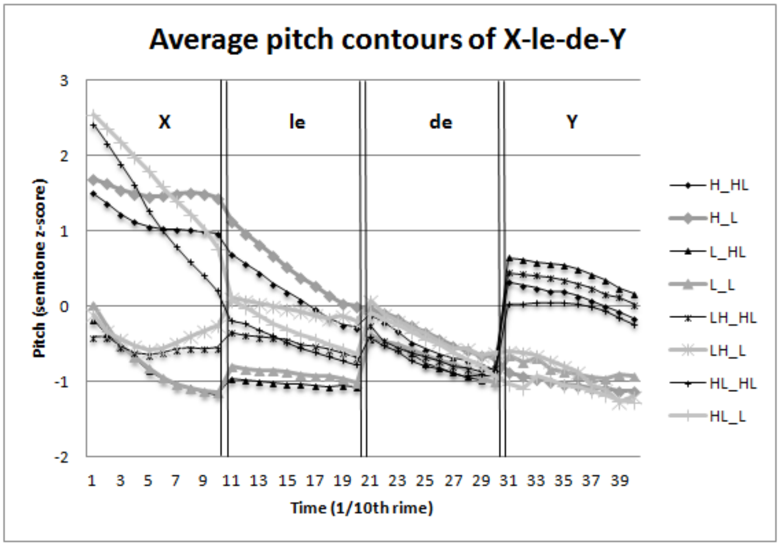


Figure 3.26 Result of two consecutive neutral tones: -PERF-NML

The result of two consecutive neutral tones with -perfective-nominalizer (-*le*-*de*) is shown in Figure 3.26. The eight pitch contours varied the four preceding tones (syllable X) and two following tones (syllable Y). The chart shows that the pitch contours at the neutral-tone perfective suffix -*le* reached from the end pitch of the preceding lexical tones toward 0 ~ -1 semitone z-score. A two-way ANOVA shows that the pitch at the mid position of the perfective suffix (*le*) was significantly influenced by the preceding tones (X) [$F(3,77) = 40.59$, $p < .001$] and surprisingly, by the following tones (Y) [$F(1,77) = 10.27$, $p = 0.0020 < .01$]. The interaction effect was not significant [$F(3,77) = 0.59$, $p = 0.62$].

In the following neutral-tone syllable (nominalizer suffix *de*), the pitch contours fell from 0 ~ -0.5 semitone z-score to -0.5~ -1 semitone z-score. We can see that the starting pitches of the second neutral-tone syllable were higher than the end pitches of the first neutral tone. This is probably due to the consonant perturbation by the initial voiceless obstruent /t/. A two-way ANOVA shows that the preceding tones (X) did not influence the pitch at the mid position [$F(3,77) = 1.82$, $p = 0.15$]. The main effect of the following tones (Y) was still significant [$F(3,77) = 8.90$, $p = 0.0038 < .01$]. The interaction effect was not significant [$F(3,77) = 0.48$, $p = 0.70$].

The result shows that the preceding tones (X) influenced the first neutral-tone syllable (*-le*), but not the second neutral-tone syllable (*-de*). The influence of the following tone (Y) is not expected because in a similar study in Standard Mandarin (Y. Chen & Xu 2006), the following tones have no influence on the pitch of the neutral tone syllables.

3.3.2.2. **-durative-nominalizer**

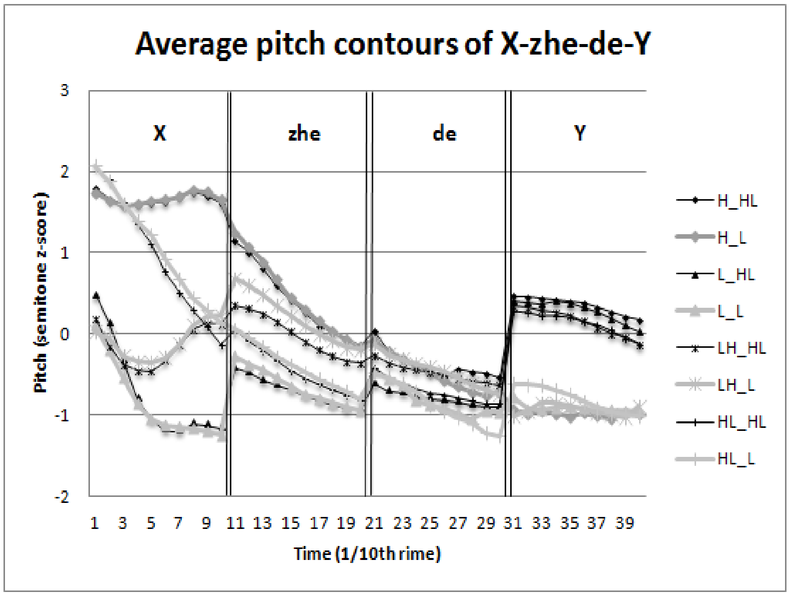


Figure 3.27 Result of two consecutive neutral tones: -DUR-NML

The result of two consecutive neutral tones -durative-nominalizer (*-zhe-de*) is shown in Figure 3.27. The eight pitch contours varied the four preceding tones (syllable X) and two following tones (syllable Y). Similar to the result of -perfective-nominalizer, the pitch at the end of the first neutral-tone durative *-zhe* reached toward 0 ~ -1 semitone z-score. A two-way ANOVA shows that the pitch at the mid position of the durative *-zhe* was influenced by the preceding lexical tone (X) [$F(3,76) = 38.19$, $p < .001$]. The influence of the following lexical tone (Y) was not significant [$F(1,76) = 0.96$, $p = 0.33$], and there was no significant interaction between the two factors [$F(3,76) = 0.23$, $p = 0.87$].

On the other hand, the second neutral tone, nominalizer *-de*, consistently reached from 0 ~ -0.6 semitone z-score toward -0.5 ~ -1 semitone z-score. However, a two-way

ANOVA shows that the preceding lexical tone (X) still influences the pitch at the mid position of the nominalizer *-de* [$F(3,76) = 4.22$, $p = 0.0082 < .01$]. There was no significant influence of the following lexical tone [$F(1,76) = 0.1638$, $p = 0.68$], nor was there any interaction between the two factors [$F(3,76) = 0.16$, $p = 0.92$].

As expected, the following lexical tones (Y) have no influence on the pitch of the neutral tones, neither were there any interaction effects. However, the influences of the preceding lexical tones (X) seems to last from the first neutral tone (*-zhe*) to the second neutral tone (*-de*).

3.3.2.3. **Consecutive neutral tones with reduplicants**

This section discusses the results of the consecutive neutral-tone syllables with reduplicants. Figure 3.28 shows the result of -reduplicant and -possessive (*de*) and Figure 3.29 shows the result of reduplicant-plural(*men*)-possessive(*de*). Both of the results show that the first neutral-tone syllables (reduplicant) show different pitch patterns. The reduplicants were high after the initial /H/, high falling after the initial /LH/, and moving to mid-low after initial /L/ and /HL/. The reduplicant seemed to have lexical tones /H/ after /H/ and /HL/ after /LH/. Only the reduplicants following /L/ and /HL/ show similar pitch patterns with other neutral-tone syllables. The pitch behavior of the reduplicant after /LH/ was different from what was found in 3.1.7.5, which had a mid-low pitch target. As noted before, all the tested word in reduplications are kinship terms. The tested /LH-Ø/ word in 3.1.7.5 was *shu-shu* /LH-Ø/ 'uncle' and the tested /LH-Ø/ word here is *po-po* /LH-Ø/ 'mother-in-law'. The [LH-HL]-like pitch contour of the word *po-po* /LH-Ø/ was probably lexicalized.

A two-way ANOVA was conducted to examine the pitch of the reduplicants. As shown in Figure 3.28, at the mid position of the reduplicant in -REDUP-POSS, a two-way ANOVA yielded a main effect for the preceding tones (X) [$F(3, 58) = 90.74$, $p < .001$]. The main effect of the following tones (Y) was not significant [$F(1, 58) = 0.15$, $p = 0.70$]. The interaction effect was not significant either [$F(3,58) = 0.65$, $p = 0.58$]. Also, at the mid position of the reduplicant in -REDUP-PL-POSS as shown in Figure 3.29, the influence of the preceding tone (X) was significant [$F(3,46) = 79.28$, $p < .001$]. The influence of the following tone (Y) was not significant [$F(1,46) = 0.26$, $p = 0.61$], and there was no interaction between the preceding tone and the following tone [$F(3,46) = 0.46$, $p = 0.71$].

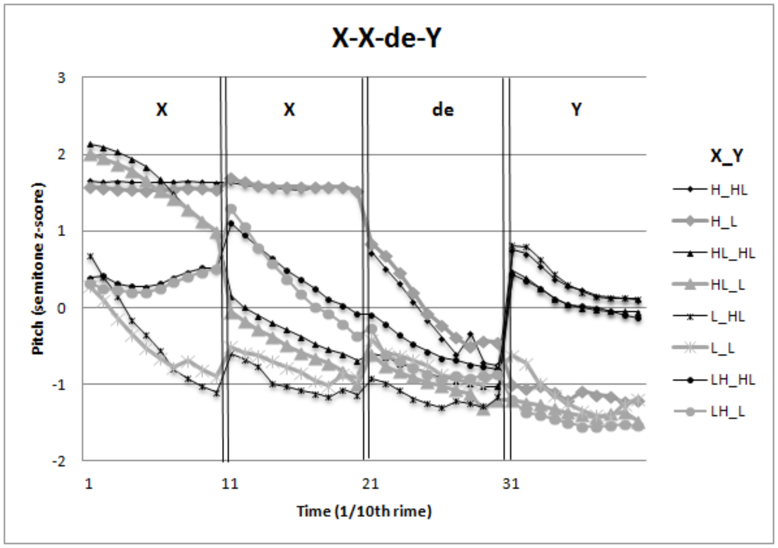


Figure 3.28 Result of two consecutive neutral tones: -REDUP-POSS

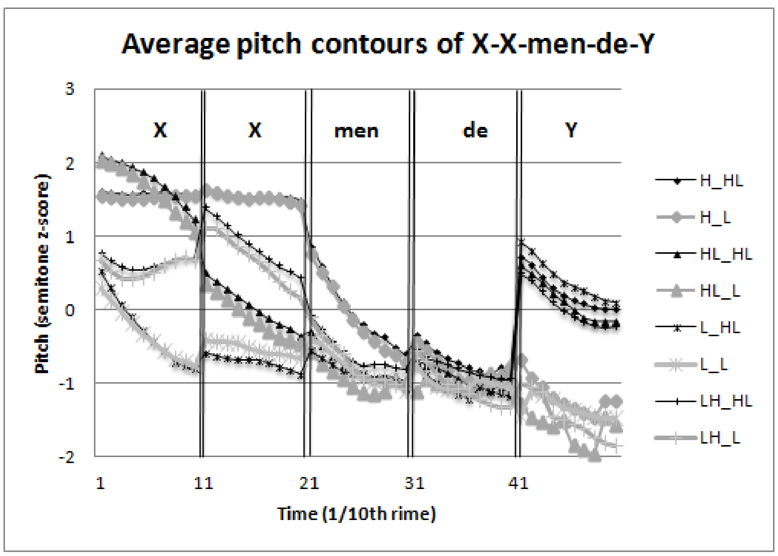
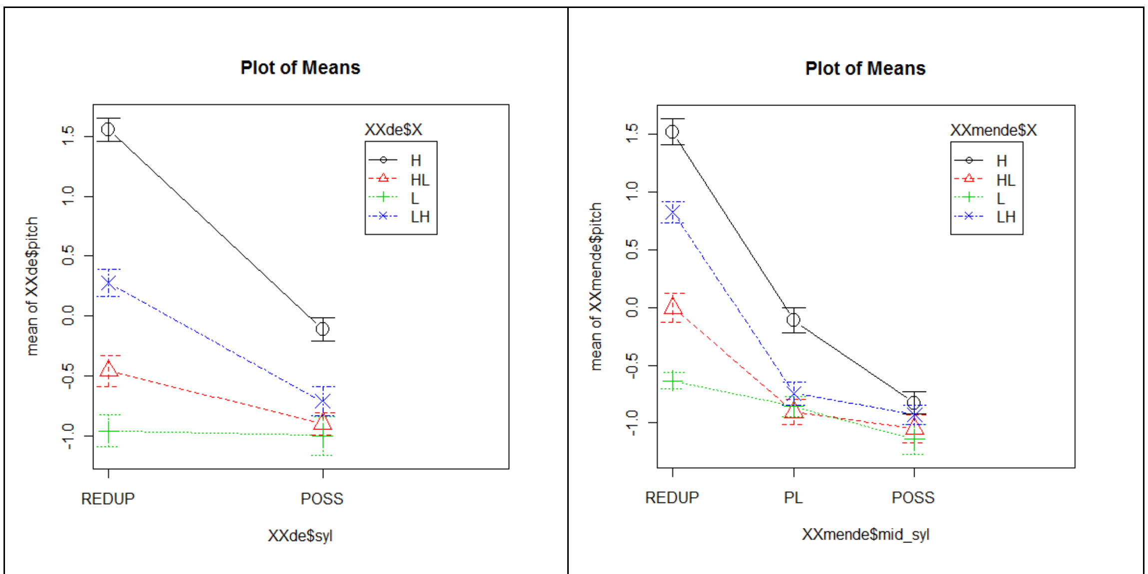


Figure 3.29 Result of three consecutive neutral tones: -REDUP-PL-POSS

As for the second neutral tone in -reduplicant-possessive, the pitch of the neutral-tone possessive (*-de*) after the reduplicant reached to -0.5 ~ -1.2 semitone z-score as shown in Figure 3.28. A two-way ANOVA shows that at the mid position of the neutral-tone possessive suffix (*de*), the main effect of the preceding tones (X) was still significant [$F(3,58) = 12.82$, $p < .001$]. The main effect of the following tones (Y) was not significant [$F(1, 58) = 0.02$, $p = 0.90$]. There was no significant interaction of the two factors [$F(3,58) = 1.88$, $p = 0.14$].

The three consecutive neutral tones -reduplicant-plural-possessive in Figure 3.29 shows that the first two neutral tones are similar to Figure 3.28. At the mid position of the second neutral tone (plural suffix -*men*), the influence of the preceding tones (X) was still significant [$F(3,46) = 13.37$, $p < .001$], and there was no significant influence of the following tone (Y) [$F(1,46) = 0.85$, $p = 0.36$] or significant interaction between the preceding tones and the following tones [$F(3,46) = 0.21$, $p = 0.89$]. However, at the mid position of the third neutral tone (possessive suffix -*de*), the influence of the preceding tones (X) was not significant [$F(3,46) = 1.87$, $p = 0.15$]. The influence of the following tones (Y) [$F(1,46) = 3.14$, $p = 0.08$] or the interaction between the two factors [$F(3,46) = 0.45$, $p = 0.72$] were not significant either. Both of these consecutive neutral tones show that the following lexical tone (Y) did not seem to have any effects on the neutral tone, and there were no interaction effect between the preceding (X) and following (Y) lexical tone. However, the preceding lexical tone (X) affected the following neutral-tone syllables.



a. -REDUP-POSS   b. -REDUP-PL-POSS

Figure 3.30 Plot of means of consecutive neutral tones at the mid position

As demonstrated in both of the consecutive neutral tones with reduplicants in Figure 3.30, the first neutral tones were heavily influenced by the preceding lexical tones. The pitch of the second neutral tones were lowered to 0 ~ -1 semitone z-score, but the influence of the preceding tones was still significant. On the third neutral tone, the influence of the preceding lexical tones was finally overcome and the mean pitches were around -1 semitone z-score.

The results of consecutive neutral tones with reduplicant are discussed above. However, some pitch contours were excluded from averaging because their pitch contours of the reduplicants were obviously different from the others. As mentioned in the previous section on reduplication, there were some tonal variations on the reduplications.

First, in X-REDUP-PL-POSS, F1 and M4 produced different pitch contours when X was /H/ and Y was /L/. While other speakers produced the reduplicated /H/ *ma* with a high pitch, F1 and M4's reduplicated /H/ *ma* had a falling pitch contour reaching to mid pitch range when Y was /L/. Figure 3.31 shows the pitch contours of X-REDUP-PL-POSS-Y when X was /H/ produced by F1 and M4. As shown in the figure, when Y was /HL/, the reduplicated X has a high pitch. On the other hand, when Y was /L/, the reduplicated X fell to a mid pitch range. This pitch pattern of high falling is similar to the pitch patterns of /H-Ø/ in Standard Mandarin. However, both of the speakers only have this pitch patterns in X-REDUP-PL-POSS before /L/. This pitch pattern is not found in the instances of *ma-ma* /H-Ø/ in words, in X-REDUP-PL-POSS before /HL/, or in X-REDUP-POSS-Y.

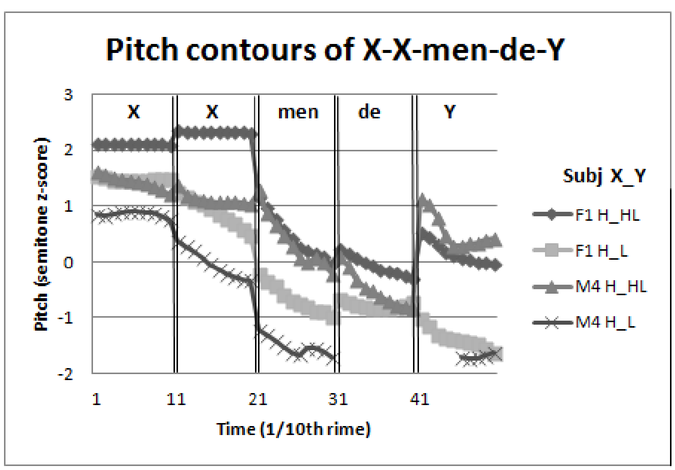
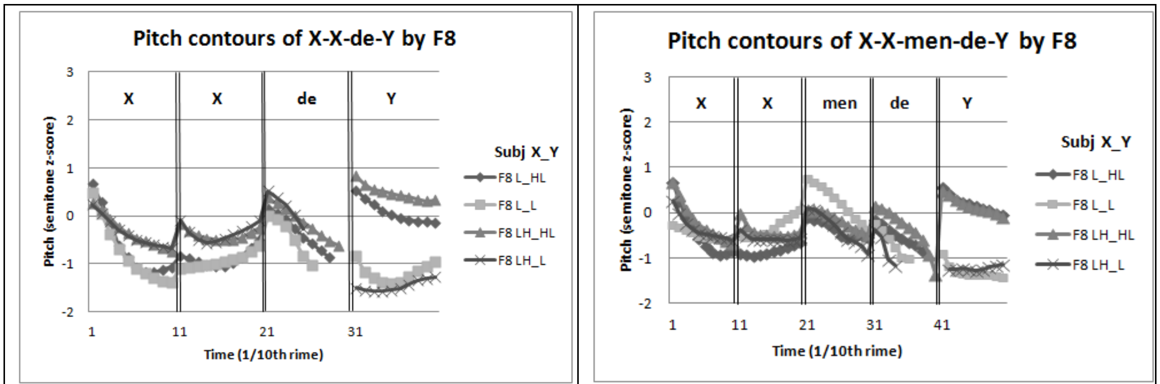


Figure 3.31 Pitch contours of X /H/-REDUP-PL-POSS-Y by Subject F1 and M4

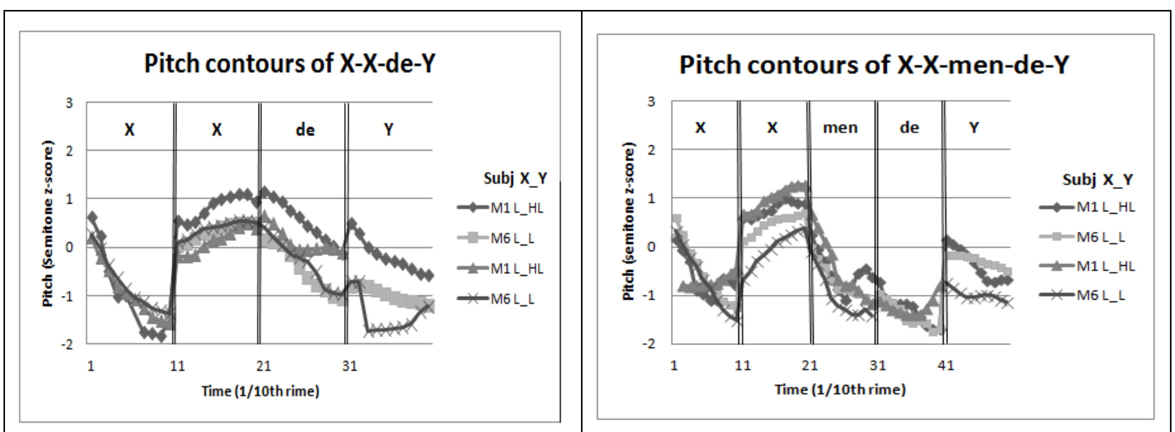
Second, in X-REDUP-PL-POSS-Y and X-REDUP-POSS-Y., subject F8 consistently produced the reduplication *nai-nai* /L-Ø/ ‘grandmother’ and *po-po* /LH-Ø/ ‘mother-in-law’ with a [L-LH] pitch contour as shown in Figure 3.32. As mentioned above in the reduplication study at 3.1.7.5, subject F8 also employed the [L-LH] pitch patterns in *shen-shen* /L-Ø/ ‘aunt’, which is a common way to show endearment.



a. X-REDUP-POSS-Y   b. X-REDUP-PL-POSS-Y

Figure 3.32 Pitch contours of consecutive neutral tones with reduplicated /L/ and /LH/ by F8

Lastly, subjects M1 and M6 produced the reduplicant *nai-nai* /L-Ø/ ‘grandmother’ with post-L rising which is a Standard Mandarin feature. The pitch contours are shown in Figure 3.33. Both of the subjects also produced *shen-shen* /L-Ø/ ‘aunt’ with a post-L rising in the previous reduplication section 3.1.7.5.



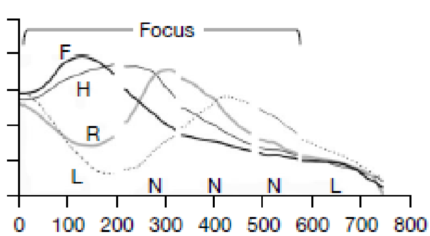
a. X-REDUP-POSS-Y   b. X-REDUP-PL-POSS-Y

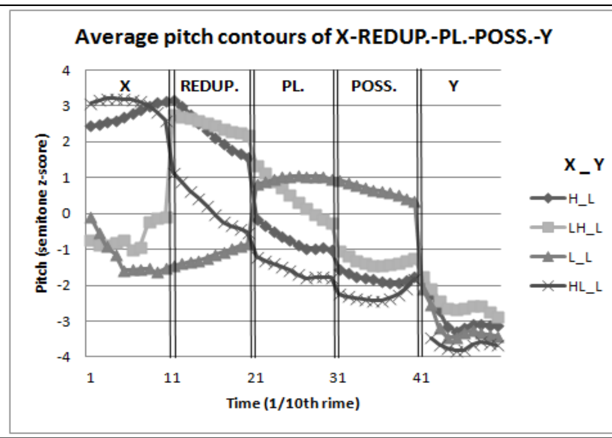
Figure 3.33 Pitch contours of consecutive neutral tones with reduplicated /L/ by M1 and M6

The exceptions shown above indicate that some speakers produced pitch patterns that resemble the pitch patterns of neutral tones in Standard Mandarin such /H-Ø/ and /L-Ø/. However, such pitch patterns were only found sporadically in certain kinship terms, in which the pitch patterns were easily lexicalized.

3.3.3. Discussion

This experiment slightly modified Chen and Xu's study (2006) on Standard Mandarin neutral tone and examined how the consecutive neutral tones differ in Taiwan Mandarin. Figure 3.34 shows the results of three consecutive neutral-tone syllables produced by SM speakers. Figure 3.34a is the result extracted from Chen and Xu (2006:54), and Figure 3.34b is the result of a male Standard Mandarin speaker who performed the exact same sentences in my study. The two results are presented slightly differently. In Chen and Xu's experiment, they used words with only sonorant initials so they could present the pitch contours of the whole syllables. However, in my experiment, the word lists that were used to create consecutive neutral-tone syllables do not always have a sonorant initial due to the limited distribution of reduplication in Taiwan Mandarin. As a result, I only measured and presented the F0 of the rimes in order to compare across different syllables.





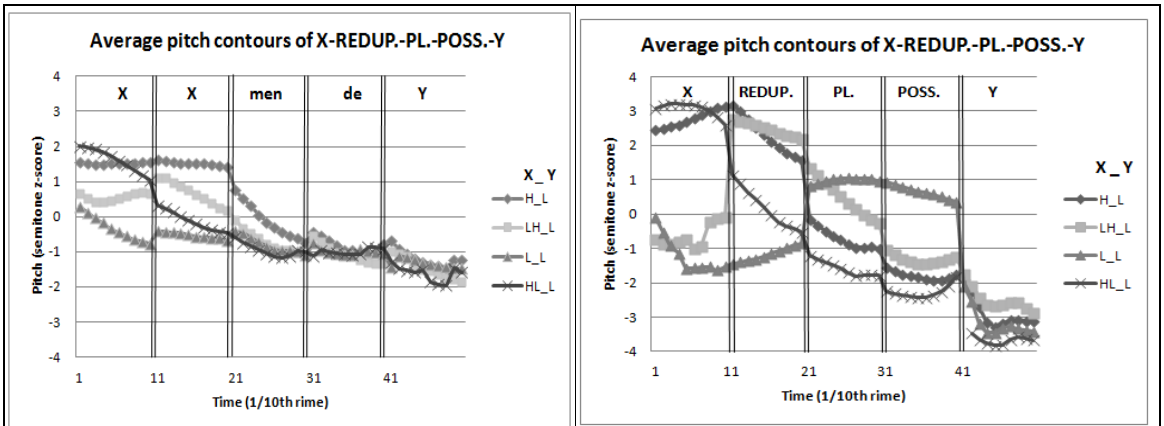
a. Chen & Xu 2006:54

b. A male Standard Mandarin speaker

Figure 3.34 X-REDUP-PL-POSS-HL produced by Standard Mandarin speakers

As we can see in the two figures, Figure 3.34b is very similar to the Standard Mandarin results shown in Figure 3.34a. The pitch movements in Figure 3.34b. resemble the second half of the syllables in Figure 3.34a. In Standard Mandarin, /H/ rises to the high pitch in the initial lexical tone (X), and the pitch gradually moves to the mid-low pitch target throughout the neutral-tone syllables; /LH/ does not rise to the higher pitch register until the first neutral-tone syllable, and then the pitch lowers to the mid-low range in the following neutral tones; /L/ first reaches to the low pitch range in the lexical tone and then the pitch gradually rises to the high register until the second neutral tone, and then the pitch slowly lowers to mid pitch range; /HL/ starts to fall from the lexical tone and has a fast pitch drop in the first neutral tone, and remains at mid-low pitches in the following neutral tones.

Figure 3.34 shows that results of my SM speaker are similar to Chen and Xu's (2006) results. I will therefore use these SM results to compare with my TM results as shown in Figure 3.35. Figure 3.35a shows the result of three consecutive neutral tones before /L/ produced by 12 TM subjects, and Figure 3.35b shows the result of the same sentences produced by an SM speaker.



a. Results of TM speakers	b. Results of SM speaker

Figure 3.35 Comparison of the results of X-REDUP-PL-POSS-Y by TM and SM

As shown in Figure 3.35, the results of the TM speakers (a) and the SM speaker (b) show great differences. The pitches of the preceding lexical tones ranged between 2 ~ -1 semitone z-score in TM, and the pitch merged to around -1 semitone z-score at the third neutral-tone syllables. However, the pitches of the preceding lexical tones ranged between 3 ~ -2 semitone z-score in SM, and the pitches at the third neutral-tone syllables still ranged between 0.5 ~ -2.5 semitone z-score. In TM, the pitch contours of the first neutral tone (reduplicant) were different from other neutral-tone syllables. The reduplicated /H/ had a high pitch and the reduplicated /LH/ had a high falling pitch. However, at the second neutral tone, the pitch contours in TM reached to a mid-low pitch target, and at the third neutral tone the four pitch contours were overlapped. On the other hand, in SM the pitch contours varied even at the third neutral tone. As mentioned in 3.3.2.3, in TM the influences of the preceding tones were significant at the mid position of the first and the second neutral tone, but not at the third neutral tone. However, Chen and Xu's (2006: 61) study showed that for SM the influences of the preceding tones were significant at the mid position of all three neutral-tone syllables. The influence of the preceding tones was even significant at the end position of the third neutral tone. Aside from the obvious different pitch patterns of the /L/ contour, the influences of the preceding /H, LH, L/ were still obvious at the third neutral tone.

Chen and Xu's study (2006) suggested that the consecutive neutral tones in Standard Mandarin converged toward a pitch target of mid value. However, the pitch target was implemented with a weak articulatory strength. As a result, the influence of the preceding tones was still effective at the third neutral tone and the pitch target was not implemented yet. On the other hand, my result shows that the neutral tone in Taiwan Mandarin has a mid-low (around -1 semitone z-score) pitch target rather than a mid pitch target as in SM. Similar to SM, the pitch target of the neutral tone in TM also needs some time to implement. My results show that the influence of the preceding lexical tones was still significant at the second neutral tone as shown in -DUR-NML, -REDUP-POSS, and -REDUP-PL-POSS. However, Figure 3.35a shows that by the third neutral tone, the pitch target was implemented completely in Taiwan Mandarin. Considering that the first neutral tone in -REDUP-PL-POSS did not reach a mid-low target like other neutral-tone syllables after /H/ and /LH/, the articulatory strength to reach to the mid-low target is rather strong. Also, the result of -PERF-NML in this study also shows that the second neutral tone was not influenced by the preceding lexical tones. The articulatory strength to implement the pitch target in the neutral tone in TM is much stronger than in SM.

3.4. Summary

The results of the three experiments discussed in this chapter show that the neutral tone in Taiwan Mandarin does not behave like the neutral tone in Standard Mandarin. The neutral tones in disyllabic words, novel occurrences, and consecutive occurrences all indicated that there is no general post-L rising and the pitch contours of the neutral tone reach to a mid-low target. Although the pitch of the neutral-tone syllables does seem to be influenced by the preceding lexical tones, the influences of the preceding tones were overcome much faster than in Standard Mandarin. The first experiment on disyllabic words shows that the influences of the preceding tones were no longer effective at the late position of the neutral-tone syllables of *-de*, *-ge*, and three reduplications. The third experiment on consecutive neutral tones further confirms the exact mid-low pitch target. The preceding lexical tones did not influence the mid position of the second neutral tone in -PERF-NML and the mid position of the third neutral tone in -REDUP-PL-POSS On the other hand, a previous study showed that in Standard Mandarin the influences of the preceding lexical tones were still significant even at the third neutral tone.

The results do show that some neutral-tone syllables have fixed tonal patterns. As shown in the disyllabic-word experiment, the neutral tone *-tou* was /LH/ and some instances of *-men* were also /LH/. These examples show that TM speakers may produce a lexical tone when the original tone is obvious. When the original tone is not obvious, they may even find a candidate tone from other cues (writing). The reason why other neutral-tone syllables such as *-zhe, -de, -le* do not have a lexical tone representation is simply because they do not have an obvious original tone. The consecutive neutral-tone experiment also shows that some reduplicants seem to have lexical tones. *ma-ma* is /H-H/, and *po-po* is /LH-HL/. This is not surprising since reduplication is no longer a productive process, and most of the reduplications for nicknames have fixed tonal patterns.

Also, my results show that frequent words and kinship terms seem to exhibit variations. For example, *she-me* /LH-Ø/ 'what', *duo-me* /H-Ø/ 'such' all show speaker variation in terms of the pitch patterns. These variations were a specific innovation in

Taiwan Mandarin that happens to frequent words. As for kinship terms, one speaker produced words like *shen-shen* /L-Ø/ 'aunt', *po-po* /LH-Ø/ 'mother-in-law/grandma', and *nai-nai* /L-Ø/ 'grandmother' with a fixed pitch pattern /L-LH/. This is also a documented pitch pattern in Taiwan Mandarin often used to show endearment (H.-c. Hsu 2006). Several speakers sporadically produced kinship terms with Standard Mandarin pitch patterns. For example /L-Ø/ was produced with post-L rising in *shen-shen* /L-Ø/ 'aunt' and *nai-nai* /L-Ø/ 'grandmother'; or *ma-ma* /H-Ø/ 'mother' was produced with a high-high falling pitch pattern. These sporadic instances suggest that some lexicalized pitch patterns of Standard Mandarin have entered the language through language contact. As a result, speakers only have the Standard Mandarin-like pitch patterns in some words or certain situations, but not in novel formations.

Some of the data suggest that for some TM speakers the neutral tone might be treated as a low tone phonologically. Two speakers produced the disyllabic word *yi-ge* /LH-Ø/ with a falling-low pitch patterns suggesting that the neutral-tone *-ge* was treated like a low tone because the low tone trigger the sandhi rule to change *yi* into a falling tone. Although these two speakers might have learned the lexicalized pitch pattern of *yi-ge* 'one', this variation is still motivated by the fact that some TM speakers treated the neutral-tone *ge* syllable like a low tone. In the second experiment, the one on novel forms, one of the two speakers actually applied Tone 3 sandhi rule to all the neutral-tone syllables. This shows that, at least for that specific speaker, the neutral tone is the same as the low tone (Tone 3) phonologically.

My results show that the neutral tone in Taiwan Mandarin is very different from the neutral tone in Standard Mandarin. The neutral tone in Taiwan Mandarin is definitely not neutralized the same way as in Standard Mandarin. The pitch target of the TM neutral-tone is different from that of SM. The pitch target is implemented much faster in TM than in SM, which is lexical-tone like. My result even suggests that some speakers treated the neutral tone like a low tone. In the next chapter I will discuss the difference between the neutral tone and the low tone in Taiwan Mandarin.

CHAPTER 4. COMPARISON BETWEEN LEXICAL-TONE AND NEUTRAL-TONE SYLLABLES

Chapter Three discussed the pitch behavior of the neutral tone in Taiwan Mandarin and showed that most of the neutral tone syllables do not behave like the neutral tone in Standard Mandarin. Most of the neutral-tone syllables have a rather low pitch target that resembles a low tone /L/ (Tone 3). This chapter compares the acoustic properties, pitch and duration, of these neutral-tone syllables to the properties of the similar low-tone syllables. Vowel quality is not in the scope of this study because the destressing of many neutral-tone syllables is a diachronic process. Synchronically the reduced vowel /ə/ is very similar to one of the stressed vowels, /ɤ/. If the acoustic properties of the two are similar, it is very likely that the neutral tone has merged into the low tone in these syllables; if the acoustic properties of the two are different, it is important to characterize the acoustic differences between the neutral tone and the low tone. Also, this chapter compares some neutral-tone syllables that seem to have a rising tone with rising-tone syllables.

4.1. Methods

The pitch contours of the neutral-tone syllables in TM seem to have a target and resemble lexical tones. Therefore it is important to compare the pitch contours of the neutral-tone syllables with the lexical-tone syllables that have similar pitch patterns. Aside from the pitch patterns, duration is also an important feature of neutral-tone syllables. This experiment also compares the durations of the neutral-tone syllables with the similar lexical-tone syllables.

4.1.1. Materials

In order to compare the neutral-tone syllables and the lexical-tone syllables, subjects in Chapter Three were also asked to read a list of disyllabic words/phrases that contain the corresponding lexical-tone syllables. These corresponding syllables all preferably have the same segments in order to control for other factors, and they have the lexical tone that is similar to the pitch contour of the neutral-tone syllables. For example, the neutral-tone diminutive suffix *tou* /Ø/ has a pitch contour that is similar to *tou* /LH/'head' or 'throw'. In this experiment, the subjects were asked to read the disyllabic words/phrases in which *tou* /LH/ 'head' or 'throw' was the second syllable and the preceding lexical tones varied: *gong-tou* /H-LH/ 'public vote', *she-tou* /LH-LH/ 'snake head', *ma-tou* /L-LH/'horse head', and *peng-tou* /HL-LH/ 'to meet up'. Table 4.1 shows the list of the corresponding syllables and these syllables in disyllabic words/phrases.

As shown in Table 4.1, the neutral-tone syllable -*men* /Ø/ 'PL' has corresponding syllables -*men* /LH/ 'door/gate' and -*ben* /L/ 'origin'. -*men* /LH/ was included because the character of -*men* /Ø/ 們 has a phonetic element 門 (*men* /LH/) which may cause the speakers to pronounce -*men* /Ø/ with a rising tone. -*ben* /L/ was included because many occurrences of -*men* /Ø/ in Taiwan Mandarin have a low pitch target. -*ben* /L/ was included instead of -*men* /L/ because -*men* /L/ does not exist in Mandarin due to a phonological gap. Also, the low-tone corresponding syllable -*ben* /L/ was only tested in the combination of /H-L/, /LH-L/, and /HL-L/ because the Tone 3 Sandhi rule applies to /L-L/ resulting in LH-L.

Similarly, the neutral-tone syllable -*tou* /Ø/ has the corresponding syllables -*tou* /LH/ 'throw; head' and -*dou* /L/ 'shiver; dipper-shape'. -*tou* /LH/ was included because it is the tone of -*tou* /Ø/ before destressing. -*dou* /L/ was included because -*tou* /L/ does not exist in Mandarin.

The neutral-tone syllable -*zi* /Ø/ 'DIM.' has a corresponding syllable -*zi* /L/ 'person; first of the 12 earthly branches; son'. The low tone -*zi* /L/ is the same character as the neutral-tone -*zi*, and the neutral-tone -*zi* /Ø/ is the unstressed -*zi* /L/.

Table 4.1 Testing corresponding control syllables with the same segment

neutral tone syllables	corresponding syllables	[H]	[LH]	[L]	[HL]
men 'PL'	*men* /LH/ 門 'door/gate'	*jia-men* 'house door'	*fo-men* 'buddhism'	*bei-men* 'north gate'	*da-men* 'gate'
	***ben* /L/** 本 'origin'	*ji-ben* 'basic'	*cheng-ben* 'cost'	-	*ke-ben* 'textbook'
tou 'DIM.'	*tou* /LH/ 投;頭 'throw; head'	*gong-tou* 'vote'	*she-tou* 'snake head'	*ma-tou* 'horse head'	*peng-tou* 'to meet up'
	***dou* /L/** 抖;斗 'shiver; dipper-shape'	*fa-dou* 'to shiver'	*bei-dou* 'Big Dipper'	-	*yun-dou* 'an iron'
zi 'DIM.'	*zi* /L/ 子 'person; first of the 12 earthly branches; son'	*jun-zi* 'gentleman'	*jia-zi* '60 years'	-	*nie-zi* 'unfilial child'
reduplication	*ma* /H/ 媽 'mom' *shu* /L/ 薯 'potatoes' *shen* /L/ 審 'judge' *di* /L/ 底 'bottom'	*ma-ma* 'mother'	*fan-shu* 'sweet potato'	*ping-shen* 'judges'	*di-di* 'underground'
le 'PERF'	N/A				
de 'POSS'	N/A				
me 'such'	N/A				
zhe 'DUR'	*zhe* /L/ 者 'person'	*shang-zhe* 'injured person'	*du-zhe* 'reader'	-	*zuo-zhe* 'author'
ge 'CL'	*ge* /L/ 葛 'a last name'	*bei-ge* 'boycott'	N/A	-	N/A

The reduplicated syllables were compared to four different corresponding syllables that have a similar pitch and the same segments. *ma-ma* /H-Ø/ was compared to *wo-ma* /L-H/ 'my Mother' because the reduplicated *ma* /Ø/ has a high pitch; *shu-shu* /LH-Ø/ was compared to *fan-shu* /H-Ø/ 'sweet potatoes' because the reduplicated *shu* /Ø/ has a low pitch target; *shen-shen* /L-Ø/ was compared to *ping-shen* /LH-L/ 'judges' because the reduplicated *shen* /Ø/ mostly has a low pitch; and *di-di* /HL-Ø/ was compared to *di-di* /HL-L/ 'underground' because the reduplicated *di* /Ø/ has a low pitch.

There are five neutral-tone syllables with a /ə/ rime, they are *le* /lə/ 'PERF', *de* /tə/ 'POSS; NML', *me* /mə/ 'such', *zhe* /ʈʂə/ 'DUR' and ge /kə/ 'CL'. Out of the five neutral-tone syllables, only *zhe* /Ø/ 'DUR' has a corresponding low-tone syllable *zhe* /L/ 'person'. The syllables *le*, *de*, *me*, *ge* do not have low-tone corresponding syllables in Mandarin due to a gap in Mandarin phonology[16]. As a result, the neutral-tone syllables *le, de, me*, *ge* were compared with other syllables with a /ə/ rime and a low tone. As shown in Table 4.2, *she* /ʂə/ /L/ and *che* /ʈʂʰə/ /L/ were included to compare with *le, de*, *ge*, and *me*. Phonetically the rimes of *che* and *she* are described as [ɤ], while the rimes of the neutral-tone *de, le, me*, *ge* are described as [ə] because they are unstressed and reduced. However, both [ɤ] and [ə] belong to the same phoneme /ə/ in Mandarin.

Table 4.2 Testing corresponding controlled syllables with the same rime

neutral tone syllables	corresponding syllables with /ə/	[H]	[LH]	[L]	[HL]
le, de, ge ,me	*she* /L/ 捨 'let go'	*shi-she* 'to give'	*nan-she* 'hard to let go'		*bu-she* 'can't let go'
	che /L/ 扯 'pull'	*la-che* 'pull and drag'	*xian-che* 'chitchat'		*luan-che* 'chitchat'

The disyllabic words/phrases in Table 4.1 and Table 4.2 were put in the same frame sentence: *qing shuo*_____*ba ci*. 'Please say______ eight times.' The pitch contours and the durations of these disyllabic words/phrases were compared to the results from the previous chapter.

4.1.2. Procedures

The same subjects in 3.1.3 were asked to read the disyllabic words/phrases in Table 4.1 and Table 4.2. All the disyllabic words were carried in the frame sentence *qingshuo X X baci* 'Please say XX eight times'. All the sentences were written in Chinese characters

16 *ge*/Ø/ does have a low-tone corresponding syllable *ge*/L/. However, the syllable is very limited in its occurrences in words. As a result, it was not included as a corresponding syllable.

and the subjects were asked to read every sentence just once. The subjects could repeat the sentences if they felt they made a mistake. The tested sentences in this section were mixed with the tested sentences from 3.1 along with the fillers. The subjects were asked to read a total of 118 sentences. The tested words and fillers were randomly ordered. The full list of the sentences is attached in Appendix B.

The recorded disyllabic words/phrases were treated the same way as described in 3.1.6 in which the rimes of the disyllabic words/phrases were hand-labeled and their pitch contours were extracted in order to compare the pitch contours of the neutral tone with the corresponding lexical tones.

In addition, the durations of the segmented rimes in Chapter Three and this chapter were measured using a Praat script written by Mietta Lennes. The script measures the duration of any labeled intervals in a given textgrid. The rimes were chosen to be measured instead of syllables because languages generally do not count onset in their calculation of syllable weight; only the rimes are relevant to syllable weight. If the stressed and unstressed syllables in TM are different, I would expect the duration differences to show on the rimes.

4.2. Results

Sections 4.2.1 to 4.2.5 compare the pitch contours and the durations of the neutral-tone -*men*, -*tou*, and -*zi* /ə/ rime, and reduplications with their lexical tone counterparts.

4.2.1. Neutral-tone -*men*

The neutral-tone -*men* ‘PL’ was compared to the low tone *ben* ‘origin’ and the rising tone *men* ‘door/gate’. -*men* /Ø/ is first compared with *ben* /L/, and then compared with *men* /LH/ as follows.

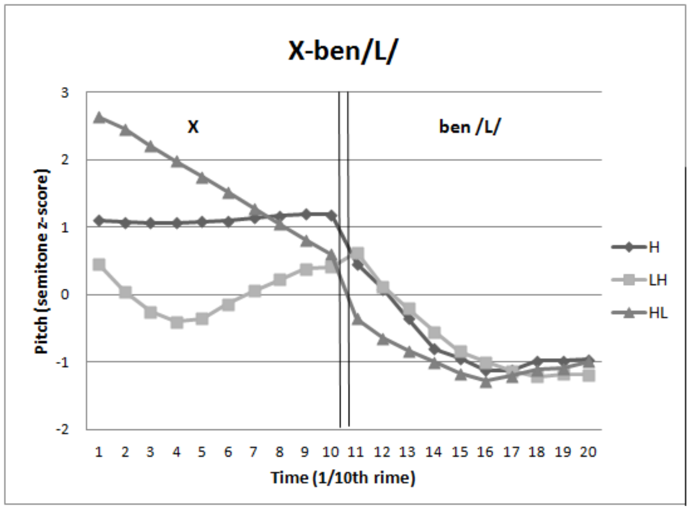


Figure 4.1 Average pitch contours of *X-ben* /L/

The average pitch contours of X-*ben* /L/ are shown in Figure 4.1. Only three different lexical tones appear in the first syllable because consecutive low tones are not allowed in Mandarin due to the Tone 3 Sandhi rule. As shown in the figure, the pitch contours merge by the midpoint of the *ben* syllable, overcoming the preceding tonal differences. While the influence of the preceding tones at the early position of the *ben* syllable was significant [$F(2,18) = 6.97$, $p = 0.0057 < .01$), there were no significant influences of the preceding tones found at the mid and late position [Mid: $F(2,18) =1.94$, $p = 0.17$; Late: $F(2,18) = 0.92$, $p = 0.42$]. The pitch contours merged to almost -1 semitone z-score at the mid position ($M = -0.98$, $SD = 0.34$) and stayed around -1 semitone z-score at the late position ($M = -1.13$, $SD = 0.31$). The pitch target of the low tone was implemented very quickly in *ben* /L/.

Figure 4.2 compares *ben* /L/ with *men* /Ø/. We can see that the pitch contours of *ben* /L/ merged in the first half of the syllable (lighter lines). However, the pitch contours of *men* /Ø/ (darker lines) were influenced by the preceding lexical tones at the early and mid position as mentioned in Chapter Three. The figure shows that the pitch contours of *ben* /L/ and *men* /Ø/ were rather different when the preceding tones were /LH/ or /H/. When after /LH/ or /H/, the pitch contours of *ben* /L/ fell to -1 semitone z-score in the first half and stayed around -1 semitone z-score. However, the pitch contours of *men* /Ø/ gradually fell to -0.5 ~ -1 semitone z-score throughout the syllable. The pitch differences between *men* /Ø/ and *ben* /L/ were significant at the mid position when they were after /LH/ [$t(18)$

= -4.13, $p < .001$] and after /H/ [$t(6) = -2.86$, $p = 0.029 < .05$]. The pitch differences between *men* /Ø/ and *ben* /L/ remained statistically significant at the late position when following /LH/ [$t(17) = -5.95$, $p < .001$]. On the other hand, when the preceding tone was /HL/ or /L/, the pitch contours of *men* /Ø/ and *ben* /L/ were quite similar. The pitch differences at all three positions were not significant.

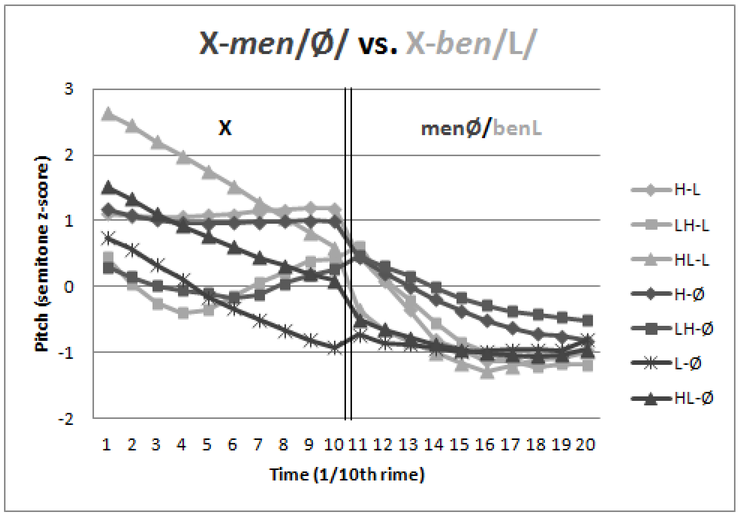


Figure 4.2 Average pitch contours of X-*men* /Ø/ and X-*ben* /L/

In addition, the numbers of the syllables that had creaky voice were different between the *ben* /L/ and *men* /Ø/. As mentioned in data analysis in 3.1.6, the pitch contours of many tested syllables were excluded when most of the F0 readings were not available. Many of the syllables without F0 readings had creaky voice that resulted from the lowering of the pitch. Therefore it is also important to compare the numbers of the absent syllables that were creaky. As shown in Table 4.3, when the preceding tone was /H/, seven out of twelve productions of *ben* /L/ had creaky voice, comparing to only one out of twelve productions of *men* /Ø/. Similarly, more instances of *ben* /L/ had creaky voice than *men* /Ø/ after /LH/ and /HL/. Interestingly, six out of twelve instances of *men* /Ø/ had a creaky voice after /L/. This shows that it was difficult to maintain the low pitch in modal voice.

Table 4.3 Numbers of excluded syllables with creaky voice – for *ben* /L/, *men* /Ø/, *men* /LH/

Preceding tone / Syllable	H	LH	L	HL
ben /L/	7/12	3/12	-	5/12
men /Ø/	1/12	1/12	6/12	3/12
men /LH/	0/12	0/12	0/12	0/12

As for the comparison of the durations, Table 4.4 shows the average durations of the rime of the first syllable (S1) and the second syllable (S2) of the tested disyllabic words. The durations are presented in milliseconds (ms). The S2/S1 ratios are also presented in the chart in order to illustrate the rhythmic information of S2 to S1. For each S2 rime, two S2/S1 ratios are presented: The S2/S1 ratio in the third column is the ratio of averaged S2 duration to averaged S1 duration; while the S2/S1 ratio in the fourth column is the average of all twelve S2/S1 ratios. The two S2/S1 ratios are provided to give an idea of speed variation. Slower speech would have more influence on the average S2/average S1 while average S2/S1 ratio tends to capture the rhythmic proportion of the disyllabic words.

Table 4.4 Duration comparison between neutral-tone -*men* and low-tone -*ben*

	S1 duration (ms)	S2 duration (ms)	avg. S2/ avg. S1	avg. S2/S1 ratio
men /Ø/	117.3	147.3	1.256	1.317
ben /L/	115.3	144.5	1.253	1.286

As shown in Table 4.4, the duration of the rime of S2 -*men* /Ø/ (M = 147.3, SD = 33.1) was very similar to that of *ben* /L/ (M = 144.5, SD = 44.8). There was no significant difference between the two average S2 durations [$t(62) = -0.318$, $p = 0.75$]. The difference between their S1 durations was not significant [$t(79) = -0.347$, p = 0.73]. The differences between their averaged S2/S1 ratios was not significant either [$t(71) = -0.342$, $p = 0.733$].

To sum up, the results above show that the neutral-tone -*men* has similar duration and rhythmic quality with the low-tone *ben*. However, their pitch contours were different from each other. The pitch target of the low tone was implemented by the time that the midpoint of *ben* was reached, while the influence of the preceding tone was still significant at the late position for the neutral-tone -*men*.

Next, I compare -*men* /Ø/ with *men* /LH/. Figure 4.3 shows the average pitch contours of *men* /LH/ after four different preceding lexical tones. As shown in the figure, the pitch contours of *men* /LH/ first reached to -1 ~ -0.3 semitone z-score and rose to -0.6 ~ 0 semitone z-score. The pitch contours show that the rising tone in TM did not seem to rise as much as in SM. This feature of the rising tone is in agreement with Fon and Chiang's (1999) finding which showed that the rising tone in TM is [323], while the rising tone in SM is described as [35]. Also, the pitch contours were influenced by the preceding lexical tones, unlike the low tone *ben*. One-way ANOVA shows that the influence of the preceding lexical tones was significant at the early, mid, and late positions [Early: $F(3, 43) = 30.04$, $p < .001$; Mid: $F(3, 43) = 13.84$, $p < .001$; Late: $F(3, 43) =11.25$, $p < .001$]. At the late position, the pitch after /LH/ ($M = -0.12$, $SD = 0.24$) was significantly higher than the other three.

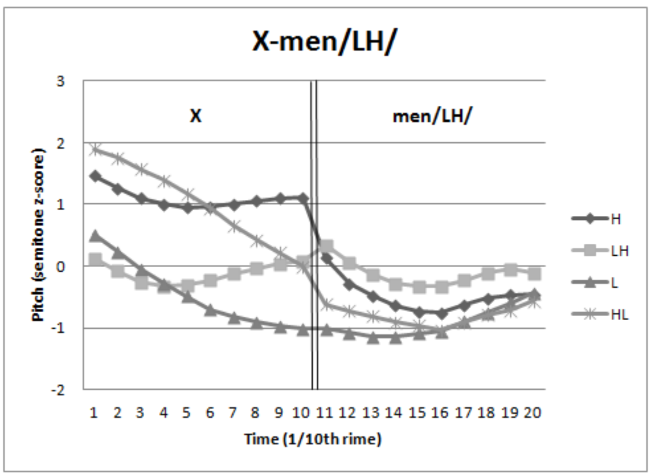


Figure 4.3 Average pitch contours of *X-men* /LH/

The pitch contours of *X-men* /Ø/ was similar to *X-men* /LH/ in terms of the influence of the preceding tones. Both *men* /LH/ and *men* /Ø/ were influenced by the preceding lexical tones. However, as shown in Figure 4.4, the shapes of the pitch contours were quite different between *men* /Ø/ and *men* /LH/: *men* /LH/ had obvious pitch rises in the second half of the syllable. The paired t-test showed that the pitch of *men* /LH/ at the late position ($M = -0.55$, $SD = 0.41$) was significantly higher than at the mid position ($M =$

-0.78, *SD* = 0.43) [*t*(46) = -5.61, *p* < .001]. On the contrary, the pitch contours of *men* /Ø/ fell in the second half of the syllable. The paired t-test showed that the pitch of *men* /Ø/ at the late position (*M* = -0.75, *SD* = 0.41) was significantly lower than at the mid position (*M* = -0.55, *SD* = 0.49) [*t*(36) = 4.50, *p* < .001].

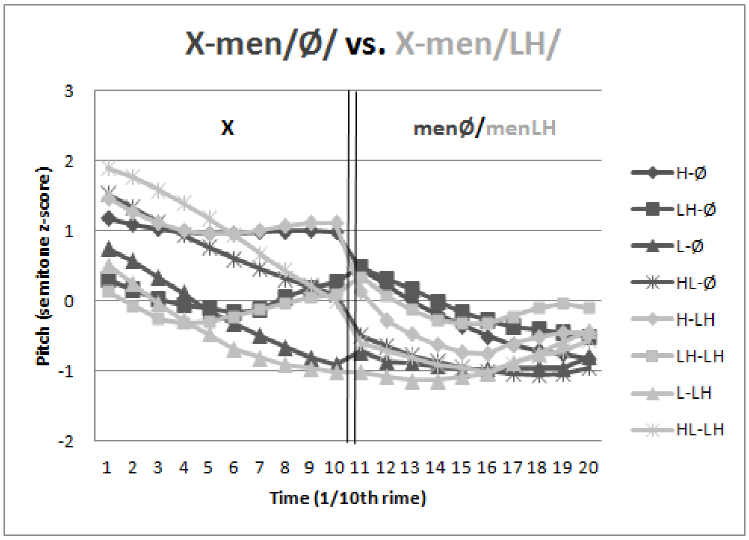


Figure 4.4 Average pitch contours of *X-men* /Ø/ and *X-men* /LH/

The differences in the shape of the pitch contours also created pitch differences between *men* /Ø/ and *men* /LH/. The pitch differences between *men* /LH/ and *men* /Ø/ were especially obvious when the preceding tone was /LH/ or /H/. When the preceding tone was /LH/, the pitch difference between *men* /LH/ and *men* /Ø/ was statistically significant at the late position [*t*(17) = 2.30, *p* = 0.034 < .05]. When the preceding tone was /H/, the pitch differences between *men* /LH/ and *men* /Ø/ were statistically significant at the early and mid position [Early: *t*(20) = -2.84, *p* = 0.010 < .05; Mid: *t*(20) = -2.62, *p* = 0.016 < .05]. However, when the preceding tone was /L/ or /HL/, the pitch differences between *men* /LH/ and *men* /Ø/ were not statistically significant in any of the three positions even though their pitch contours looked different.

Also, the pitch contour shape also affected the numbers of syllables that had creaky voice. As shown in Table 4.3 above, none of the *men* /LH/ syllables had creaky voice, while some of the *men* /Ø/ had creaky voice, especially after /L/ and /HL/.

As for the durational differences, Table 4.5 shows the average duration of S2- *men* /LH/ (M = 165.9, SD = 27.8) was longer than that of *men* /Ø/ (M = 147.3, SD = 33.1).

The difference was statistically significant [$t(91) = 2.99$, $p = 0.0036 < .01$]. This finding is consistent with the finding of Deng et al. (2008) who found that in Taiwan Mandarin, Tone 2 (rising) is longer than Tone 3 (low) in sentences. However, the differences between their average S2/S1 ratios were not statistically significant [$t(76) = -1.35$, $p = 0.18$]. The neutral-tone *men* were not lighter in disyllabic words. The S2 duration of *men* /Ø/ was actually longer than its S1 duration. This shows that the *men* /Ø/ is not shortened as an unstressed syllable in trochaic words.

Table 4.5 Duration comparison between neutral-tone -*men* and rising-tone -*men*

	S1 duration (ms)	S2 duration (ms)	avg. S2/ avg. S1	avg. S2/S1 ratio
men /Ø/	117.3	147.3	1.256	1.317
men /LH/	137.4	165.9	1.207	1.228

The results presented above show that the neutral-tone *men* was different from the rising-tone *men*. The pitch contours of *men* /LH/ had obvious pitch rise while this pitch shape was not found in *men* /Ø/. Also, the durations of *men* /LH/ were longer than the durations of *men* /Ø/.

This section compares *men* /Ø/ with *men* /LH/ and *ben* /L/. The results show that *men* /Ø/ had a similar duration to *ben* /L/, but it is shorter than *men* /LH/. However, the S2/S1 ratios also showed that the *men* /Ø/ syllable was not weak in disyllabic words. As for pitch, the pitch contours of *men* /Ø/ were also different from *men* /LH/ and *ben* /L/: *men* /Ø/ did not have the pitch rise like *men* /LH/, and *men* /Ø/'s pitch target were not implemented as fast as in *ben* /L/. The pitches of *men* /Ø/ were still influenced by the preceding lexical tones at the late position.

4.2.2. Neutral-tone -*tou*

The neutral-tone -*tou* /Ø/ was compared with the low-tone *dou* /L/ and the rising tone *tou* /LH/. In this section, I first compare *tou* /Ø/ with *dou* /L/, and then I compare *tou* /Ø/ with *tou* /LH/.

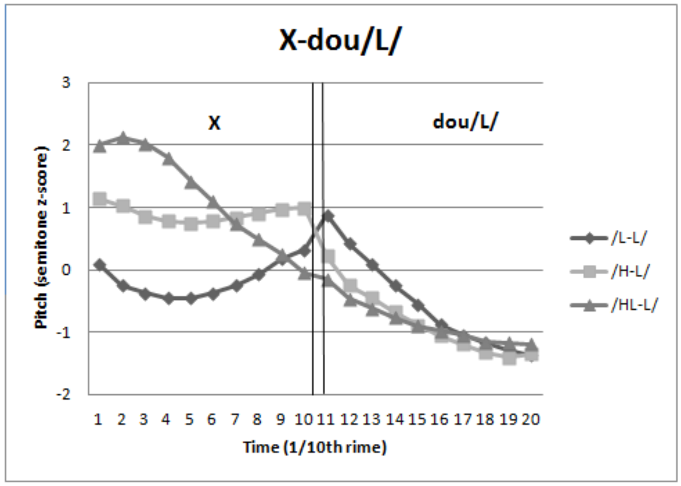


Figure 4.5 Average pitch contours of X-*dou* /L/

Figure 4.5 shows the average pitch contours of the disyllabic X-*dou* /L/ after /H, L, HL/ ([H, LH, HL]). The three average pitch contours of *dou* /L/ merged at the second half of the *dou* syllable and fell to -1~ -1.5 semitone z-score. The pitch differences resulting from the preceding lexical tone were overcome in the first half of the *dou* syllable. One-way ANOVA shows that the influence of the preceding lexical tones was only significant at the early position [$F(2,18) = 5.61$, $p = 0.013 < .05$], but not at the mid position or the late position [mid: $F(2,18) = 1.84$, $p = 0.19$; late: $F(2,18) = 0.73$, $p = 0.50$]

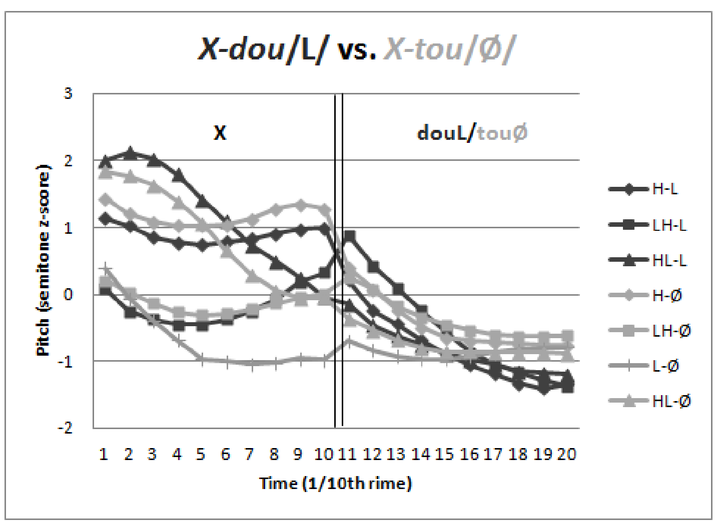


Figure 4.6 Average pitch contours of X-*tou* /Ø/ and X-*dou* /L/

Figure 4.6 compares the average pitch contours of X-*tou* /Ø/ with X-*dou* /L/. There was a clear pitch difference between *tou* /Ø/ and *dou* /L/ at the second half of the syllables. The pitch contours of *tou* /Ø/ reached to -0.5 ~ -1 semitone z-score while the pitch contours of *dou* /L/ reached to -1 ~ -1.5 semitone z-score. The difference between the average pitch of *tou* /Ø/ and *dou* /L/ at the late position was statistically significant [$t(45) = -5.44$, $p < .001$; *dou* /L/: $M = -1.22$, $SD = 0.30$; *tou* /Ø/: $M = -0.75$, $SD = 0.32$]. Also, as shown in Table 4.6, nearly half of the *dou* /L/ syllables had creaky voice, while only three instances of creaky voice were found in *tou* /Ø/ syllables.

Table 4.6 Numbers of excluded syllables with creaky voice - *dou* /L/, *tou* /Ø/, *tou* /LH/

Preceding tone / Syllables	H	LH	L	HL
dou /L/	5/12	5/12	-	5/12
tou /Ø/	3/12	0/12	0/12	0/12
tou /LH/	0/12	0/12	0/12	0/12

The duration of the rime of -*tou* /Ø/ was compared with the duration of the rime of -*dou* /L/. As shown in Table 4.7, the average duration of S2 -*tou* /Ø/ (M = 113.0, SD = 24.3) was slightly shorter than the duration of S2 *dou* /L/ (M = 121.1, SD = 22.3). However, the differences between *dou* /L/ and *tou* /Ø/ was not statistically significant [$t(79) = 1.57$, $p = 0.119$]. Their S1 durations were not significantly different [$t(76) = 0.32$, $p = 0.75$]. Furthermore, their average S2/S1 ratios were not significantly different either [($t(77)$=1.12, $p = 0.264$]. This result shows that in TM there is no obvious rhythmic difference between full-full X-*dou* /L/ and full-neutral X-*tou* /Ø/. Interestingly, both -*dou* /L/ and -*tou* /Ø/ have S2 durations that were shorter than their S1 durations. This trochaic rhythmic pattern is expected for full-neutral disyllabic words, but not expected for the full-full disyllabic words such as X-*dou* /L/.

Table 4.7 Duration comparison between the neutral-tone -*tou* and the low-tone -*dou*

	S1 duration (ms)	S2 duration (ms)	S2/S1 ratio	avg. S2/S1 ratio
tou /Ø/	151.6	113.0	0.745	0.761
dou /L/	153.6	121.1	0.788	0.804

The results above show that *tou* /Ø/ was not shorter than *dou* /L/ and there was no rhythmic difference between X-*tou* /Ø/ and X-*dou* /L/ words. However, the two differed in terms of the pitch target. The pitch target of -*tou* /L/ was significantly lower than that of -*tou* /Ø/. Next, I compare the neutral-tone *tou* /Ø/ with the rising-tone *tou* /LH/.

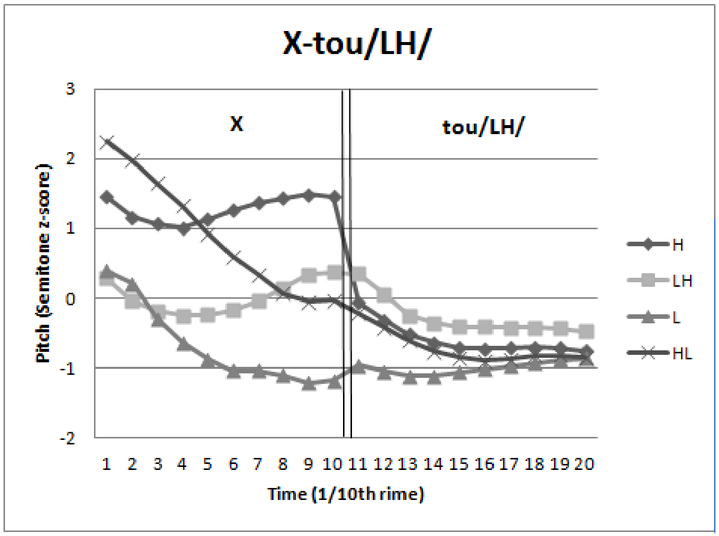


Figure 4.7 Average pitch contours of X-*tou* /LH/

Figure 4.7 shows the pitch contours of *tou* /LH/. The pitch contours show that the rising tone did not seem to have obvious pitch rise. As shown in the figure, the pitch contours almost stayed flat between -0.5~ -1 semitone z-score in the second half of the *tou* /LH/ syllable. Also, the pitch contours of *tou* /LH/ seemed to be influenced by the preceding lexical tones. One-way ANOVA shows that the influence of the preceding lexical tones was significant at the early, mid, and late positions [Early: $F(3,41) = 13.93$, $p < .001$; Mid: $F(3,41) = 6.69$, $p < .001$; Late: $F(3,41) = 4.99$, $p = 0.0048 < .01$].

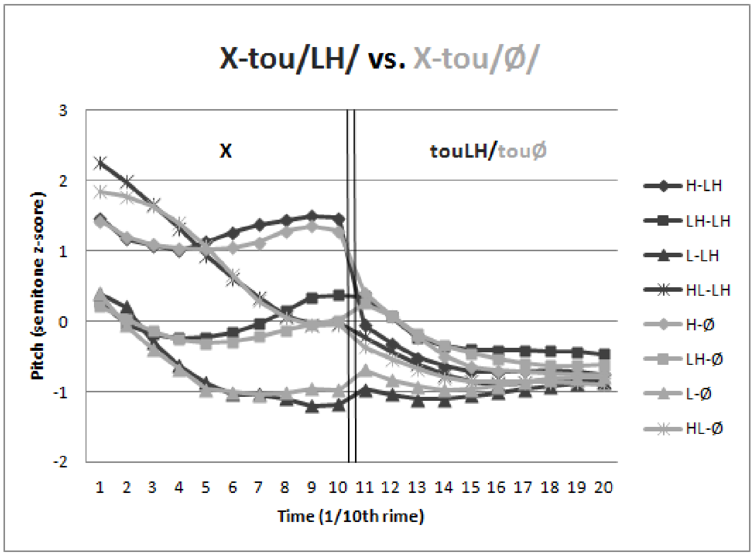


Figure 4.8 Average pitch contours of X-*tou* /LH/ and X-*tou* /Ø/

Figure 4.8 compares the average pitch contours of X-*tou* /Ø/ with X-*tou* /LH/. Both *tou* /Ø/ and *tou* /LH/ stayed flat around -0.5 ~ -1 semitone z-score. Their average pitch contours overlapped and the t-tests show that there were no statistical differences at early, mid, or late positions [Early: $t(84) = -0.66$, $p = 0.51$; Mid: $t(82) = -0.31$, $p = 0.76$; Late: $t(81) = 0.45$, $p = 0.66$]. There were also no statistical differences between the pitches of *tou* /LH/ and *tou* /Ø/ at any of the three positions when they were after the same tone. All 12 two-tailed t-tests (4 preceding tones * 3 positions) showed no statistical differences at the 95% confidence level. Also, both *tou* /LH/ and *tou* /Ø/ showed almost no creaky voice as shown in Table 4.6 above, which was another similarity between the two syllables.

Table 4.8 Duration comparison between the neutral-tone -*tou* and the rising-tone -*tou*

	S1 duration (ms)	S2 duration (ms)	S2/S1 ratio	avg. S2/S1 ratio
tou /Ø/	151.6	113.0	0.745	0.761
tou /LH/	154.1	122.4	0.794	0.797

The comparison of the durations in Table 4.8 shows that the average duration of -*tou* /Ø/ (M = 113.0, SD = 24.3) was slightly shorter than the duration of *tou* /LH/ (M = 122.4, SD = 28.0). However, the difference was not statistically significant [$t(92) = 1.74$, $p = 0.083$]. Their S1 durations were not significantly different either [$t(90) = 0.46$, $p = 0.65$]. Unsurprisingly, the average S2/S1 ratios were not different between *tou* /LH/ and *tou* /Ø/

[$t(91) = 1.10$, $p = 0.27$]. The full-full X-*tou* /LH/ also has a trochaic rhythm as found in X-*dou* /L/ above.

The comparisons between -*tou* /Ø/ and -*tou* /LH/ show that the pitch contours of these two syllables were very similar. Their durations were also very similar. It is very likely that the TM speakers produce the neutral-tone -*tou* /Ø/ with the rising tone, although the phonetic realizations of -*tou* /Ø/ and -*tou* /LH/ do not show obvious rise. Although -*tou* /Ø/ had similar duration with -*dou* /L/, the pitch target of -*dou* /L/ was significantly lower than -*tou* /Ø/.

4.2.3. Neutral-tone -*zi*

The neutral-tone *zi* was compared to the low tone *zi*. Figure 4.9 shows the pitch contours of X-*zi* /L/ varying three preceding tones. As shown in the figure, the pitch contours overcame the preceding tone and merged to -1.5 semitone z-score. One-way ANOVA shows that the influence of the preceding tones was statistically significant only at the early position [$F(2,11) = 5.98$, $p = 0.017 < .05$]. The influences of the preceding tones were not significant at the mid position [$F(2,11) = 1.21$, $p = 0.34$] or at the late position [$F(2,11) = 0.23$, $p = 0.80$].

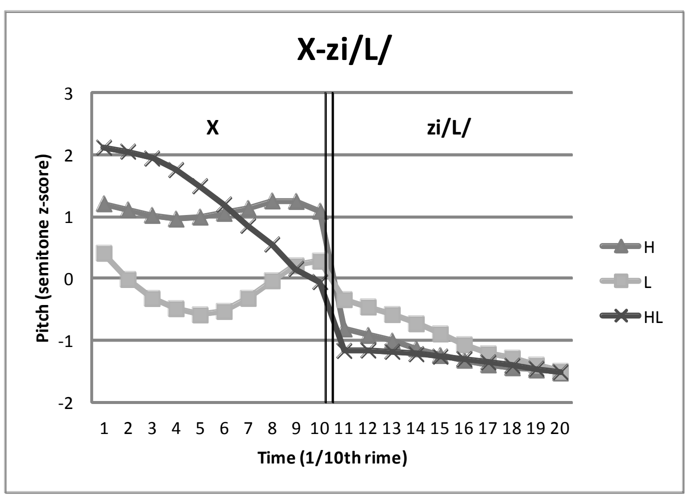


Figure 4.9 Average pitch contours of X-*zi* /L/

Figure 4.10 compares the pitch contours of X-*zi* /L/ and X-*zi* /Ø/. As shown in the figure, the pitch contours of *zi* /Ø/ (darker lines) were rather flat and stayed around 0 ~ -1 semitone z-score. Also, as discussed in the previous chapter, the pitch of zi /Ø/ was influenced by the preceding lexical tone. On the other hand, the pitch contours of *zi* /L/ (lighter lines) overcame the preceding tones and merged to -1.5 semitone z-score. The t-tests show that the differences between the average pitch of *zi* /Ø/ and *zi* /L/ were statistically significant at the early, mid, and late positions [Early: $t(28) = -4.24$, $p < .001$; Mid: $t(25) = -5.41$, $p < .001$; Late: $t(28) = -6.76$, $p < .001$]. The pitch contours of *zi* /L/ were statistically lower than the pitch contours of *zi* /Ø/. Even though *zi* /Ø/ after /L/ was lower than *zi* /Ø/ after /H, LH, HL/, the pitch of *zi* /Ø/ after /L/ ($M = -0.99$, $SD = 0.38$) at the late position was still higher than the pitch of *zi*/L/ ($M = -1.37$, $SD = 0.37$). T-test shows that the difference was statistically significant [$t(26) = -2.66$, $p = 0.013 < .05$].

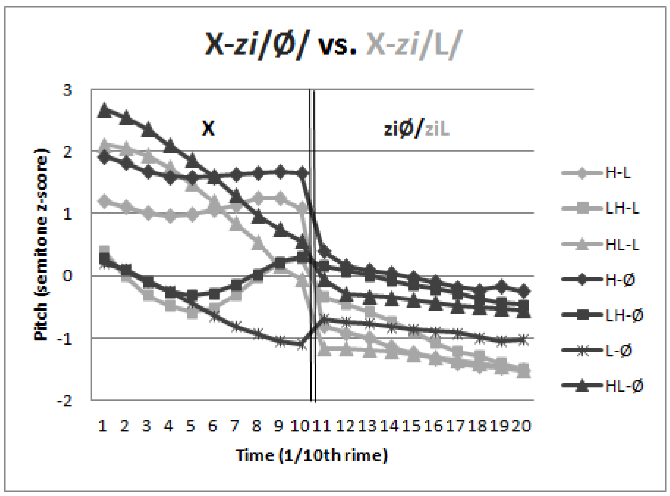


Figure 4.10 Average pitch contours of X-*zi* /L/ and X-*zi* /Ø/

The comparison of the numbers of the creaky syllables also shows that *zi* /L/ had a lower pitch target than *zi* /Ø/. As shown in Table 4.9, more than half of the *zi* /L/ syllables were excluded because of their creaky voice. On the other hand, fewer *zi* /Ø/ were excluded especially when the preceding tone was /H/ or /LH/. However, the numbers of *zi* /Ø/ with creaky voice were not low either. 25% of them were creaky after /H/ and /LH/ (3/12; 3/12), and more than half of them were creaky after /L/ and /HL/ (7/12; 8/12).

Table 4.9 Numbers of excluded syllables with creaky voice - *zi* /L/, *zi* /Ø/

Preceding tone / Syllable	H	LH	L	HL
zi /L/	7/12	6/12	-	9/12
zi /Ø/	3/12	3/12	7/12	8/12

As for the durations, Table 4.10 shows that the average duration of the rime of *zi* /Ø/ (M = 58.4, SD = 22.4) was shorter than the average duration of the rime of *zi* /L/ (M = 78.0, SD = 28.6). The difference was statistically significant [$t(64) = 3.41$, $p = 0.0011 < .01$]. Even though the S1 duration of *zi* /L/ was significantly longer than the S1 duration of *zi* /Ø/ [$t(76)=2.25$, $p = 0.028 < .05$], the difference between their average S2/S1 ratio were still statistically significant [$t(71) = 2.18$, $p = 0.032 < .05$]. The S2/S1 ratio of *zi* /Ø/ was 0.415 and the S2/S1 ratio of *zi* /L/ was 0.497.

Table 4.10 Duration comparison between neutral-tone -*zi* and full-tone -*zi*

	S1 duration (ms)	S2 duration (ms)	S2/S1 ratio	avg. S2/S1 ratio
zi /Ø/	144.8	58.4	0.403	0.415
zi /L/	157.7	78.0	0.495	0.497

The comparison of -*zi* /Ø/ and *zi* /L/ above shows that the pitch target of -*zi* /Ø/ was not as low as -*zi* /L/, and the pitch target of -*zi* /Ø/ was not implemented as fast as that of -*zi* /L/. Also, the duration of the rime of -*zi* /Ø/ was shorter than that of -*zi* /L/, and the differences between their S2/S1 ratios were also significant, suggesting a possible rhythmic difference between X-*zi* /Ø/ and X-*zi* /L/.

4.2.4. Neutral-tone syllables with /ə/ rime

In this section, -*zhe* /Ø/ 'DUR' is compared with the low-tone syllable *zhe* /L/ 'person'. The rest of the neutral-tone syllables with /ə/ rime: *le* /lə/ 'PERF', *de* /tə/ 'POSS; NML', *me* /mə/ 'such', and *ge* /kə/ 'CL' are compared with *she* /ʂə/ /L/ 'to let go' and *che* /$ʈʂ^hə$/ /L/ 'to pull'.

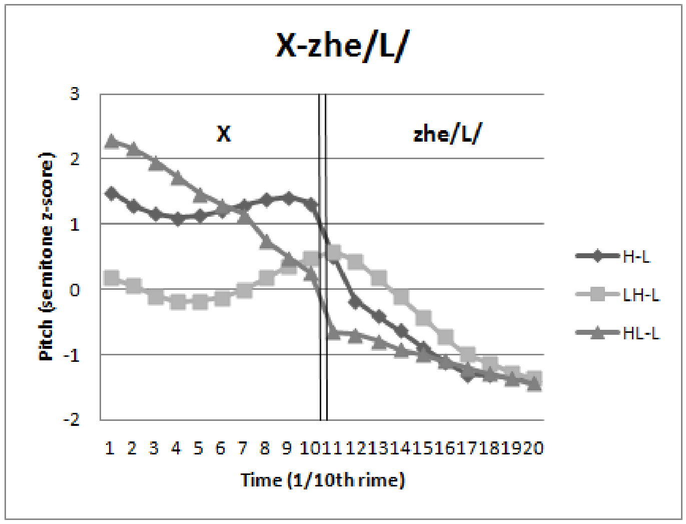


Figure 4.11 Average pitch contours of X-*zhe* /L/

Figure 4.11 shows the average pitch contours of X-*zhe* /L/ varying the three preceding tones. We can see that the three pitch contours of *zhe* /L/ merged to -1.5 semitone z-score in the end of the syllable. One-way ANOVA tests shows that the influence of the preceding tones was significant at the early position [$F(2,22) = 11.87$, $p < .001$]. The influence of the preceding tones lasted through the mid position [$F(2,22) = 4.95$, $p = 0.0168 < .05$], but the influence was not statistically significant at the late position [$F(2,22) = 0.40$, $p = 0.67$].

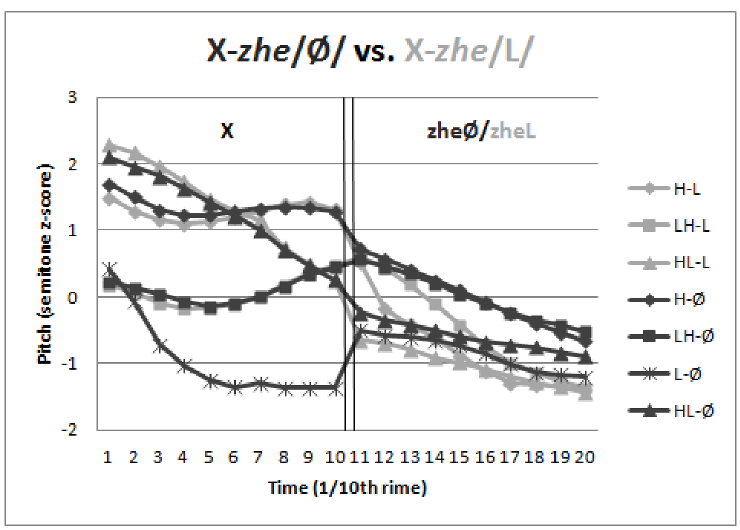


Figure 4.12 Pitch contours of X-*zhe* /L/ and X-*zhe* /Ø/

The pitch contours of *zhe* /L/ and *zhe* /Ø/ differed. As shown in Figure 4.12, the pitch contours of *zhe* /Ø/ ended between -0.5 ~ -1.2 semitone z-score while the pitch contours of *zhe* /L/ ended at -1.5 semitone z-score. The t-tests show that the average pitches of *zhe* /L/ and *zhe* /Ø/ significantly different at the mid and late position [Mid: $t(52) = -3.96$, $p < .001$; Late: $t(53) = -5.55$, $p < .001$].

Figure 4.12 also shows that the pitch contours of *zhe* /Ø/ varied depending on the preceding tones. At the late position, the pitch of *zhe* /Ø/ after /H, LH, HL/ was higher than the average pitch of *zhe* /L/ after /H, LH, HL/ respectively. T-test shows that the pitch differences at the late position were all significant [after /H/: $t(14) = -5.25$, $p < .001$; after /LH/: $t(18) = -3.68$, $p = 0.0017 < .01$; after /HL/: $t(14) = -3.12$, $p = 0.0075 < .01$]. However, when the preceding tone was /L/, the pitch contour of *zhe* /Ø/ was similar to the pitch contours of *zhe* /L/. At the late position, the difference between the average pitch of *zhe* /Ø/ after /L/ ($M = -1.14$, $SD = 0.40$) and the average pitch of *zhe* /L/ ($M = -1.23$, $SD = 0.46$) were not statistically significant [$t(11) = 0.52$, $p = 0.62$].

Also the numbers of syllables with creaky voice differed between *zhe* /L/ and *zhe* /Ø/. As shown in Table 4.11, *zhe* /L/ had more creaky syllables than *zhe* /Ø/. While about half of the *zhe* /L/ syllables had creaky voice and were excluded, fewer zhe /Ø/ syllables had creaky voice and were excluded. We can also see that *zhe* /Ø/ after /L/ and /HL/ have more creaky syllables than after /H/ and /LH/ (5/12; 4/12), which is parallel with the patterns of the pitch contours.

Table 4.11 Numbers of excluded syllables with creaky voice - *zhe* /L/, *zhe* /Ø/

Preceding tone / Syllable	H	LH	L	HL
zhe /L/	5/12	6/12	-	8/12
zhe /Ø/	2/12	2/12	5/12	4/12

As for the durations, Table 4.12 shows that the average duration of -*zhe* /Ø/ (M = 80.0, SD = 31.2) was shorter than *zhe* /L/ (M = 111.7, SD = 30.4). A t-test shows that the differences was statistically significant [$t(77) = 4.69$, $p < .001$]. As their S1 durations were not significantly different [$t(76) = 1.09$, $p = 0.28$], their average S2/S1 ratios were also significantly different [$t(82) = 4.11$, $p < .001$] (0.580 vs. 0.759).

Table 4.12 Duration comparison between neutral-tone -*zhe* and full-tone -*zhe*

	S1 duration (ms)	S2 duration (ms)	S2/S1 ratio	avg. S2/S1 ratio
zhe /Ø/	140.7	80.0	0.569	0.580
zhe /L/	147.9	111.7	0.755	0.759

The comparison between *zhe* /Ø/ and *zhe* /L/ shows that the two syllables had different pitch targets, and that the pitch target of *zhe* /L/ was implemented faster than that of *zhe* /Ø/. Also, the duration of *zhe* /Ø/ was shorter than the duration of *zhe* /L/. The rhythmic differences were also statistically significant.

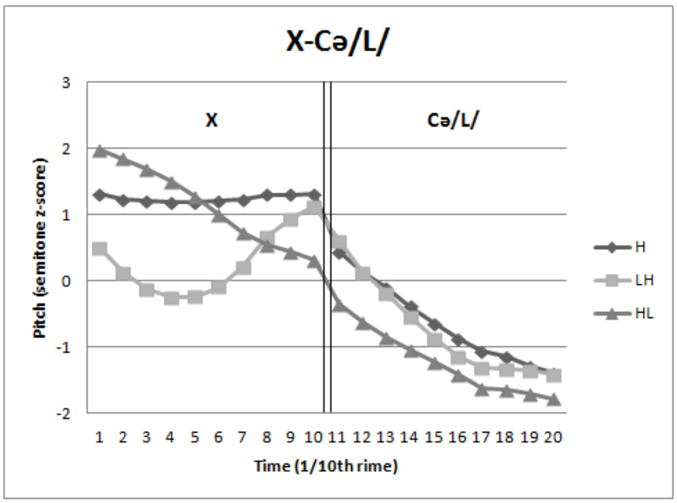


Figure 4.13 Pitch contours of X-Cə /L/

Due to the syllable gap in Mandarin, the rest of the neutral tone syllables with /ə/ rime were compared with the low tone *che* and *she* (Cə /L/). The average pitch contours of Cə /L/ varying the preceding tones are shown in Figure 4.13. Like other low-tone syllables described above, the pitch contours of Cə /L/ all lowered to less than -1.5 semitone z-score. However, the pitch contours of Cə /L/ did not merge in the second half of the syllable like other low-tone syllables. The pitches were influenced by the preceding tones. Two-way ANOVAs show that there was a main effect of the preceding tones at all three positions [Early: $F(2,33) = 10.58$, $p < .001$; Mid: $F(2,33) = 9.37$, $p < .001$; Late: $F(2,33) = 5.79$, $p = 0.007 < .01$]. The main effect of syllable (*che* or *she*) was not significant at any of the three positions [Early: $F(1,33) = 0.21$, $p = 0.65$; Mid: $F(2,33) =$

0.52, p = 0.48; Late: $F(1,33)$ = 2.18, p = 0.15]. The interaction effect was also not significant at any of the three positions [Early: $F(2,33)$ = 0.16, p = 0.85; Mid: $F(2,33)$ = 2.41, p = 0.11; Late: $F(2,33)$ = 0.02, p = 0.97]. The statistical analysis shows that the pitches of Cə /L/ were influenced by the preceding lexical tones. However, there were no influences of the syllables (*che* or *she*) or any interaction effect.

Figure 4.14 compares the pitch contours of X-Cə /L/ (*che* and *she*) with the pitch contours of X-Cə /Ø/ (*de*, *le*, *me*, *ge*). As shown in the figure, the average pitch contours of Cə /Ø/ (darker lines) were higher than the average pitch contours of Cə /L/ in the second half of the syllables (lighter lines). At the mid position, the pitch of Cə /L/ (M = -0.95, SD = 0.43) was lower than the pitch of Cə /Ø/ (M = -0.53, SD = 0.49), the difference was statistically significant [$t(70)$ = 5.28, p < .001]. At the late position, the pitch of Cə /L/ (M = -1.40, SD = 0.43) was even lower compared to the pitch of Cə /Ø/ (M = -0.80, SD = 0.45). The difference was also statistically significant [$t(64)$ = 7.63, p < .001]. Even though the pitch of Cə /Ø/ after /L/ was lower than the pitch of Cə /Ø/ after /H, LH, HL/, the pitch of Cə /Ø/ after /L/ (M = -1.16, SD = 0.43) was still higher than the pitch of Cə /L/ (M = -1.40, SD = 0.43) at the late position. The differences were marginally significant [$t(68)$ = 2.36, p = 0.021 < .05].

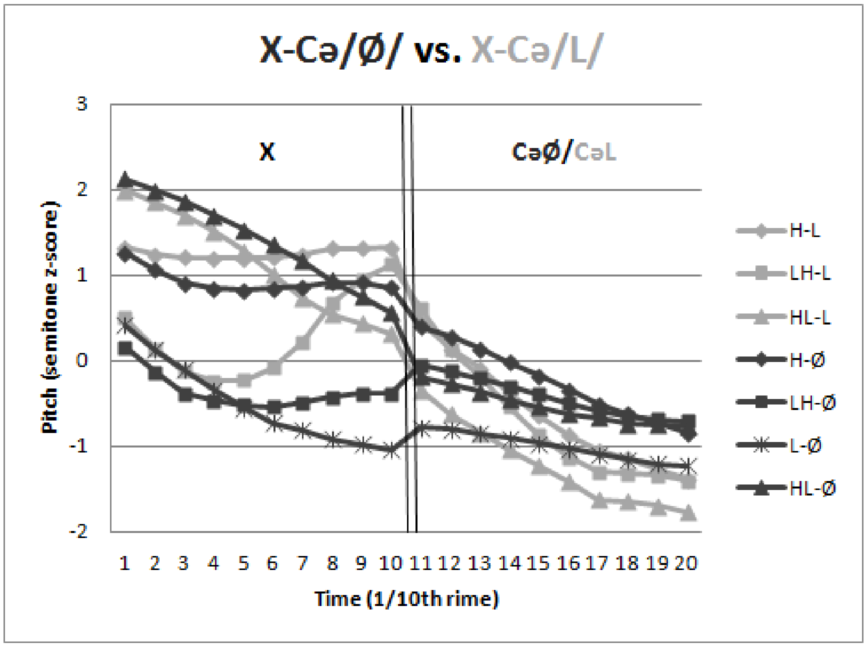


Figure 4.14 Pitch contours of X-Cə /L/ and X-Cə /Ø/

As for the syllables that had creaky voice and were excluded, Cə /L/ also had higher percentage than Cə /Ø/. Table 4.13 shows the number of syllables that were excluded due to creaky voice. In the bottom of the table, the percentages were also calculated in terms of Cə /L/ and Cə /Ø/. *me* /Ø/ was not listed because many of the *me* /Ø/ showed different pitch contours as mentioned in the previous chapter. The data here shows that nearly half of the Cə /L/ had creaky voice and were excluded. On the other hand, under 20% of the Cə /Ø/ after /H, LH/ had creaky voice and around 30% of the Cə /Ø/ after /L, HL/ had creaky voice.

Table 4.13 Numbers of excluded syllables with creaky voice - *Cə* /L/, *Cə* /Ø/

Preceding tone / Syllable	H	LH	L	HL
che /L/	7/12	5/12	-	4/12
she /L/	6/12	6/12	-	5/12
de /Ø/	1/12	1/12	6/12	5/12
ge /Ø/	3/12	2/12	2/12	4/12
le /Ø/	1/12	3/12	2/12	2/12
Cə /L/	54.2%	45.8%	-	37.5%
Cə /Ø/	13.9%	16.7%	27.8%	30.6%

The durations of the rimes of the neutral-tone *de*, *le*, *me*, *ge* were compared with the rimes of the low-tone *che* and *she*. As shown in Table 4.14, the average duration of *de* /Ø/ (M = 80.0, SD = 29.5) was shorter than the average duration of Cə /L/ (M = 111.7, SD = 23.8). The difference was statistically significant [$t(86) = 6.22$, $p < .001$]. The comparison of the S2 durations also shows that the rimes of Cə /L/ was significantly longer than the rimes of *le* (M = 86.3, SD = 21.1), *me* (M = 92.8, SD = 24.9), and *ge* (M = 81.3, SD = 31.7) [Cə /L/ vs. *le*: $t(109) = 6.14$, $p < .001$; vs. *me*: $t(55) = 3.57$, $p < .001$; vs. *ge*: $t(75) = 5.537$, $p < .001$].

However, the results were quite different when the duration of the preceding syllable (S1) is taken into account. As shown in Table 4.14, the average S2/S1 ratio of Cə /L/ was 0.851. Only *de* /Ø/ ($M = 0.576$) and *le* /Ø/ ($M = 0.694$) had significantly lower S2/S1 ratio [vs. *de* /Ø/: $t(118) = 4.14$, $p < .001$; vs. *le* /Ø/: $t(115) = 2.54$, $p = 0.0124 < .05$]. However, the average S2/S1 ratio of *ge* /Ø/ ($M = 0.818$) was not significantly different from Cə /L/

[*t*(103) = 0.43, *p* = 0.67]. The average S2/S1 ratio of *me* /Ø/ (*M* = 1.016) was even marginally longer than Cə /L/ [*t*(60) = -1.86, *p* = 0.068 < 0.1].

Table 4.14 Duration comparison between neutral-tone -*Ce* and full-tone -*Ce*

	S1 duration (ms)	S2 duration (ms)	S2/S1 ratio	avg. S2/S1 ratio
de /Ø/	148.7	80.0	0.538	0.576
le /Ø/	132.3	86.3	0.652	0.694
me /Ø/	99.0	92.8	0.938	1.016
ge /Ø/	108.6	81.3	0.753	0.818
Cə /L/	146.0	111.7	0.765	0.851

The reason why *me* /Ø/ and *ge* /Ø/ had higher S2/S1 ratios was probably because their S1 rimes were shorter (*M* = 99.0 & 108.6 ms). The testing disyllabic words of *me* /Ø/ were *duo-me* /H-Ø/, *she-me* /LH-Ø/, *ze-me* /L-Ø/, and *na-me* /HL-Ø/ with S1 as [twɔ, ʂɤ, tsɤ, nɑ] respectively. The testing disyllabic words of *ge* /Ø/ were *ba-ge* /H-Ø/, *yi-ge* /LH-Ø/, *wu-ge* /L-Ø/, *si-ge* /HL-Ø/ with S1 as [pɑ, i, u, sɨ] respectively. Almost all of these S1 have monophthongs, which are inherently shorter than the rimes with final nasals. As a result, the S2/S1 ratios of *me* /Ø/ and *ge* /Ø/ were higher. On the other hand, the testing words of Cə /L/ were *la-che* /H-Ø/, *xian-che* /LH-Ø/, *luan-che* /HL-Ø/, *shi-she* /H-Ø/, *nan-she* /LH-Ø/, and *bu-she* /HL-Ø/. Their S1 were [lɑ, ɕjan, lwan, ʂɨ, nan, pu]. Half of them have final nasals, which makes the rimes longer. As a result, the average duration of S1 of Cə /L/ was much longer (146.0 ms).

The comparison of the neutral-tone and the low-tone syllables with a /ə/ rime shows that the pitch of the low tone was much lower than that of the neutral-tone. Also, the low-tone /ə/ syllables had longer duration than the neutral-tone /ə/ syllables. However, the rhythmic differences of the disyllabic X-Cə were not always significant. Only *de* /Ø/ and *le* /Ø/ had significantly smaller S2/S1 ratios; the S2/S1 ratios of *me* /Ø/ and *ge* /Ø/ were not significantly smaller than Cə /L/.

4.2.5. Reduplications

The pitch contours of reduplications were compared to several words to see whether the reduplicated neutral-tone syllable has a low tone. The reduplicated *shu* /Ø/ in *shu-shu* /LH-Ø/ 'uncle' was compared to *shu* /L/ in *fan-shu* /H-L/ 'sweet potatoes'; *shen* /Ø/ in *shen-shen* /L-Ø/ 'aunt' was compared to *shen* /L/ in *ping-shen* /LH-L/ 'reviewer'; and *di* /Ø/ in *di-di* /HL-Ø/ was compared to *di* /L/ in *di-di* /HL-L/ 'underground'. The reduplicated *ma* /Ø/ was not included because it had a very obvious high tone as mentioned in the production chapter.

Figure 4.15 shows the pitch contours of reduplication and their low-tone counterparts. The three corresponding low-tone syllables shared rather similar pitch contours (lighter lines). The pitch contours lowered to less than -1 semitone z-score in the second half of the syllables. A one-way ANOVA shows that the pitches of the three low-tone syllables were significantly different at the early position [$F(2,22) = 4.16$, $p = 0.02933 < .05$]. However, the pitch differences were not significant at the mid position [$F(2,22) = 0.09$, $p = 0.91$] and the late position [$F(2,22) = 0.28$, $p = 0.76$].

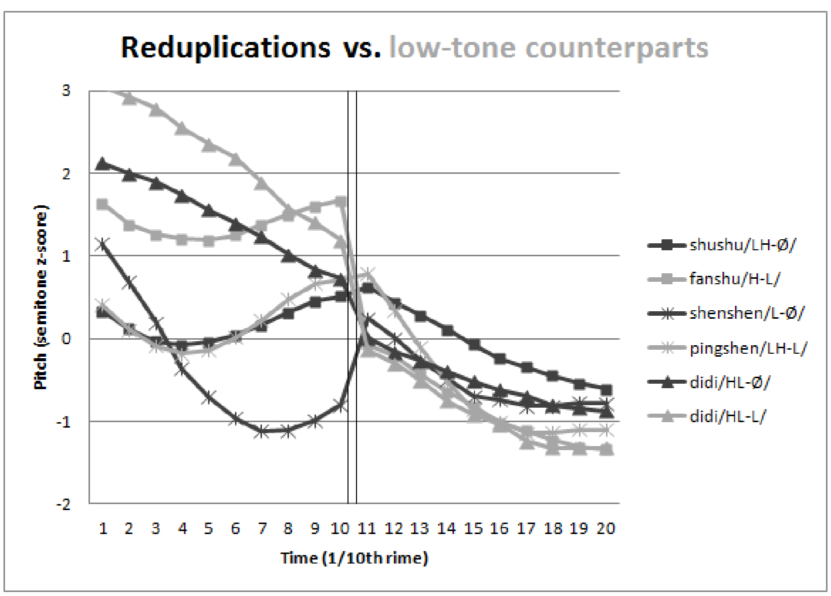


Figure 4.15 Pitch contours of reduplications and their full-tone counterparts

Figure 4.15 also shows the pitch contours of the reduplications. The pitch contours of the neutral-tone reduplicated syllables were generally higher. At the late position, the pitch of *shu* /Ø/ (M = -0.44, SD = 0.31) was higher than the pitch of *shu* /L/ (M = -1.23, SD = 0.37). The difference was statistically significant [$t(16) = -5.04$, $p < .001$]. However, even though the pitch of the neutral-tone *di* and *shen* (*di* /Ø/: M = -0.81, SD = 0.64; *shen* /Ø/: M = -0.82, SD = 0.65) were higher than the pitch of the low-tone *di* and *shen* (*di* /L/: M = -1.32, SD = 0.78; *shen* /Ø/: M = -1.13, SD = 0.25), these differences were not statistically significant [*di*: $t(13) = -1.52$, $p = 0.15$; *shen*: $t(11) = -1.34$, $p = 0.21$].

Also, the corresponding low-tone syllables were more likely to have creaky voice. As shown in Table 4.15, about 25% of the low-tone syllables had creaky voice and were excluded from the average, while nearly no neutral-tone reduplicated syllables had creaky voice.

Table 4.15 Numbers of excluded syllables with creaky voice – *Comparison between the reduplications with their corresponding low tone syllables*

Syllable / Tone	*shu*	*shen*	*di*
/L/	3/12	3/12	4/12
/Ø/	0/12	0/12	1/12

As for the duration comparison, the results of the reduplicants and their counterparts are shown in Table 4.16. The reduplicated *ma* /Ø/ was compared to the high tone *ma* ‘mom’. The average durations of *ma* /Ø/ was 144.2 ms (SD = 30.3) and the average durations of *ma* /H/ was 152.2 ms (SD = 25.8). However, paired t-test showed the difference was not statistically significant [$t(11) = 0.86$, $p = 0.41$]. Their average S2/S1 ratios were also similar (1.215 vs. 1.156). The difference was not statistically significant [$t(11) = -0.82$, $p = 0.43$].

Table 4.16 Duration comparison between the reduplications and their corresponding full-tone words

Comparison	Disyllabic words	S1 duration (ms)	S2 duration (ms)	S2/S1 ratio	avg. S2/S1 ratio
ma	*ma-ma* /H-Ø/	119.2	144.2	1.210	1.215
	wo-ma /L-H/	134.8	152.2	1.129	1.156
shu	*shu-shu* /LH-Ø/	114.1	91.7	0.804	0.808
	fan-shu /H-L/	170.6	87.3	0.512	0.512
shen	*shen-shen* /L-Ø/	139.2	146.1	1.050	1.052
	ping-shen /LH-L/	145.3	150.9	1.038	1.066
di	*di-di* /HL-Ø/	99.6	96.4	0.968	0.966
	di-di /HL-L/	116.8	102.2	0.875	0.878

The reduplicated *shu* /Ø/ in *shu-shu* 'uncle' is compared with *shu* /L/ in *fan-shu* 'sweet potatoes'. The average duration of *shu* /Ø/ (M = 91.7, SD = 20.3) was similar to the average duration of *shu* /L/ (M = 87.3, SD = 27.1). Paired t-test shows that the difference between the two was not significant [$t(11) = -0.60$, $p = 0.56$]. However, the comparison of the S2/S1 ratios shows that the S2/S1 ratio of *shu-shu* /LH-Ø/ (0.808) was much bigger than the S2/S1 ratio of *fan-shu* /H-L/ (0.512). Paired t-test shows that the difference was statistically significant [$t(11) = -5.87$, $p < .001$]. This is opposite to what was expected because a strong-weak disyllabic word is supposed to have a lower S2/S1 ratio. In this case the strong-strong disyllabic word had a lower S2/S1 ratio. This is probably because the *fan* (S1) in *fan-shu* had a final nasal. The duration of *fan* (S1) in *fan-shu* was therefore much longer than that of *shu* (S1) in *shu-shu* (170.6 ms vs 114.1 ms), resulting in a lower S2/S1 ratio.

The reduplicated *shen* /Ø/ in *shen-shen* /L-Ø/ 'aunt' was compared with *shen* /L/ in *ping-shen* /LH-L/ 'reviewer'. The result shows that the average duration of *shen* /Ø/ (M = 146.1, SD = 28.2) and *shen* /L/ (M = 150.9, SD = 39.3) were similar. Paired t-test shows that the difference was not significant [$t(11) = 0.51$, $p = 0.62$]. There was also no significant difference between their average S2/S1 ratios [$t(11) = 0.15$, $p = 0.88$]. This result further suggests that the difference between the S2/S1 ratios of *fan-shu* and *shu-shu* above is due to the S1 duration difference. As the S1 of *ping-shen* and *shen-shen* here both have VN rimes, their S1 durations were similar (139.2 ms vs. 145.3 ms); and their S2/S1 ratios were similar too.

Luckily, it was possible to compare the reduplication *di-di* /HL-Ø/ 'brother' with its minimal-pair *di-di* /HL-L/ 'underground'. The result shows that the second syllable of *di-di* /HL-L/ (M = 102.2, SD = 20.4) was slightly longer than the second syllable of *di-di* /HL-Ø/ (M = 96.4, SD = 33.9). However, this difference was not statistically significant [$t(11) = 0.82$, $p = 0.43$]. Their average S2/S1 ratios also did not show significant difference [$t(11) = -1.32$, $p = 0.21$].

The comparison of the neutral-tone reduplicated syllables and their counterparts shows that except for the reduplicated high tone syllable, other reduplicated syllables have a low pitch target. Statistical analysis shows that *shu* /Ø/ in *shu-shu* /LH-Ø/ was significantly higher than its low-tone corresponding syllable *shu* /L/. However, *shen* /Ø/ in *shen-shen* /L-Ø/ and *di* /Ø/ in *di-di* /HL-Ø/ were not significantly higher than their low-tone corresponding syllables *shen* /L/ and *di* /L/. This result is parallel with the finding of *men* /Ø/ in which the preceding tones play an important role. The neutral tone after /H, LH/ had higher pitch than the neutral tone after /L, HL/. As a result, the neutral tone was significantly higher than the low tone when following /H, LH/, but the difference between the low tone and the neutral tone after /L, HL/ was not significant. As for the duration, the results show that the durations of the reduplicated syllables were similar to their low-tone corresponding syllables. The rhythms of the reduplications were also similar to the rhythms of their corresponding disyllabic words.

4.3. Discussion

In this chapter, the neutral-tone syllables were compared with similar lexical-tone syllables. All of the tested neutral-tone syllables were compared with low-tone corresponding syllables that share similar segments/rimes. Neutral-tone *-men* /Ø/ and *-tou* /Ø/ were also compared with the rising-tone corresponding syllables.

Table 4.17 Summary of the results of the comparisons

Comparisons	Pitch contours	S2 durations	S1/S1 ratios
-tou /Ø/ vs. *-tou* /LH/	Similar	Similar	Similar
-tou /Ø/ vs. *-dou* /L/	Different	Similar	Similar
-men /Ø/ vs. *-men* /LH/	Different	Different	Similar
-men /Ø/ vs. *-ben* /L/	Different	Similar	Similar
-zi /Ø/ vs. *-zi* /L/	Different	Different	Different
-Cə /Ø/ vs. -Cə /L/	Different	Different	Different
-zhe /Ø/ vs. *-zhe* /L/	Different	Different	Different
Reduplications	Different	Similar	Similar

The comparison between the neutral tone and the rising-tone corresponding syllables showed that the neutral-tone syllable *-tou* /Ø/ did not de-stress in Taiwan Mandarin. The prescriptive neutral-tone *-tou* /Ø/ was acoustically very similar to the rising tone *tou* /LH/ in terms of their pitch contours and their durations. On the other hand, the neutral-tone *-men* /Ø/ had different pitch contours from *-men* /LH/ and the durations of *-men* /Ø/ were shorter than *men* /LH/, which demonstrates that *-men* /Ø/ and *men* /LH/ are not the same prosodically.

The rest of the neutral-tone syllables were compared with their low-tone corresponding syllables. All of these low-tone syllables share some features with each other. First, the influences of the preceding tones were overcome faster than in the neutral tone. The influences of the preceding tones were not found at the mid and the late position in *ben* /L/, *zi* /L/, or the reduplication counterparts; and the influences of the preceding tones were overcome in the late position in *zhe* /L/. However, *che* /L/ and *she* /L/ seemed to be influenced by the preceding tones even at the late position, which is not what I expected. Second, the pitch target of the low tone is lower than the neutral tone. The end pitches of the neutral-tone syllables generally fell between -0.5 ~ -1 semitone z-score, but all of these low-tone syllables lowered their pitch to under -1 semitone z-score. In fact, except for the syllables with final nasals such as *-ben* /L/and *-shen* /L/, where the F0 was raised by the final nasals, the rest of the low-tone syllables all reached to almost -1.5 semitone z-score. Lastly, the low-tone syllables have higher numbers of creaky voice which interferes with F0 readings. 58% of the *zi* /L/ syllables were excluded because of the creaky voice. 42% ~ 52% of the low-tone syllables with /ə/ rime and -*ben* /L/ were excluded. Other low-tone syllables also have at least 25% of their syllables excluded due to creaky voice.

As for the neutral-tone syllables, the previous production chapter showed that although the neutral tone in Taiwan Mandarin had its own pitch target, the influences of the preceding lexical tones were overcome slower than in the low tone. The influences of the preceding tones were overcome at the late position only in -*de*, -*ge*, and three of the four reduplicated items, but not in the other seven neutral-tone syllables. This difference between the neutral tone and the low tone might be explained by Xu's study (1997) on carry-over tonal influence. The study found that in Standard Mandarin, while the carry-over effect is found in all the lexical tones, the influence of the end pitch of the preceding lexical tone lasted until the end of Tone 1 /H/ and Tone 2 /LH/, but was not found in the end of Tone 3 /L/ and Tone 4 /HL/. The study claimed that a high tonal target is more susceptible to contextual influences than a low tonal target, probably because a low tonal target is more difficult to reach due to phonatory constraints. It is possible that the neutral tone in TM had a stronger influence from the preceding lexical tone because its pitch target is higher than the low tone.

In this chapter, the comparison shows that the pitch target of the neutral tone in TM was not as low as the low tone. At the late position, the pitch of *zi* /Ø/ and Cə /Ø/ were all statistically higher than the pitch of *zi* /L/ and Cə /L/ respectively. The results also show that at the late position, the pitch of *men* /Ø/ after /H, LH/ was statistically higher than *men* /L/; *zhe* /Ø/ after /H, LH, HL/ was higher than *zhe* /L/; and the neutral-tone reduplicant after /LH/ was higher than its corresponding low-tone syllable.

Even though generally the pitch of the neutral tone is higher than the low tone, the pitch contours of the two tones seem to overlap sometimes. Many of the pitches of the neutral tone in the late position were not significantly different from the pitch of the low tone in the late position, especially when their preceding lexical tone had a low offset. The results show that in the late position, the pitch of *men* /Ø/ after /HL, L/ was not significantly different from *ben* /L/; reduplicated syllable after /HL, L/ was not significantly different from their low-tone corresponding syllables; and *zhe* /Ø/ after /L/ was not significantly different from *zhe* /L/. The number of syllables with creaky voice quality (and thus excluded) also shows similar trends. While the low-tone syllables generally have a higher percentage of creaky voice than the neutral-tone syllables, the neutral tone after /HL, L/ was found with more creaky-voice syllables than the neutral tone after /LH, H/. For example, half of the instances of *men* /Ø/ after /L/ had creaky voice and were excluded (6/12); more than half of the instances of *zi* /Ø/ and *zhe* /Ø/ after /L, HL/ had creaky voice and were excluded (*zi* /Ø/: 7/12 when after /L/, 8/12 when after /HL/; *zhe* /Ø/: 7/12 when after /L/, 8/12 when after /HL/).

Overall, the comparison of the pitch contours between the neutral tone and the low tone in TM shows that the low tone has a lower pitch and higher chance of having creaky voice. However, some neutral-tone syllables have pitch contours as low as the low tone and creaky voice, especially when they are after /HL/ or /L/. This can explain why some speakers might treat the neutral tone as a low tone, as discovered in the previous chapter. Also, it remains unclear whether the pitch differences were perceptible by the listeners.

As for the comparison of the duration and the S2/S1 ratio, the results show different patterns depending on the tested neutral-tone syllables. *men* /Ø/ and the neutral-tone reduplicants were not statistically different from their low-tone corresponding syllables, and the comparison of S2/S1 ratios shows no rhythmic differences either. On the other hand, the durations of the rest of the neutral-tone syllables: *zi*, *zhe*, *de*, *le*, *ge*, and *me* were all significantly shorter than their low-tone corresponding syllables. The S2/S1 ratios of these neutral-tone syllables were also significantly lower than their low-tone counterparts except for *me* and *ge*, which was likely because their S1 all had shorter rimes.

However, even though many neutral-tone syllables in TM were significantly shorter than their low-tone corresponding syllables, the duration difference was small. *zi* /Ø/ was 20 ms shorter than *zi* /L/ (58.4 vs. 78.0 ms); and the durations of Cə /Ø/ and *zhe* /Ø/ (80 ~ 92 ms) were 20 to 30 ms shorter than the duration of Cə /L/ and *zhe* /L/ (111 ms). In terms of S2/S1 ratios, X-*zi* /Ø/ was 0.415 and X-*zi* /L/ was 0.497; and X-*zhe* /Ø/ was 0.612 and X-*zhe* /L/ was 0.769. The S2/S1 ratios of X-Cə /Ø/ were between 0.576~ 1.016, while the S2/S1 ratio of X-Cə /L/ was 0.851. The differences of the S2/S1 ratios between the pairs were generally around 0.1~ 0.2. In contrast, in a study done by Lin and Yan (1990), they compared the tonal minimal pair of full-full and full-neutral in Standard Mandarin. They found that when the full-tone S2 was /HL/, the average duration of S2 rime was 214 ms, which is almost twice as long as the average duration of the neutral-tone S2 rime (100 ms). The S2/S1 ratio of full-full was 0.89 and the S2/S1 ratio of full-neutral was 0.47. The differences between these two S2/S1 ratios were 0.42. Although they used a different lexical tone to compare, the result is still important for us because when a syllable is in sentences, /HL/ is the shortest tone in Standard Mandarin while /L/ is the shortest tone in Taiwan Mandarin (Deng et al. 2008). Lin and Yan's study and mine both compared the neutral tone with the shortest possible lexical tone, and the results indicate that the durational and rhythmical differences in SM are much larger than in TM.

The duration results also show that there is no general syllable reduction in the neutral-tone syllables in Taiwan Mandarin. The neutral-tone syllables that were significantly shorter all have the rimes [ə] or [ɨ]; these vowels are inherently short, and they are often the result of diachronic reduction. In contrast, the rest of the neutral-tone syllables (rimes [ən, u, i]) show no significant differences from their low-tone counterparts. If all neutral-tone syllables were reduced in Taiwan Mandarin, I would have found -*men* /Ø/ and the reduplicants (rimes [ən, u, i]) to be shorter than their counterparts too.

It is likely that neutral-tone syllables with rimes [ə] or [ɨ] are shorter than their low-tone counterparts because these vowels are intrinsically short. Therefore when other neutral-tone syllables were lengthened to be as long as the lexical-tone syllables, it was harder for these syllables to be lengthened. Another possible motivation for not lengthening is that these neutral-tone syllables with /ə/ or [ɨ] (*zi* 'DIM', *zhe* 'DUR', *de* 'POSS, NML', *le* 'PERF', *ge* 'CL', and *me* 'such') are all grammatical morphemes with weak semantic content, which are more susceptible to reduction. On the other hand, even though the plural suffix -*men* /Ø/ is productive, 1) it can only attach to humans, 2) it is only obligatory with pronouns but not with nouns, and 3) it is not compatible with numerical classifiers (Ramsey 1987, 64; Norman 1988, 159). It is possible that plural pronouns like *wo-men* '1.PL', *ni-men* '2.PL' and *ta-men* '3.PL' are lexicalized, and the morpheme boundaries are no longer apparent; and the -*men* suffix in words like *laoshi-men* 'teacher-PL' is used to emphasize the plurality. Therefore -*men* /Ø/ is lengthened like other neutral-tone syllables. As for the reduplications, this result shows that TM speakers do not reduce the second syllables of the reduplicants. It is likely that they treat all the reduplicated kinship terms as monomorphemic roots.

The results of this chapter suggest that the neutral-tone syllables in Taiwan Mandarin are not unstressed like in SM. In SM phonology, the stress decides the treatment of the syllable. An unstressed syllable undergoes vowel reduction and tone loss, therefore it possesses a neutral tone. However, as shown in Chapter Three, the pitch contours of the neutral tone in TM have a lexical-tone-like pitch target, therefore the neutral tone in TM is not toneless. Furthermore, this chapter shows that there is no general syllable reduction in Taiwan Mandarin. As a result, there is no strong evidence to argue that the neutral-tone syllables are unstressed and have a neutral tone. In TM, these syllables should be characterized as having a lexical tone. This chapter further shows that this lexical tone is not as low as the low tone. Therefore these “neutral-tone” syllables should be characterized as having a fifth tone with a mid-low pitch target.

The duration results suggest that those neutral-tone syllables with [ə] and [ɨ] in TM are probably in the process of changing from an unstressed syllable to a stressed syllable. These syllables have a lexical tone instead of being toneless. However, they are still shorter than other syllables with [ə] and [ɨ]. Nevertheless, the durational differences between these rimes and their corresponding full-tone syllables’ rimes were much smaller than in SM, which is another piece of evidence that TM is more syllable-timed than SM.

This chapter finds pitch and durational differences between the neutral tone and the low tone in Taiwan Mandarin, and the differences in many cases were statistically significant. However, whether the differences are perceptible by the TM listeners is another question, the next chapter conducts a perception test to see if the differences are perceived and to find out the possible perceptual cues.

CHAPTER 5. PERCEPTION TESTS

As shown in Chapter Four, some neutral-tone syllables share similar duration and pitch patterns with the corresponding lexical-tone syllables. However, some neutral-tone syllables were quite different from the lexical-tone syllables. It remains uncertain whether these differences are actually perceived by TM speakers. This section investigates whether the neutral-tone syllables are perceptually distinct from the lexical-tone syllables. There are two perception experiments in this study. Experiment 1 uses unaltered speech from four different TM speakers. Experiment 2 uses re-synthesized speech from only one TM speaker to further investigate the acoustic cues that are used to help the TM listeners to distinguish the neutral tone from the full tone.

5.1. Perception Experiment 1

I would like to explore whether the differences between the neutral-tone syllables and the lexical-tone syllables are perceptible by TM listeners. I hypothesize that TM listeners have difficulty distinguishing the neutral tone from the lexical tone, so that some of the differences between neutral-tone and lexical-tone syllables in TM are perceptible by TM listeners, but some of the differences are not.

In this experiment, TM listeners were asked to identify the word/phrase, choosing between the minimal pair of full-full and full-neutral. The tested tokens were produced by four different TM speakers and there were six tonal minimal pairs being tested. The six tonal minimal pairs contrast between *zi* (/Ø/ vs. /H/), *zi* (/Ø/ vs. /L/), *tou* (/Ø/ vs. /LH/), *men* (/Ø/ vs. /LH/), and two of them contrast between *zhe* (/Ø/ vs. /L/). Based on the results of Chapters Three and Four, I expect high accuracy rates for the pair contrasting *zi* (/Ø/ vs. /H/) because there is no post-L rising in TM; and I expect low accuracy rates for the pair contrasting *tou* (/Ø/ vs. /LH/) and. It is unclear whether the TM listeners can distinguish the rest of the pairs.

5.1.1. Subjects

The subjects were 40 Taiwan Mandarin speakers who are 18-55 years old and who have no speaking or hearing impairment. The subjects all grew up in Taiwan and their parents are all native speakers of Southern Min.

5.1.2. Task

Subjects were asked to listen to disyllabic words/phrases that were recorded previously by two male and two female TM speakers, and to identify what the words/phrases are. The word was played twice in a frame sentence *qing-shuo X X ba-ci* 'please say X X eight times'. Listeners were presented with the two options: the word with full-full tones and the word with full-neutral tones. For example, one of the two recordings of *she-tou* in the frame sentence was played twice: *qing-shuo she-tou ba-ci* 'Please say *she-tou* eight times'. The subjects had to choose between 舌頭 *she-tou* [LH-Ø] 'tongue', and 蛇頭 *she-tou* [LH-LH] 'snake head'. The two options were both written in Chinese characters without any phonetic spellings. Subjects were asked to circle the word/phrase that they heard.

5.1.3. Materials

There were only six tested minimal pairs (Table 5.1). These minimal pairs were chosen from a list of possible neutral tone minimal pairs (See Appendix E) after the pilot study. Because I would like to control the total experiment time of the perception test, I only chose one to two minimal pairs out of each syllable contrast.

Pair 1 and pair 2 contrast the diminutive suffix -*zi* [Ø] (full-neutral) and the full-tone words *zi* [H] and *zi* [L]. Pair 1 contrasts *yi-zi* /L-Ø/ with /L-H/. Based on the result from the production study, TM speakers do not have post-L rising. The neutral tone here is expected to have a low pitch. As a result, I expect pair 1 to be very distinct and both full-full and full-neutral words to have very high accuracy rate. Pair 2 contrast *jia-zi* /LH-Ø/ with /LH-L/. The neutral tone is expected to have a low pitch, so listeners might have difficulty distinguishing them.

Table 5.1 Minimal pairs for the perception test

	pinyin	IPA	full-neutral (FN)	full-full (FF)
1	*yi-zi*	[i-tsɨ]	椅子 [L-Ø] 'chair'	以茲 [L-H] 'hereby'
2	*jia-zi*	[tɕja-tsɨ]	夾子 [LH-Ø] 'clip'	甲子 [LH-L] '60 years'
3	*she-tou*	[ʂɤ-tʰou]	舌頭 [LH-Ø] 'tongue'	蛇頭 [LH-LH] 'snake head'
4	*gui-men*	[kue-mən]	鬼們 [L-Ø] 'ghosts'	鬼門 [L-LH] 'ghost gate'
5	*zuo-zhe*	[tswo-ʈʂə]	坐著 [HL-Ø] 'sitting'	作者 [HL-L] 'author'
6	*du-zhe*	[tu-ʈʂə]	讀著 [LH-Ø] 'reading'	讀者 [LH-L] 'reader'

Pair 3 shows the contrast between the fossilized diminutive suffix *-tou* [Ø] and the word *tou* [LH] 'head'. As shown in Chapter Four, the suffix *-tou* in TM usually retains the full tone [LH]. Therefore I expect listeners to have difficulty distinguishing the pair.

Pair 4 exhibits the contrast between a noun 'ghost-PL' [L-Ø] and a noun phrase 'ghost-gate' [L-LH]. The plural suffix *-men* in TM has a neutral tone with a low target. However, sometimes TM speakers pronounce the suffix *-men* with a full tone [LH] based on the phonetic element of the character[17]. If the speakers produce the suffix *-men* with a low-target pitch, I expect the listeners to distinguish the pair. On the other hand, if the speakers produced the suffix as [LH], I expect the listeners to perform poorly.

Pairs 5-6 contain the durative suffix *-zhe* [Ø]. The full-full patterns are all combined with *-zhe* [L] 'person' (a non-productive nominalizer suffix). Based on the production study in the previous chapters, both the full tone and the neutral tone *-zhe* have a low pitch target, but the pitch contours and pitch target are different. It is therefore important to know whether the distinction is perceptible.

There were also filler questions which were also tonal minimal pairs. These minimal pairs differ only in their lexical tones. For example, the listeners have to answer whether they hear *ling-qian* /LH-LH/ 'changes' or *ling-qian* /L-LH/ 'to withdraw money'. I expected the listeners to be able to answer the filler questions without any difficulty. The list of the filler minimal pairs can be found in Appendix F.

[17] It is possible that the speakers produce the /LH/ tone because the original lexical tone of the *-men* suffix is /LH/, but most likely the speakers do so because the right side of the character 們 *-men* /Ø/ 'plural suffix' (門) is pronounced *men*/LH/. Chinese speakers tend to pronounce the phonetic element of the character when they are not sure about the pronunciation.

Both the tested pairs and the filler pairs in the frame sentences were recorded by two male TM speakers and two female TM speakers. The TM speakers were four of the twelve speakers in the production test. The speakers were asked to read each sentence just once to get the most natural result. There were a total of 6 (minimal pairs) * 2 (items per pair) *4 (speakers) = 48 tested stimuli. The tested stimuli were mixed with 32 fillers that were produced by the same speakers. Therefore there were 80 sentences for the subjects to listen to and answer.

The stimuli were separated into two pools. Each tested minimal pair was separated into the two different pools. That is, one item from the pair went into each pool. Each pool had 24 tested stimuli and 16 fillers produced by the four speakers (6 tested stimuli and 4 fillers produced by each speaker). The first half of the 80 test sentences was drawn from one pool, and the second half was drawn from the other. Each listener heard all 80 sentences. The test sentences alternated in terms of speakers' gender to prevent the interference from the previous speaker, and the two items from the same minimal pairs did not follow right after each other. Other than these conditions, the test sentences were drawn from the pool randomly. The order of the 80 stimuli is attached in Appendix G.

5.1.4. Procedures

Each word was played in the frame sentences *qing-shuo X X ba-ci* 'please say X X eight times'. The recording included a 150 millisecond (ms) beep (220 Hz sine wave), followed by a 150 ms silence, the sentence, and then a 300 ms silence. The sentence was repeated again and after that there was a 3-second silence for the subjects to choose the answer. Every 20 questions, there was a short break. The short break consisted of a 1-second silence, followed by a 500 ms beep and a 500 ms silence. The beep and the silence were repeated three times. The following question resumed after a 2.5-second silence. There was a practice section before the test. 8 questions with fillers were included in the practice section. The 8 trial questions were voices from the 4 speakers in the test in order to familiarize the listeners with their voices. The entire test took about 10 minutes.

5.1.5. Result

The results are shown in Figure 5.1. The x-axis shows the six different pairs. For each pair, the left bar is the result of the full-full item and the right bar is full-neutral. The y-axis shows the percentage of the forms being correctly identified. The accuracy rates consist of the averages between the results from the four different speakers.

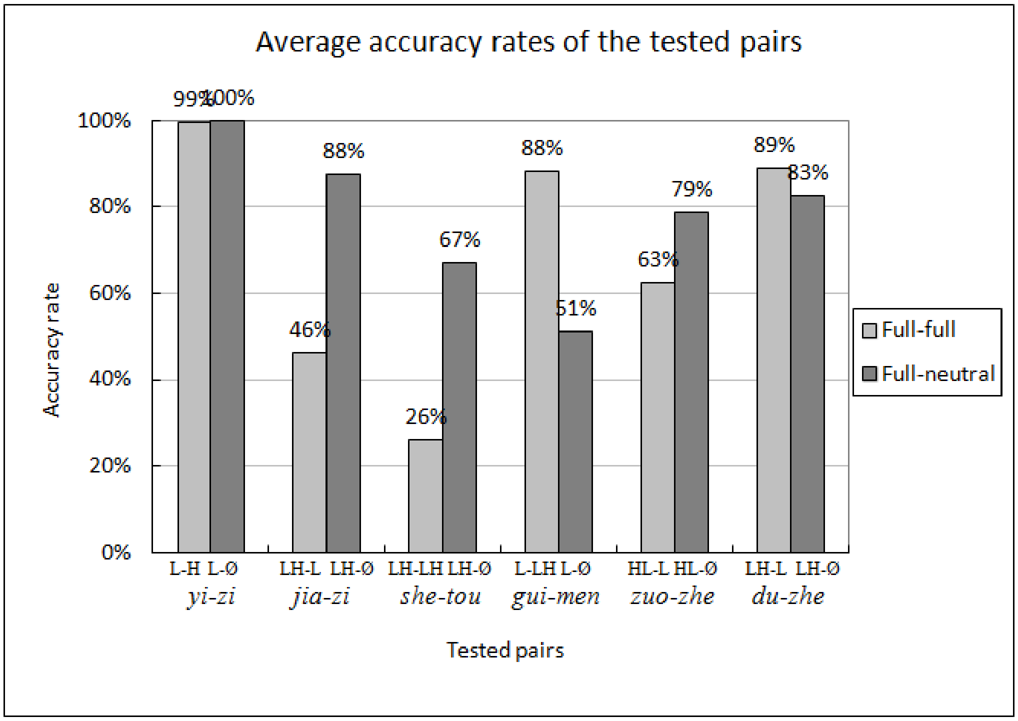


Figure 5.1 Average accuracy rates of the six tested pairs

The results show that the listeners generally have some difficulties differentiating the full-full and full-neutral tonal minimal pairs. Among the six pairs, the listeners accurately distinguished the *yi-zi* pair as expected (99% and 100%). The listeners also performed relatively well on the *zuo-zhe* and *du-zhe* pairs; the average rate of correctly identifying the full-full *zuo-zhe* (63%) is more than the chance level (n = 160, p = 0.002). On the other hand, the listeners had difficulty identifying the full-full in the *jia-zi* pair (46%) and the full-neutral in the *gui-men* pair (51%), both percentages are around the chance level (two-tail p = 0.3846, p = 0.8126). The listeners had the most difficulties with the *she-tou* pair, with the tendency to identify both of the pair as full-neutral (100-26=74% and 67%, p < 0.0001).

The results above are the averages of the four TM speakers. However, the average accuracy rates cannot reflect the variation across the different speakers. In the following section, the accuracy rates of the forms from each speaker are shown and further discussed. The details of the results and its implication are shown as follows.

5.1.5.1. **The *yi-zi* pair**

The *yi*-*zi* pair contrasts between the full-full [L-H] 'hereby' and full-neutral [L-Ø] 'chair'. In SM, the neutral tone -*zi* has a mid-high level pitch after a low tone *yi* (post-L rising), which would be hard to distinguish from *yi*-*zi* [L-H]. In this experiment, I expect the listeners to be able to distinguish the two because the neutral tone -*zi* in TM usually has a low tonal target.

Table 5.2 Percent correct responses for the *yi*-*zi* pair

	F2	F6	M3	M6	Avg.
yi-*zi* /L-H/ (F-F)	100%	100%	100%	98%	99%
yi-*zi* /L-Ø/ (F-N)	100%	100%	100%	100%	100%

The results in Table 5.2 show that the TM listeners could distinguish the *yi*-*zi* pair produced by all four speakers. As shown in Figure 5.2, the pitch contours of *zi* in the full-full (F-F) *yi*-*zi* were above 0 semitone z-score, while the pitch contours of *zi* in the full-neutral (F-N) *yi*-*zi* were below 0 semitone z-score or without pitch reading due to the creakiness such as M6. The overwhelmingly high accuracy rates suggest that the TM listeners were able to distinguish the neutral tone -*zi* from the high tone -*zi* likely because of the pitch height differences produced by the TM speakers. This result shows that the TM listeners do not expect post-L rising in TM.

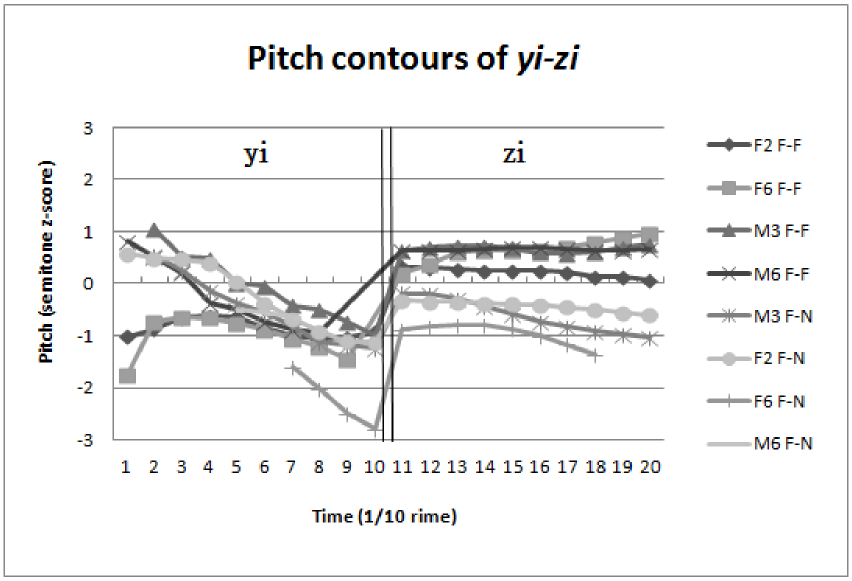


Figure 5.2 Pitch contours of the *yi-zi* pair by four TM speakers

5.1.5.2. **The *jia-zi* pair**

The *jia-zi* pair contrasts between the full-full [LH-L] ‘sixty years’ and the full-neutral [LH-Ø] ‘clip’. Since the neutral-tone *-zi* usually has a low pitch target in TM, I expect the TM listeners to have difficulties distinguishing the pair.

As shown in Table 5.3 and Figure 5.3, the TM listeners only distinguished the pairs produced by speaker F6 (Correct identification rates 73% and 70%). The TM listeners successfully identified the full-neutral *jia-zi* by all four speakers. As for the full-full pattern, F6’s full-full *jia-zi* is the only token that was identified correctly. F2’s and M6’s full-full *jia-zi* were both identified within the chance level, marked by the ** sign[18] in Table 5.3, while M3’s full-full jia-zi were identified most often as full-neutral. Since this test was a forced choice test, this result also means that the TM listeners tended to choose full-neutral.

[18] The sign ** signals that this percentage is statistically within the chance level. Based on the binomial test, only the accuracy rates that are equal to or more than 67.5% and the accuracy rates that are equal to or smaller than 32.5% are significantly different from 50% chances (n = 40; p = 0.0285 < 0.05). If the accuracy rate is within 35% ~ 65%, statistically we cannot claim that the listeners tend to choose one or the other.

Table 5.3 Percent correct responses for the *jia-zi* pair

	F2	F6	M3	M6	Avg.
jia-zi [LH-L] (F-F)	**35%	73%	28%	**50%	46%
jia-zi [LH-Ø] (F-N)	88%	70%	95%	98%	88%

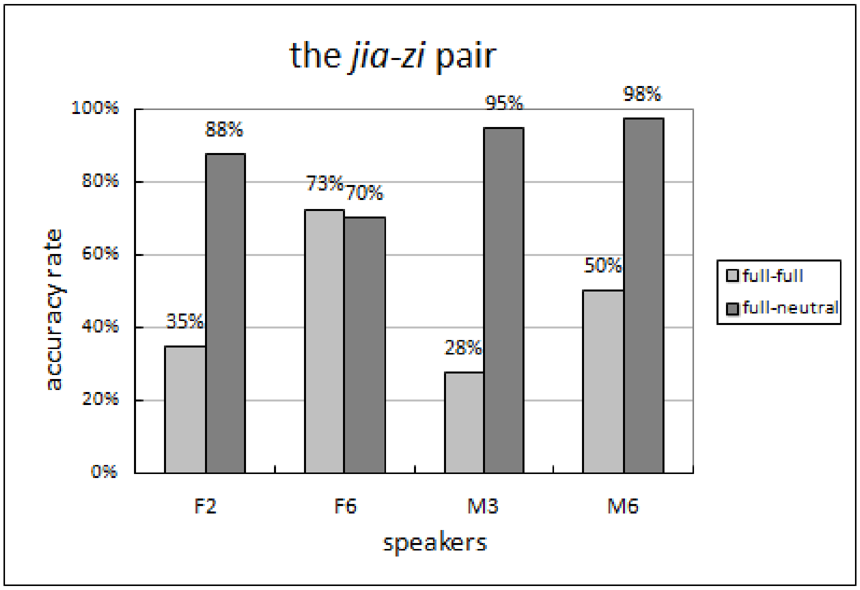


Figure 5.3 Percent correct responses for the *jia-zi* pair

The tendency to choose full-neutral is very likely related to the duration of the *zi* syllable. As shown in Figure 5.4, the S2/S1 ratio[19] on the x-axis represents the duration of *zi*, and the y-axis shows the percentage of the stimuli being identified as having a full-neutral form. The eight points in the chart are the distribution of the eight *jia-zi* stimuli (4 speakers*2 forms). As shown in the chart, F6's full-full form (F-F) has a much longer duration compared to other productions, and it is identified as having a full-full form (73% identified it as full-full). The rest of the stimuli had much shorter *zi* durations and tended to be identified as having a full-neutral form. This figure also shows that S2/S1 ratio is not the only acoustic cue used to identify neutral tones. When the S2/S1 ratios were around 0.2 ~ 0.5, the neutral-tone identification rate ranged between 50% ~ 95%. There seem to be other acoustic cues.

[19] S2/S1 ratio is the duration of the second syllable's rime divided by the duration of the first syllable's rime. S2/S1 ratio is used to represent duration in order to control for speed variation.

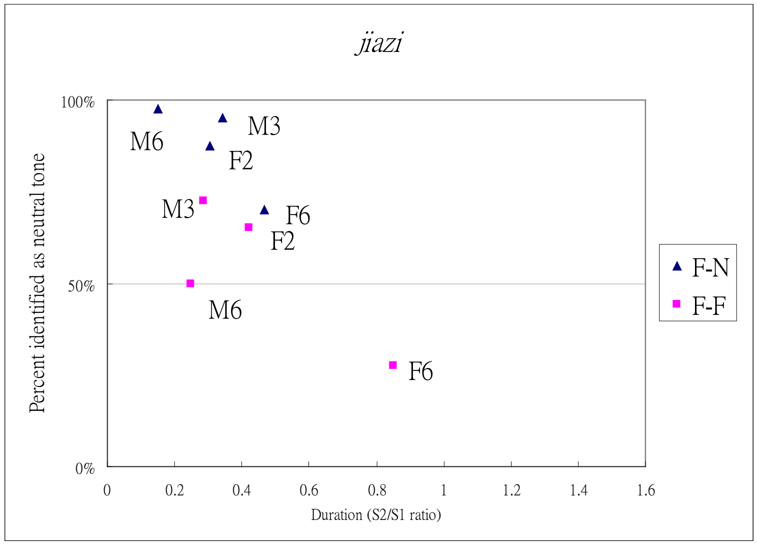


Figure 5.4 The *jia-zi* pair: relationship between duration and identification as neutral tone

Aside from the longer duration in F6's full-full, the voice quality of *zi* may also be an important acoustic cue. As shown in Figure 5.5, the second half of the *zi* syllable is slightly creaky, which is often used as an acoustic cue for the low tone in Mandarin.

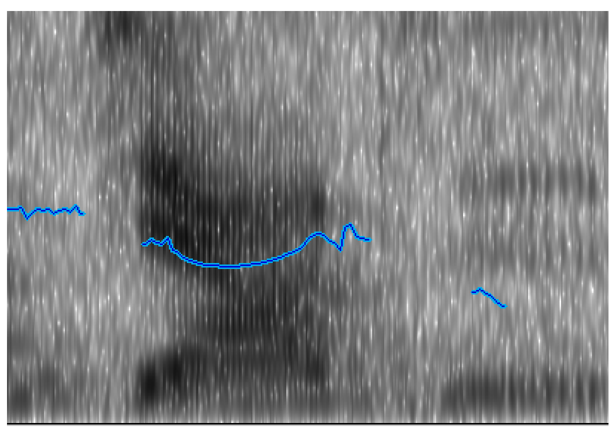


Figure 5.5 Spectrogram of the full-full *jia-zi* by speaker F6.

Although three out of four full-full items were not correctly identified, they were not identified as full-neutral with high percentages. As shown in Figure 5.4, even with similar durations, the full-neutral items (triangle symbols) have a higher full-neutral identification rate (70%~98%) than the full-full items (square symbols; 50%~73%). This shows that duration is not the only acoustic cue. There are other acoustic cues that help some of the listeners to distinguish the full-full from the full-neutral.

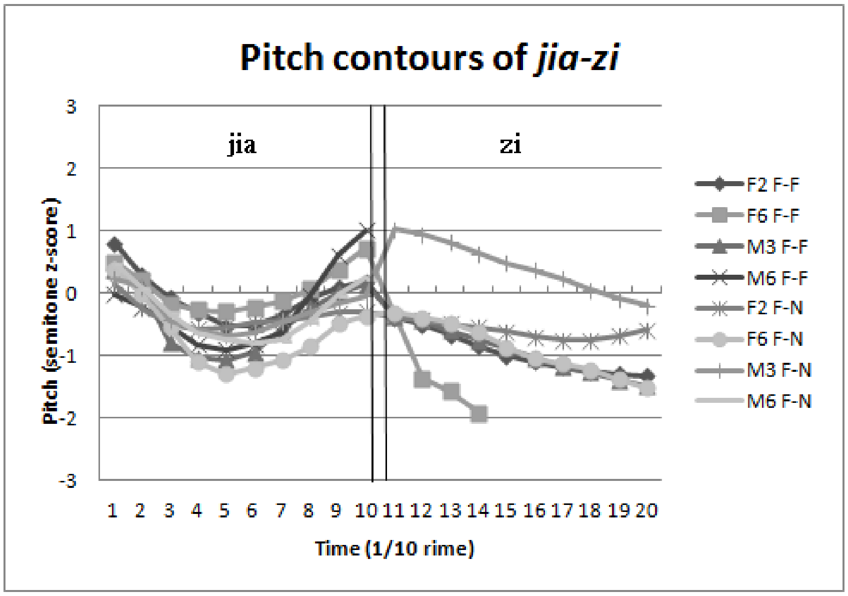


Figure 5.6 Pitch contours of the *jia-zi* pair by four TM speakers

Figure 5.6 shows the pitch contours of all the *jia-zi* stimuli. As shown in the *zi* rime, the full-neutral items have higher pitch than the full-full items. Furthermore, when the pitch contour is higher, the stimuli is more likely to be identified as full-neutral: M3's full-neutral has the highest pitch contour (95% identification as F-N); F2's full-neutral has the second highest pitch contour (88% as F-N); and F6's full-neutral, F2's full-full, and M3's full-full share similar pitch contour (70%, 65%, 73% as F-N). On the other hand, F6's full-full has a low pitch and was creaky as shown above in Figure 5.5. Therefore it was successfully identified as full-full (73%).

M6's full-full and full-neutral has no pitch reading because the consonant perturbation was so strong that both of the rimes are short and very breathy. However, M6's full-neutral sounds a lot higher than his full-full, which explain the high accuracy of M6's full-neutral (98%). On the other hand, M6's full-full is short but sounds very low. As a result of these contradictory cues, M6's full-full's accuracy rate is at the chance level (50%). Similarly, F2's full-full has a low pitch and a short duration. As a result, F2 full-full's accuracy rate is at the chance level too (35%).

The acoustic analyses above suggest that the listeners expect full-full to have long duration (S2/S1 ratio > 0.5) and low pitch at *zi*. Aside from that, creaky voice also leads the listeners to identify the stimuli as having a full tone. However, three out of four speakers did not produce the full-tone *zi* long enough, and the pitches of the full-tone *zi* were not low enough for the listeners.

When listening to the *jia-zi* pairs together, the acoustic differences between the pair seem fairly noticeable. However, the listeners had difficulty identifying the full-full patterns when they were played separately. This suggests that 1) the listeners in their past experience had heard the full-neutral *jia-zi* produced with a low-pitch neutral tone. Therefore even with the lower pitch, they were not sure whether it was a full tone; 2) the listeners expected the full tone to have a longer duration. This can be supported by the fact that F6's full-full was successfully identified. Another possible explanation is that the full-full *jia-zi* '60 years' is a low frequency word compared to the full-neutral *jia-zi* 'clip', and the listeners were more likely to choose the higher frequency word (full-neutral) when they could not hear the difference.

The result for this pair shows that the intended phonetic differences between the pairs are not always perceptible by the TM listeners. Although the speakers know the phonological differences, their phonetic outcomes are not always clear enough to be perceptible. The result suggests that the full-full *jia-zi* has gradually merged with the full-neutral *jia-zi* in Taiwan Mandarin.

5.1.5.3. **The *she-tou* pair**

The *she-tou* pair prescriptively contrasts the full-full [LH-LH] 'snake head' and the full-neutral [LH-Ø] 'tongue'. In TM, the neutral-tone *-tou* usually is not destressed and it keeps the original lexical tone [LH]. Therefore, unless the speakers produce the *-tou* with a neutral tone, I expect the listeners to have a correct identification rate that is close to 50%.

Table 5.4 Percent correct responses for the *she-tou* pair.

correct	F2	F6	M3	M6	Avg.
she-tou /LH-LH/ (F-F)	5%	**50%	20%	30%	26%
she-tou /LH-Ø/ (F-N)	90%	68%	**58%	**53%	67%

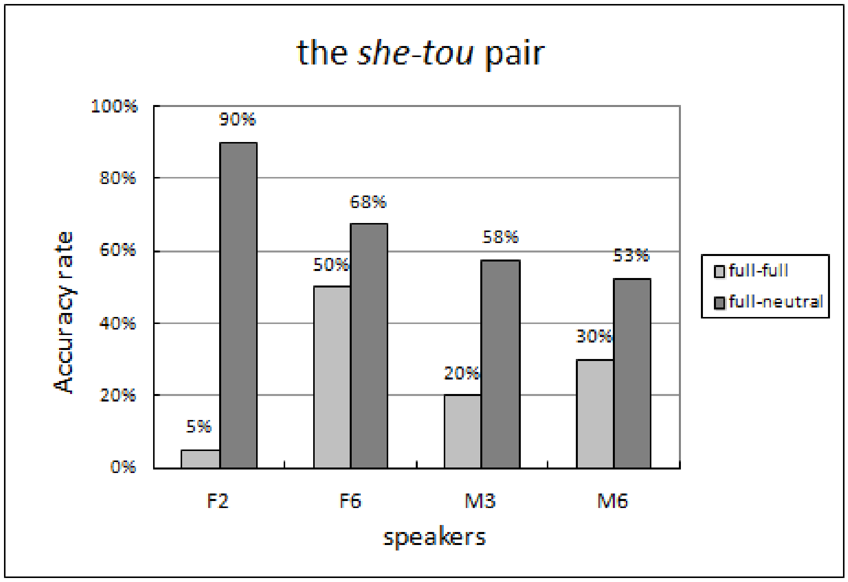


Figure 5.7 Percent correct responses for the *she-tou* pair.

As shown in Table 5.4 and Figure 5.7, most of the listeners have difficulty correctly identifying the *she-tou* pair. As expected, M3 and M6's full-neutral and F6's full-full items have chance-level accuracy rates (58%, 53%, and 50%). The rest of the items have a tendency to be identified as full-neutral regardless of the intended forms. F2's full-full and full-neutral were both identified as full-neutral at a fairly high rate (95% and 90% as full-neutral); F6's full-neutral was successfully identified (68%), while M3 and M6's full-full were identified as full-neutral (80% and 70% as full-neutral).

When the -*tou* is pronounced with a rising tone (flat pitch contour), the listeners cannot decide whether it is the full tone or the neutral tone item. The preference for the full-neutral form is possibly because the full-full *she-tou* 'snake head' is a much lower frequency word than the full-neutral *she-tou* 'tongue'.

As for the possible acoustic cues that were used by the listeners, Figure 5.8 shows that there is a medium correlation between the duration of the second syllable -*tou* and the neutral-tone identification rate (Pearson correlation coefficient: -0.71659). When the duration is shorter, the listeners are more likely to identify the item as full-neutral; when the duration is longer, the listeners are unsure about their choices. However, the correlation is rather weak, especially when the S2/S1 ratio is around 0.8. This suggests that duration was not the only acoustic cue.

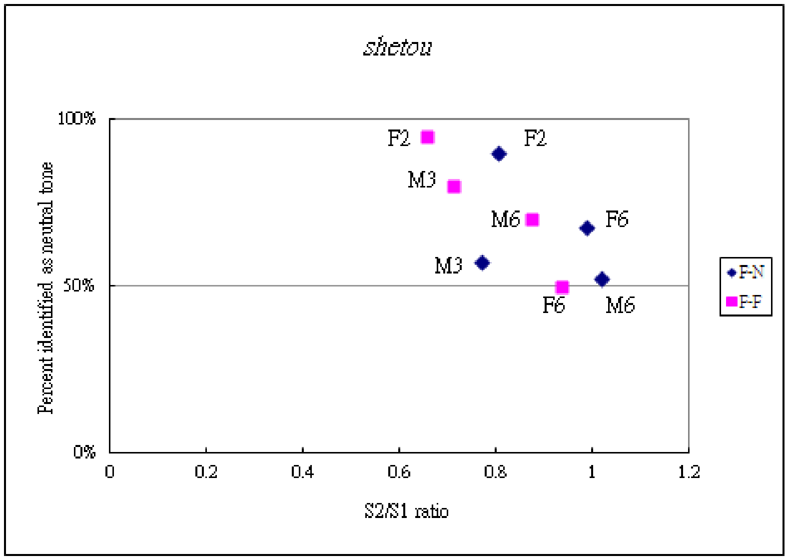


Figure 5.8 The *she-tou* pair: relationship between duration and identification as neutral tone

Pitch contours also might be able to explain some results. As shown in Figure 5.9, F2's full-full (F-F) and full-neutral (F-N) both have pitch contours that were different from the other three speaker's productions: F2's full-full and full-neutral -*tou* did not fall quickly in the beginning and stay flat or slightly rise like others. These two items also have fairly low S2/S1 ratios as shown in Figure 5.8. As a result, these two items produced by speaker F2 both have a very high percentage identified as full-neutral (95% and 90%). The rest of the pitch contours are rather similar, and they were not able to explain the variance of their full-neutral identification rates.

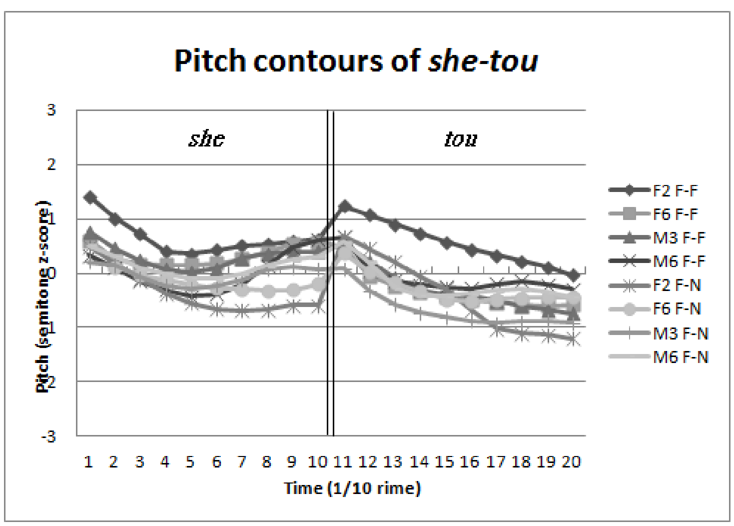
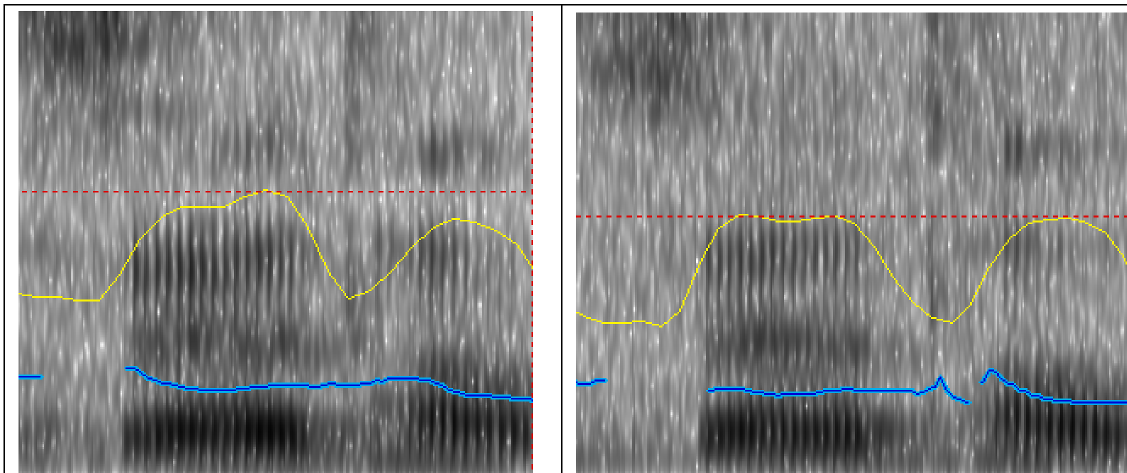


Figure 5.9 Pitch contours of the *she-tou* pairs by four TM speakers

Both F2's full-full and full-neutral have a falling pitch contour at the second syllables, which do not look like a rising tone. These pitch contours of F2 are not expected. The second syllable in full-full actually looks like a neutral-tone pitch contour in SM. However, the fact that F2 produced a falling pitch for both items suggests that F2 does not behave like SM speakers because F2 does not distinguish the pair. F2 has merged the *she-tou* pair.

After the experiment, many of the subjects said that the *she-tou* pair was the hardest of all the questions because they think the two words are homophones. Therefore, the listeners may have used a more subtle detail to help them make a decision. Figure 5.10 shows a comparison of the spectrograms of the pair for speaker M3. The pitch and duration of the syllables are almost identical, and the listeners may have used intensity as a cue. As shown, the second syllable *tou* in Figure 5.10a. has lower intensity than the first syllable *she*. This is likely to be the reason why 80% of the listeners incorrectly identified Figure 5.10a. (M3's full-full *she-tou*) as full-neutral and only 58% of the speakers identified Figure 5.10b. (M3's full-neutral *she-tou*) as full-neutral.



a. *she-tou* /LH-LH/ of M3	b. *she-tou* /LH-Ø/ of M3

Figure 5.10 Spectrograms of the *she-tou* pair produced by M3

5.1.5.4. **The *gui-men* pair**

This pair contrasts the full-full *gui-men* [L-LH] ‘ghost gate’ with the full-neutral *gui-men* [L-Ø] ‘ghosts’. In TM, the neutral-tone -*men* is either produced with a low pitch target, or produced with a /LH/ tone quality. I expect the listeners to identify the full-neutral -*men* quite easily if it is produced with a low target. On the other hand, I expect the pair to merge into the full-full if the speaker produces the neutral-tone -*men* with a /LH/ tone.

The results of the *gui-men* pair are shown in Table 5.5 and Figure 5.11. The result shows that F2 and M6 merged the pair into full-full ([L-LH]). Their full-neutral were misidentified as full-full by majority of the listeners (90% and 85% chose full-full). On the other hand, speaker F6 still distinguishes the pair. Both of her full-full and full-neutral items have high accuracy rates. As for speaker M3, most of the listeners identified his full-neutral (83%), but his full-full was rather confusing for the listeners (63% accuracy rate). The results show that F2 and M6 merged the pair into full-full, and F6 and M3 still distinguish the pair.

Table 5.5 Percent correct responses for the *gui-men* pair

correct	F2	F6	M3	M6	Avg.
gui-men /L-LH/ (F-F)	98%	100%	**63%	93%	88%
gui-men /L-Ø/ (F-N)	10%	98%	83%	15%	51%

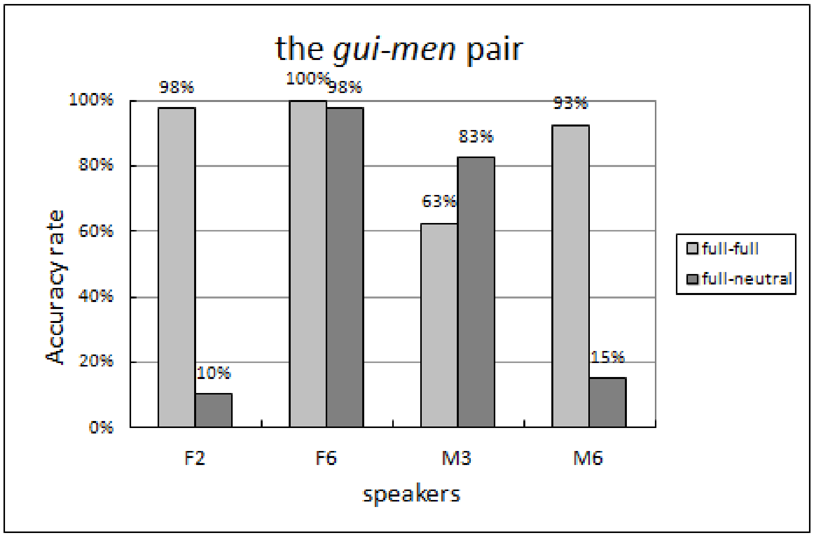


Figure 5.11 Percent correct responses for the *gui-men* pair

The listeners did not seem to use duration as an acoustic cue for their choice. As shown in Figure 5.12, the correlation between the duration and the full-neutral identification rate was weak (Pearson correlation coefficient: -0.56229). The lower S2/S1 ratio did not necessary result in higher full-neutral identification rate.

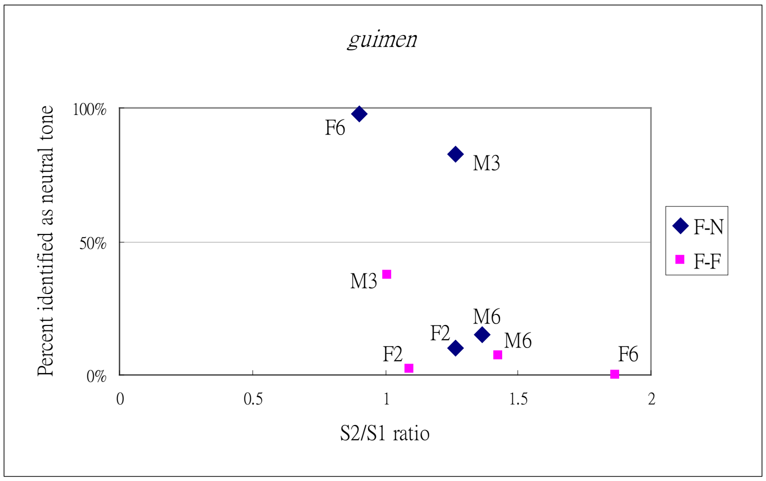


Figure 5.12 The *gui-men* pair: relationship between duration and identification as neutral tone

On the other hand, the listeners used pitch as an acoustic cue. As shown in Figure 5.13, M3's full-neutral (F-N) has a falling pitch contour on *-men*, and F6's full-neutral *-men* has only partial pitch contour due to the slight creaky voice which associates with the lowering of the pitch. These two items were identified as full-neutral with confidence (98% and 83%). On the other hand, the other pitch contours all rise in the second half of *-men*, resulting in high percentage of being identified as full-full ([L-LH]).

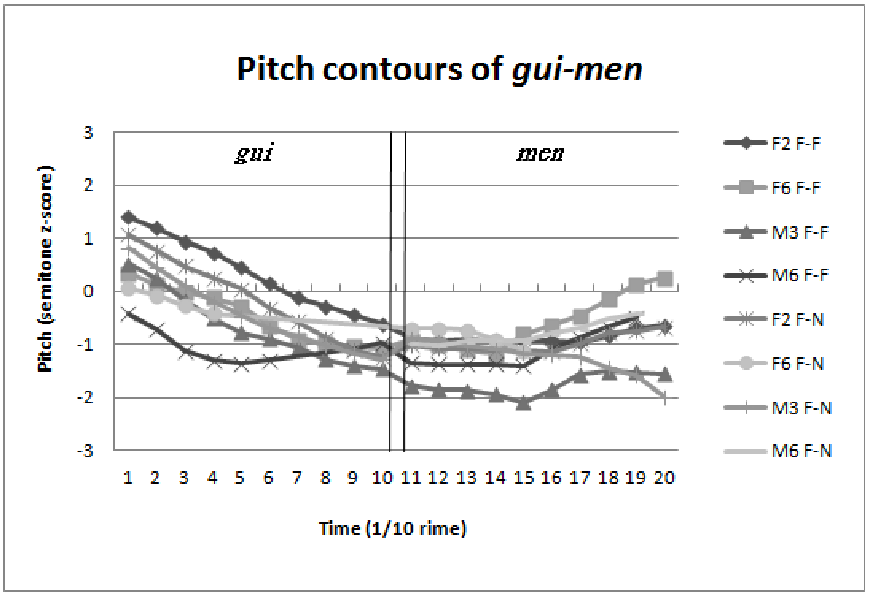
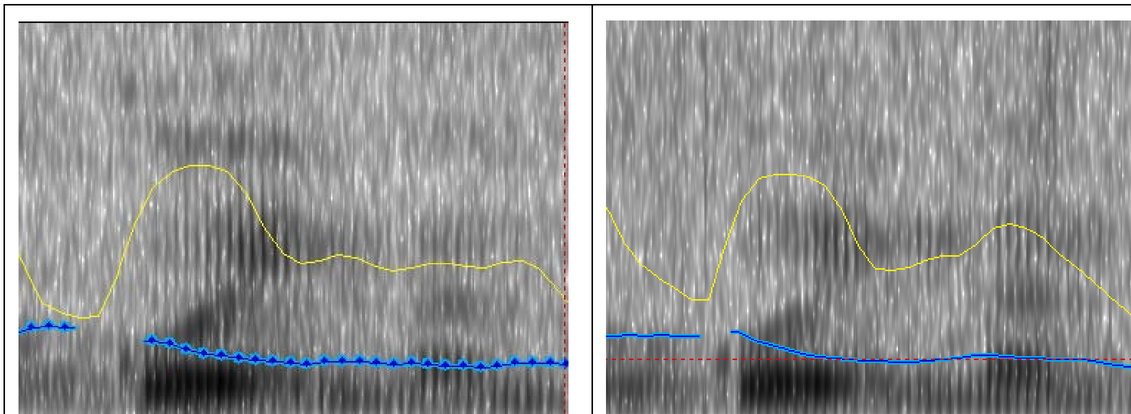


Figure 5.13 Pitch contours of the *gui-men* pair by four TM speakers

Although M3's full-full has a rising contour in the second half of *-men*, only 63% of the listeners correctly identified it as full-full. This is probably because his *-men* is noticeably shorter in duration (see Figure 5.12) and lower in intensity (Figure 5.14a). The listeners received conflicting signals from M3's full-full item. Therefore the accuracy rate is within chance level. For M3, he maintained the pitch distinction between the *gui-men* pair, but he also has a tendency to make the second syllable lighter as shown in both of the spectrograms in Figure 5.14. Therefore the listeners could identify his full-neutral but not his full-full.



a. full-full *gui-men* [L-LH] by M3	b. full-neutral *gui-men* [L-Ø] by M3

Figure 5.14 Spectrograms of the *gui-men* pair produced by M3

The result shows that some speakers (F2 and M6) merged the pair and some speakers (F6 and M3) distinguished the pair. The either very high or very low accuracy rates also suggest that when the *-men* is produced without a low pitch target, they feel fairly comfortable to choose full-full. This suggests that listeners expect the neutral-tone *-men* to have a low pitch target even if speakers do not usually produce the low pitch target.

5.1.5.5. **The *zuo-zhe* pair**

This pair contrasts the full-full *zuo-zhe* [HL-L] 'reporter' and the full-neutral *zuo-zhe* [HL-Ø] 'sitting'. I expect the listeners might have some difficulties distinguishing the *zuo-zhe* pair because the neutral-tone *zhe* also has a low target.

As shown in Table 5.6 and Figure 5.15, none of the speakers produced the *zuo-zhe* pairs that were distinct to the listeners. Only half of the items were correctly identified, including one full-full item (F6) and three full-neutral items (F2, M3, and M6). The other half of the items were ambiguous or even misidentified (M6's full-full). The results suggest that the distinction is not maintained well by these TM speakers. The results show that the pair tended to merge to full-neutral except for F6.

Table 5.6 Percent correct responses for the *zuo-zhe* pair

Speaker	F2	F6	M3	M6	Avg.
zuo-zhe /HL-L/ (F-F)	**63%	98%	**63%	28%	63%
zuo-zhe /HL-Ø/ (F-N)	100%	**35%	80%	100%	79%

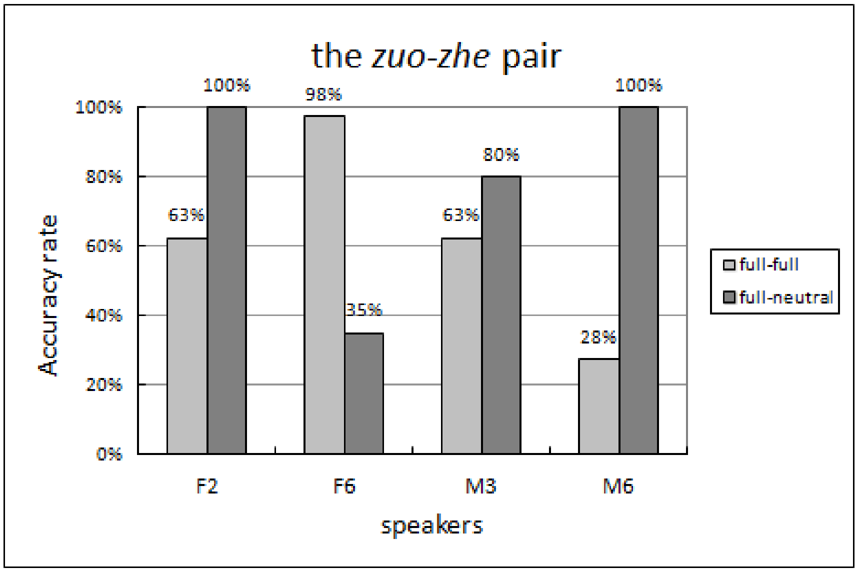


Figure 5.15 Percent correct responses for the *zuo-zhe* pair

Duration seems to be an important acoustic cue used by the TM listeners for this pair. As shown in Figure 5.16, when the S2/S1 ratio is lower, the listeners are more likely to identify the pair as full-neutral. The S2/S1 ratio and the percentage identified as neutral tone are highly correlated (Correlation coefficient: -0.83371). However, we can also see that the speakers did not necessarily produce full-neutral (F-N) with shorter duration or produce full-full (F-F) with longer duration, such as M6's full-full and F6's full-neutral.

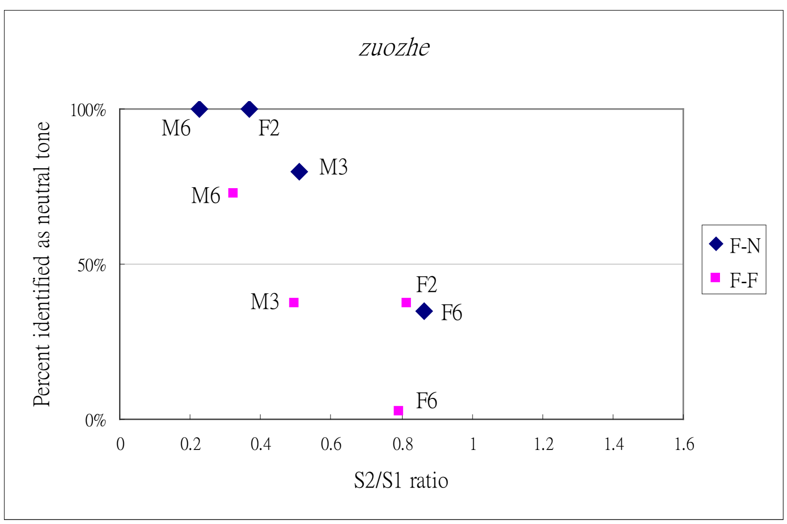


Figure 5.16 The *zuo-zhe* pair: relationship between duration and identification as neutral tone

Although duration is an important acoustic cue for the listeners, it is not the only cue. Figure 5.16 also shows that similar S2/S1 ratios could result in different percentages identified as neutral tone. It is very likely that listeners used pitch as an additional perceptual cue. Higher pitch in *-zhe* resulted in a higher neutral-tone identification rate. For example, as shown in Figure 5.16, F2's full-neutral, M3's full-neutral, and M3's full-full all have S2/S1 ratios that are around 0.4, but the three items all have different neutral-tone identification rates. Figure 5.17 shows that the pitch level of *-zhe* are highest for F2's full-neutral (above -1 semitone z-score), resulting in highest neutral-tone identification rate (100%). The pitch contour of M3's full-neutral is lower (80% identified it as neutral), and the pitch contour of M3's full-full is the lowest (63% identified it as neutral).

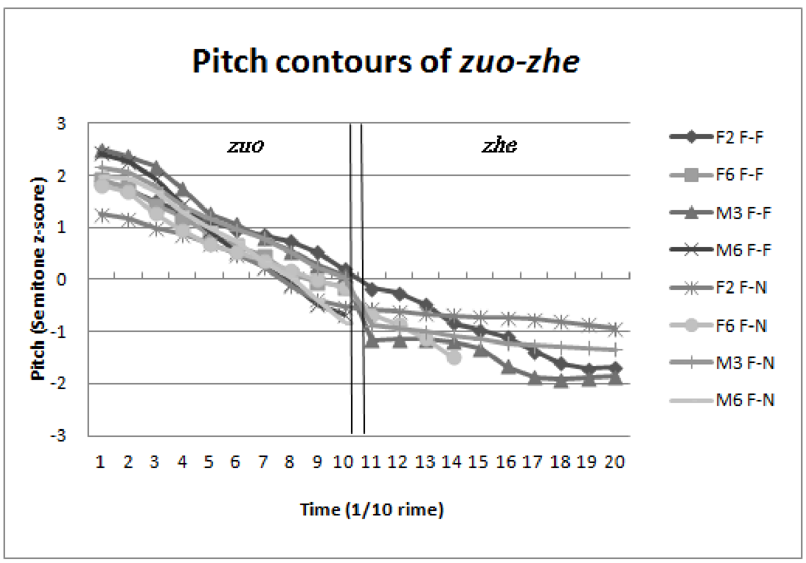


Figure 5.17 Pitch contours of the *zuo-zhe* pair by four TM speakers

Unfortunately not all the *zuo-zhe* items have pitch readings in the rime of *-zhe* in Figure 5.17. Although F6's *zuo-zhe* pair both have a similar S2/S1 ratio, their identification rates are quite different. As shown in Figure 5.18, F6's *-zhe* in full-full (a) is completely creaky, while her *-zhe* in full-neutral (b) is only creaky in the second half. Since creaky voice is usually associated with the low tone, 98% of the listeners identified F6's full-full as full-full, and 65% (100-35%) of the listeners identified F6's full-neutral as full-full. It seems that the voicing quality is also an important acoustic cue.

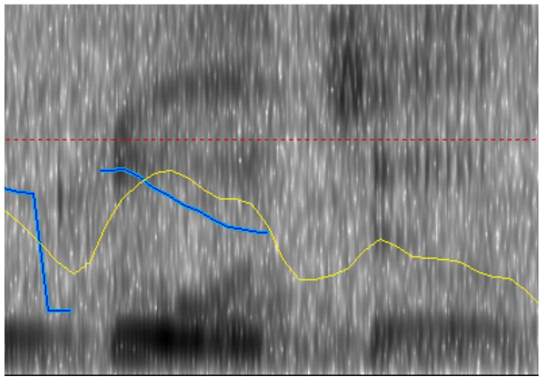
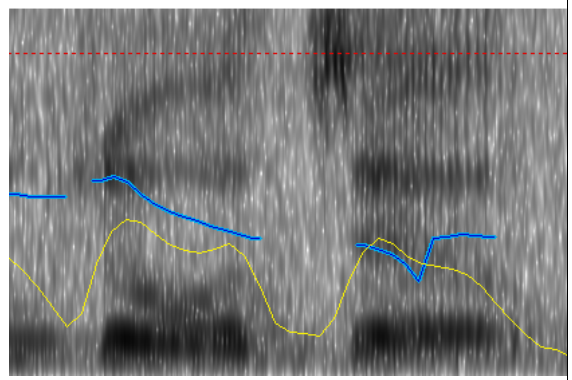
	
a. full-full *zuo-zhe* [HL-L] by F6	b. full-neutral *zuo-zhe* [HL-Ø] by F6

Figure 5.18 Spectrograms of the *zuo-zhe* pair produced by F6.

M6's full-full and full-neutral are another two items without pitch readings on the -*zhe* syllables. As shown in Figure 5.19a, although M6's full-full -*zhe* is really creaky, it is as short as only 3 glottal pulses (49ms). Therefore only 28% of the listeners think it is full-full. On the other hand, M6's full-neutral -*zhe* (See Figure 5.19b) is so weak and short that it is almost a syllabic consonant. As a result 100% of the speaker identified it as full-neutral.

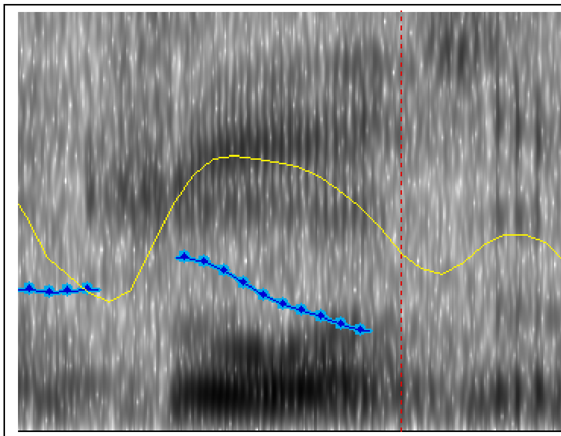
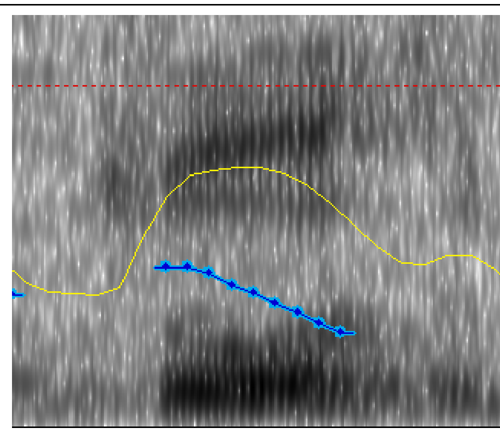
	
a. full-full *zuo-zhe* [HL-L] by M6	b. full-neutral *zuo-zhe* [HL-Ø] by M6

Figure 5.19 Spectrograms of the *zuo-zhe* pair produced by M6

The importance of voicing quality can also be found in the result of F2's full-full. F2's full-full -*zhe* has a rather long duration (Figure 5.16) and rather low pitch contour (Figure 5.17). However, it seems that long duration and low pitch are not sufficient acoustic cues. Only 63% of the speaker identified F2's full-full item as full-full. Perhaps the absence of the creaky voice on the *-zhe* syllable (Figure 5.20) makes it hard to identify as a full tone.

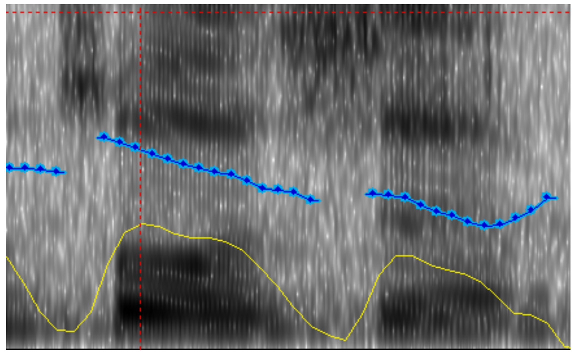


Figure 5.20 Spectrogram of the full-full *zuo-zhe* by speaker F2

The result shows that the full-neutral pattern is much more likely to be identified, which is opposite of what I expected. Higher pitch and short duration seem to be the acoustic cues for the full-neutral pattern, and most of the speakers were able to produce that. On the other hand, creaky voice and longer duration seem to be the acoustic cue for the full-full pattern. However, three out of four speakers in this study did not produce that. This result suggests that TM listeners expect there to be distinctions between the pair. However, the TM speakers do not consistently produce the distinction that the listeners expect.

The higher tendency to choose full-neutral is likely not due to the frequency of the words. Both ‘writer’ and ‘sitting’ are very common words. One possible explanation is that listeners probably had heard the full-neutral *zuo-zhe* /HL-Ø/ produced like the full-full *zuo-zhe* /HL-L/ before. Therefore unless there was a very strong acoustic cue such as long creaky voice, the listeners were not sure what the stimulus is.

5.1.5.6. **The *du-zhe* pair**

Like the *zuo-zhe* pair, the *du-zhe* pair contrasts the same morpheme *zhe*. The two pairs differ in their lexical tones of the first syllables. The listeners need to identify either the full-full *du-zhe* [LH-L] ‘reader’ or the full-neutral *du-zhe* [LH-Ø] ‘reading’.

Table 5.7 Percent correct responses for the *du-zhe* pair

correct	F2	F6	M3	M6	Avg.
du-zhe /LH-L/ (F-F)	93%	100%	**65%	98%	89%
du-zhe /LH-Ø/ (F-N)	98%	**55%	85%	93%	83%

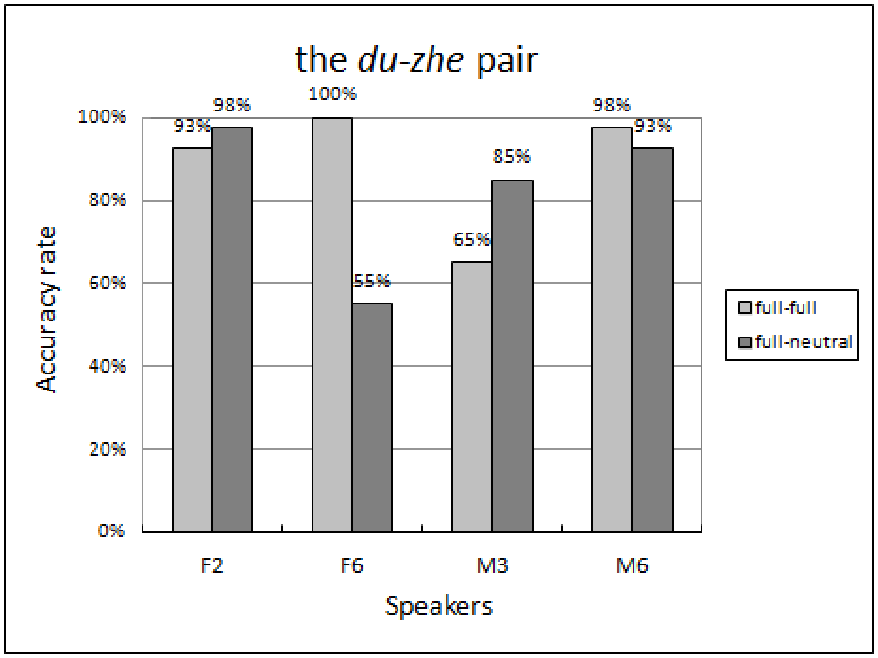


Figure 5.21 Percent correct responses for the *du-zhe* pair

The results are shown in Table 5.7 and Figure 5.21. Both F2 and M6 made a distinction between the pair, and the distinctions were perceived by the listeners. F6's full-full was correctly identified, but her full-neutral was ambiguous. M3's full-full was ambiguous, but his full-neutral was correctly identified. Generally the accuracy rates were very high.

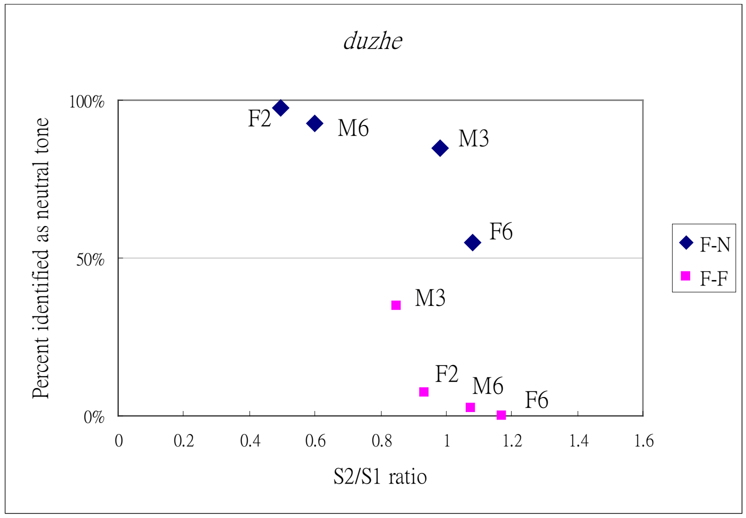


Figure 5.22 The *du-zhe* pair: correlation between duration and identification as neutral tone

As shown in Figure 5.22, duration seems to be an important acoustic cue for the listeners. When the S2/S1 ratio of *du-zhe* was higher, the listeners were less likely to choose full-neutral. The S2/S1 ratio and the percentage identified as neutral tone are negatively correlated (correlation coefficient: -0.73196). However, when we look at the range between 0.8-1.2 S2/S1 ratio, there seem to be other acoustic cues governing the choice between full-full and full-neutral.

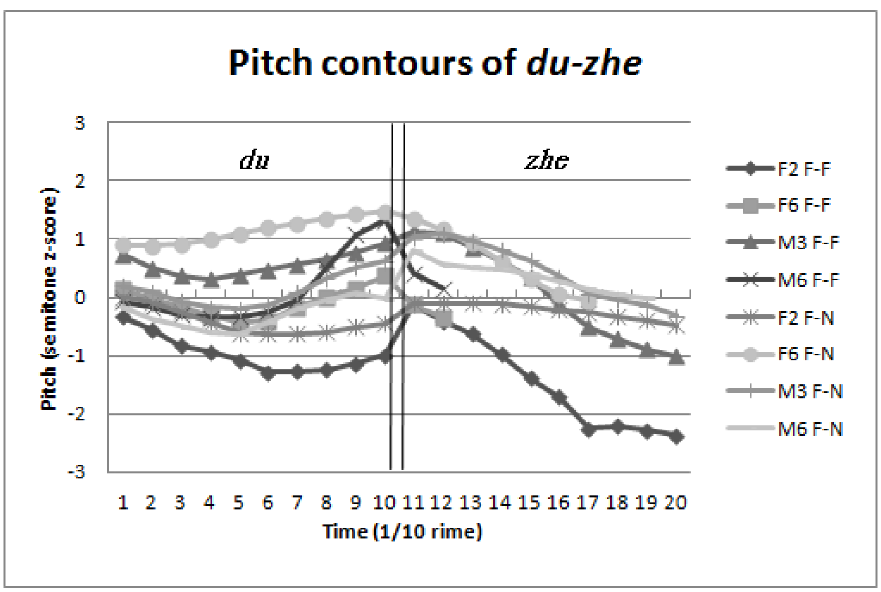


Figure 5.23 Pitch contours of the *du-zhe* pair by four TM speakers

Figure 5.23 shows that pitch might be an important acoustic cue for the listeners. As we can see in the *zhe* syllable, F2, M3, and M6's full-neutral (F-N) all have rather high pitches, and all three items were mostly identified as full-neutral (98%, 85%, and 93%). F6's full-neutral also has a similar pitch contour, but the second half of her *zhe* syllable has creaky voice (See the partial pitch contour in Figure 5.23 and the spectrogram in Figure 5.24). Also, the duration of F6's neutral-tone *zhe* syllable is longer than others (See Figure 5.22). As a result, her full-neutral item was ambiguous, only 55% of the listeners identified it as full-neutral.

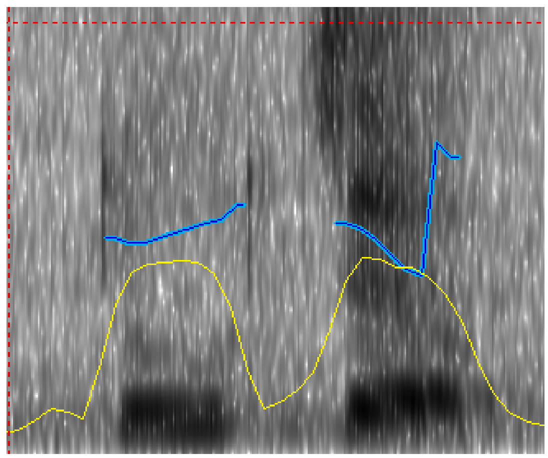
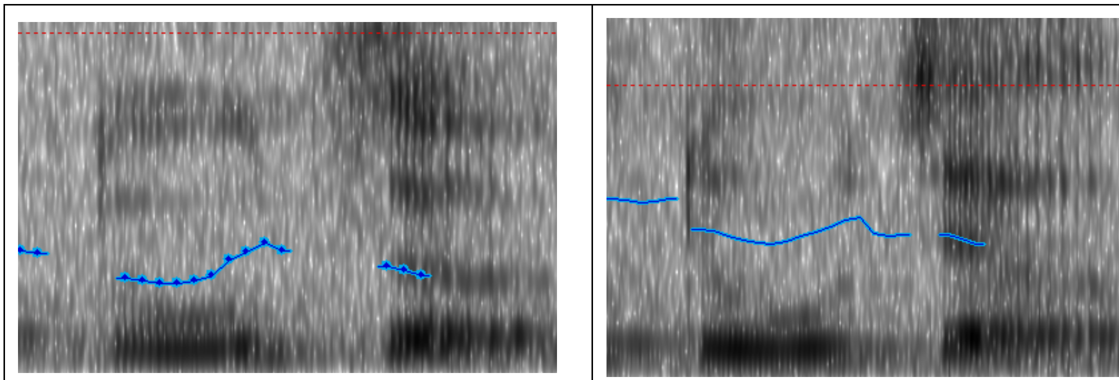


Figure 5.24 Full-neutral *du-zhe* [LH-Ø] produced by F6

In addition, Figure 5.23 shows that F2's full-full and M3's full-full have lower pitch contours in the *zhe* syllable. Although M3's full-full has a lower pitch contour reaching toward -1 semitone z-score, it was not as low as F2's full-full. With similar durations of these two items (See Figure 5.22), pitch is probably the reason why only 65% of the listeners correctly identified M3's full-full, while 93% of the listeners correctly identified F2's full-full.

F6 and M6's full-full did not have pitch readings for most of their *zhe* syllables. The pitch readings were unavailable because F6 and M6's full-tone *zhe* syllables have creaky voice as shown in Figure 5.25. The creaky voice is very likely to be the acoustic cue. 100% and 98% of the listeners correctly identified F6's and M6's full-full *du-zhe* respectedly.



a. *du-zhe* [LH-L] produced by M6 | b. *du-zhe* [LH-L] produced by F6

Figure 5.25 Spectrograms of the full-full *du-zhe* items produced by M6 and F6

The result of the *du-zhe* pair shows that generally the accuracy rates are higher than *zuo-zhe*. It is likely because the pitch of the *du* syllable ends higher than *zuo*; the pitch fall in *-zhe* after *du* is more obvious in distinguishing the tones. Therefore many of the items were successfully identified. However, in some cases the speakers produce items that are not obvious, suggesting that even with the knowledge of the items being different, the speakers does not always make the distinction when producing the items.

5.1.6. Cross-speaker variation

The results in 5.1.5 compare the accuracy rates by the tested pairs. However, there seem to be certain patterns that are speaker-specific. Table 5.8 shows the average accuracy rates for each speaker's productions. Both the accuracy rates of the tested pairs and fillers are included. Both the tested pairs and fillers are tonal minimal pairs. However, it is obvious that the accuracy rates for the tested minimal pairs are lower. The results also show that there were significant cross-speaker differences in the tested pair [χ^2= 11.51, *df* = 3, p = .009 < .01]. Speaker F6's tested pairs were identified with the highest accuracy rate (79%). On the other hand, the accuracy rates for the other three speakers ranged between 70 ~ 73%, and their differences from each other were not statistically significant [χ^2= 1.33, *df* = 2, p = 0.51]. The cross-speaker difference also influenced the accuracy rate for the fillers [χ^2= 39.16, *df* = 3, p < .001]. F6's fillers were identified with an accuracy rate of 100%. This shows that she probably maintains more distinctions or pronounces things more carefully. On the other hand, the other three speakers do not have accuracy rates as high as speaker F6. Especially for speaker M3, the accuracy rate of his fillers is lower than others, showing that distinctions in his speech are harder for the listeners to perceive.

Table 5.8 Cross speaker differences

Average accuracy rates	F2	F6	M3	M6	Average
Tested pairs	73%	79%	70%	71%	73%
Fillers	97%	100%	92%	99%	97%

The results organized by speakers with the detail of the pairs and tonal patterns are provided in Figure 5.26 and Table 5.9. Since the *yi-zi* pair has high accuracy rates across the speakers, the figure below only shows the average accuracy rates of the other five pairs. The cross-speaker comparison shows that F2, M3, and M6 all have higher accuracy rates on the full-neutral items than the full-full items. On the other hand, F6's full-full items have higher accuracy rates than her full-neutral items.

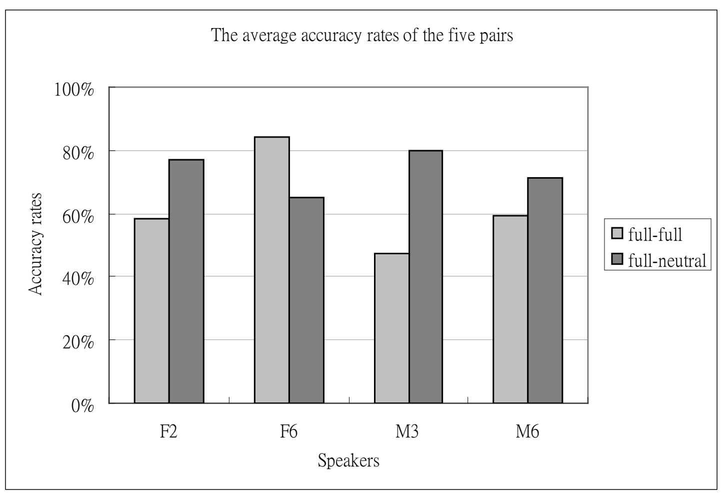


Figure 5.26 Average percent correct responses by speakers

Table 5.9 presents each speaker's accuracy rates in detail. The ** sign indicates that the percentage is at the chance level as mentioned in footnote 18 on page 116. As shown in the table, we can see that F2 and M6 shared a similar pattern. Both of the speakers distinguished the *du-zhe* pair, and both of them merged the *gui-men* pair into full-full. Overall, F2 and M6 both generally had high accuracy rates for full-neutral forms except for *gui-men*[20], while their full-full items were not consistently identified.

F6 generally made the distinction clear. Only the full-full *she-tou* and the full-neutral *zuo-zhe* and *du-zhe* were not successfully identified. Compared to other speakers, her full-full items had much higher accuracy rates, and her full-neutral items generally had lower accuracy rates except for *gui-men*. This suggests that F6 used a hyperarticulate speech style.

M3 was different from the other speakers. His full-full items were either confusing or misleading. His full-neutral items on the other hand had rather high accuracy rates except for *she-tou*. This shows that he had a tendency to merge full-full into full-neutral.

Table 5.9 The accuracy rates of the five pairs by speakers

	F2		F6		M3		M6	
	F-F	F-N	F-F	F-N	F-F	F-N	F-F	F-N
jia-zi	**35%	88%	73%	70%	28%	95%	**50%	98%
she-tou	5%	90%	**50%	68%	20%	**58%	30%	**53%
gui-men	98%	10%	100%	98%	**63%	83%	93%	15%
zuo-zhe	**63%	100%	98%	35%	**63%	80%	28%	100%
du-zhe	93%	98%	100%	**55%	**65%	85%	98%	93%
Avg.	59%	77%	84%	65%	48%	80%	60%	72%

Before the subjects recorded the speech, they were asked to complete the linguistic background forms as mentioned in 3.1.3. The information they provided show that M3 is the only speaker who considered his Mandarin “not so standard”, while the other three speakers all considered their Mandarin “fairly standard”. It is worth noting that M3 is also the speaker who treated the neutral tone as the low tone in novel productions. The distinction between the full-full and the full-neutral pair was definitely difficult for him.

In summary, F6 was the most conservative speaker who had a more obvious distinction between full-full and full-neutral. F2 and M6 generally had the tendency to merge the pair into full-neutral. However, they still maintained the distinction between the *du-zhe* pair. M3 had the strongest tendency to merge into full-neutral.

[20] And M6’s full-neutral *she-tou* has a mid accuracy rate.

5.1.7. Discussion

The result of Experiment 1 is presented in Table 5.10. The shaded area shows that the pair produced by the speaker was distinct. As we can see, the *yi-zi* pairs produced by all four speakers were distinct. F6 produced a clear distinction in the *jia-zi* and *gui-men* pairs, and F2 and M6 produced a clear distinction in the *du-zhe* pairs. For the *yi-zi* pair, since the feature of the TM neutral tone diverges from its full-full counterpart, the result is what I expect – a clear distinction. For the rest of the pairs, I expect the TM neutral tone to converge with its full-full counterpart, and the result supports our hypothesis. The listeners had difficulty distinguishing the full-full and full-neutral pair. Out of the twenty pair instances (four speakers producing five pairs), only four of them were distinct for the listeners. The TM listeners could not distinguish 80% of the pair instances. This suggests that the phonological distinction between the full tone and the neutral tone is very weak. The neutral tone is on the verge of merging with the full tone. Certain pairs such as *du-zhe* still retain the distinction in some cases. Some speakers, such as F6, also keep the distinction better than others.

Table 5.10 Summary of the results of Experiment 1

	F2	F6	M3	M6
yi-zi	distinct	distinct	distinct	distinct
jia-zi	F-N	distinct	F-N (M)	F-N
she-tou	F-N (M)	none	none	none
gui-men	F-F (M)	distinct	F-N	F-F (M)
zuo-zhe	F-N	F-F	F-N	F-N (M)
du-zhe	distinct	F-F	F-N	distinct

For the pair instances that were not distinct, "F-F" (full-full) and "F-N" (full-neutral) in the table show that the written items were correctly identified. "None" means that neither member of the pair was correctly identified; and (M) means the pair merged. For example, "F-N (M)" means that the full-neutral item was correctly identified, and the full-full item was misidentified as full-neutral. "F-N" means that the full-neutral item was correctly identified, but the full-full item was confusing.

Out of the five pairs, the *she-tou* pair was the most confusing for the listeners. The listeners apparently expected the neutral-tone to be produced like a rising tone. As a result, not even the full-full items were correctly identified. Speaker F2's full-neutral is correctly identified because she produced the neutral-tone *-tou* with a low pitch target, which is an SM feature. Interestingly, her full-full item was misidentified as full-neutral because her full-full item also had a low pitch target. This is probably a speech error due to the low frequency of the word because neither TM nor SM speakers would reduce *-tou* in *she-tou* 'snake-head'. On the other hand, the *gui-men* pair shows different patterns. The listeners expected the neutral-tone *-men* to have a low pitch which is different from the full tone *-men* /LH/. Therefore the listeners were fairly comfortable identifying the full-full items. The full-full items generally have high accuracy rates except for M3, who has the tendency to shorten the second syllable. We can also clearly observe that F2 and M6 merged the full-neutral into the full-full while F6 retained the distinction between the pair.

As for *jia-zi*, *zuo-zhe*, and *du-zhe*, both the listeners and the speakers knew that there are distinctions between the pairs. However, the listeners did not always correctly identify the items produced by the speakers. The pairs produced by F6 had higher accuracy rates, which suggests that the distinction she made was better perceived by the listeners. The result also shows that *du-zhe* seems to be the pair for which the distinction was easier to perceive. It is probably because the neutral-tone *-zhe* morpheme is a productive suffix, and generally suffixes are more susceptible to reduction than stem morphemes. Therefore its neutral-tone quality such as short duration is easier to maintain. On the other hand, the neutral-tone *-zi* is an obsolete suffix, and the full-neutral word with *-zi* is lexicalized. This makes it more likely to acquire the properties of full-tone. Also, the pitch of the *zuo* /HL/ in *zuo-zhe* ends in a low pitch and the pitch of the *du* /LH/ in *du-zhe* ends in a high pitch. Both the full-tone and the neutral-tone *zhe* have a low pitch target; the pitch fall is bigger for the *zhe* syllables when following *du*. The large pitch fall probably made it easier to distinguish the *du-zhe* pair than the *zuo-zhe* pair.

For these three pairs, when the pair instances were not distinct, the listeners tended to choose full-neutral (except for F6). Most of the full-neutral items were correctly identified, and it was the full-full items that were often confusing or misidentified as full-neutral. A possible explanation is that these full-neutral items in TM have become more and more similar to the full-full. As a result, when the listeners hear a full-full item, their experiences suggests that these items are still likely to be full-neutral. These full-full items therefore have lower accuracy rates even though the speakers intended to produce the distinction.

Overall, the results of Experiment 1 show the instability of the full-full and full-neutral distinction. The neutral-tone and the full-tone are merging together slowly: some suffixes are faster than others, and some speakers are faster than others. The phonological difference is known by both the speakers and the listeners, but the intended speech is not always successfully identified by the listeners. Speakers are not all producing the same cues, and listeners are not all listening for the same cues. As a result, only 20% of the minimal pairs were distinct. The directions of the merger were not consistent.

This experiment also shows that duration, pitch contours, intensity, and voice quality may all be important acoustic cues. The next experiment investigates the acoustic cues used by the listeners.

5.2. Perception Experiment 2

The preliminary acoustic analyses in Perception Experiment 1 show that the TM listeners rely on duration, pitch, voice quality, and perhaps intensity to make their decisions. It looks as if long duration, lower pitch and creaky voice result in choosing full-full patterns. In Perception Experiment 2, I further investigate the acoustic cues used by TM listeners to distinguish between a neutral-tone syllable and a full-tone syllable, with the focus on pitch.

5.2.1. Research Design

From the results of Perception Experiment 1, we can see that pitch was one of the perceptual cues used by the listeners. Higher pitch contours seemed to result in the identification of a neutral tone. However, it was unclear how high the pitches have to be, which part of the pitch contour was used as a cue, or what kind of pitch contour shapes were used as a cue. Furthermore, the data from Perception Experiment 1 was not controlled. It was unclear whether the listeners were using some other perceptual cues to make their decisions.

In this experiment, I examine the pitch cues used to distinguish the neutral tone from the low tone. The method of this perception experiment was the same as Perception Experiment 1: the listeners were asked to identify one of the tonal minimal pair of full-full vs. full-neutral. However, in order to control for other possible perceptual cues, this experiment use re-synthesized speech to manipulate the pitch contours. The duration was controlled so as to be the same and the pitch contours was manipulated into different pitch shapes and pitch heights. Two tonal minimal pairs were included in this experiment. The minimal pair *ji-zhe* /HL-L/ 'reporter' vs. *ji-zhe* /HL-Ø/ 'remembering' varied the beginning pitch and the initial pitch fall; and the minimal pair *wei-zhe* /LH-L/ 'violator' vs. *wei-zhe* /LH-Ø/ 'surrounding' varied the pitch contour shape and the end pitch.

I expected to see that the pitch contours with a higher beginning pitch, a smaller pitch fall, a higher end pitch, and a slow-fall pitch contour would be more likely to be identified as a neutral tone. On the other hand, the pitch contours with a lower beginning pitch, a large pitch fall, a lower end pitch, and a quick-fall pitch contour would be more likely to be identified as a low tone.

The subjects and the task in Experiment 2 were the same as in Perception Experiment 1 as described in 0. The same subjects were asked to complete Experiment 2 after finishing Experiment 1. The task is the same as described in 5.1.2 where subjects were asked to choose between the minimal pair of full-full tones and full-neutral tones.

5.2.2. Materials

The words *ji-zhe* /HL-L/ 'reporter' and *ji-zhe* /HL-Ø/ 'remembering', and *wei-zhe* /LH-L/ 'violator' and *wei-zhe* /LH-Ø/ 'surrounding', produced by a TM female speaker with a loud and clear voice, were selected. Both the full-full and the full-neutral productions were included in this experiment because I wanted to keep the stimuli as natural as possible. Also, this way I could test whether the listeners used other perceptual cues, most likely to be vowel or voice qualities of the *zhe* syllables, to make their decisions. Figure 5.27 plots the vowels of *zhe* /Ø/ and *zhe* /L/ in both *ji-zhe* and both *wei-zhe* in the vowel space of this TM female speaker. As shown in the figure, the vowel of *zhe* /L/ was a little more back than the vowel of *zhe* /Ø/. The F2 of *zhe* /Ø/ was about 125 Hz higher than the F2 of *zhe* /L/ as illustrated in Table 5.11. This difference of vowel backness might be used as a perceptual cue. Also, there were some voice-quality differences between the minimal pairs, which might be used as a perceptual cue. I will discuss the voice-quality differences in the result section 5.2.4.

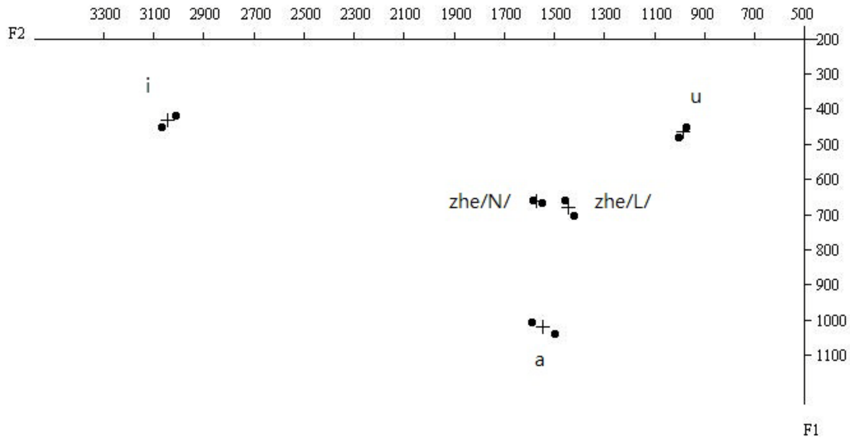


Figure 5.27 Vowel quality of zhe /Ø/ and zhe /L/ by the female speaker

The durations and the pitch contours of the *zhe* syllables in both *ji-zhe* and both *wei-zhe*, all four syllables, were manipulated using Praat's function "To manipulate". After the manipulation, the *zhe* syllables were put in the same frame sentences: *qing-shuo wei-X ba-ci* 'please say *wei-zhe* eight times' or *qing-shuo ji-X ba-ci* 'please say *ji-zhe* eight times' in order to control the sentence environment.

Table 5.11 F1, F2, and the durations of the rimes of *zhe* used in Perception Experiment 2

Rimes of *zhe*	*wei-zhe* /L/	*wei-zhe* /Ø/	*ji-zhe* /L/	*ji-zhe* /Ø/
F1 (Hz)	700	664	656	656
F2 (Hz)	1424	1551	1462	1590
Duration (ms)	133	75	130	80

The durations of the *zhe* syllables were controlled to be the same, which was the average between the minimal pairs. The original durations of all four *zhe* rimes are shown in Table 5.11 above. After the manipulation, all the *wei-zhe* stimuli have a rime of *zhe* with a duration of 104 ms, while all the *ji-zhe* stimuli have a rime of *zhe* with a duration of 105 ms.

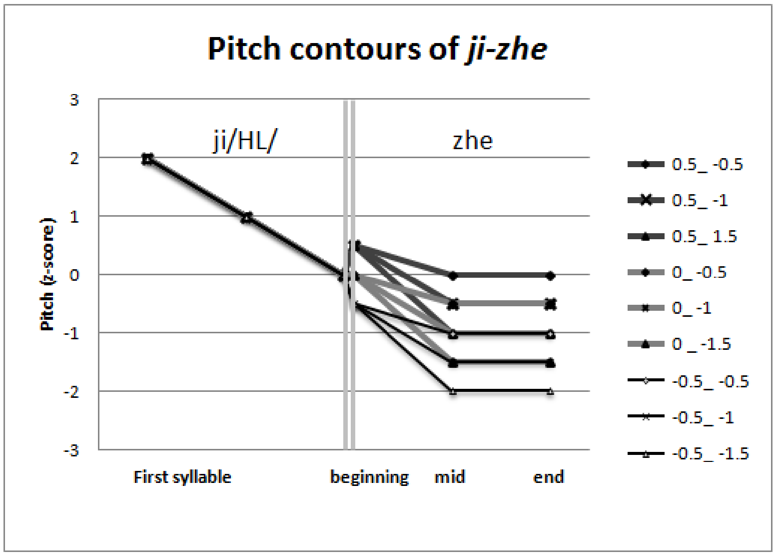


Figure 5.28 Schematized illustrations of synthesized stimuli *ji-zhe*

The pitch contours of *zhe* in *ji-zhe* /HL-L/ 'reporter' and /HL-Ø/ 'remembering' were re-synthesized into 9 pitch contours. I varied the beginning pitch and the initial pitch fall because they are both cues to identify the low tone in TM (Fon et al. 2004). In Figure 5.28, the x-axis shows an illustration of the time dimension of both of the syllables; the y-axis shows the pitch, which was presented in z-score[21]. The pitch of TM speakers usually ranges from -2 to 2 z-score. As shown in the figure, three starting pitches of the second syllables differed: +0.5, 0, and -0.5 z-score. The pitch falls of the first half also varied between 0.5, 1, and 1.5 z-scores. In total, 2 (original words of /HL-L/ and /HL-Ø/)* 3 (beginning pitch)*3 (pitch fall) = 18 stimuli were created.

As for *wei-zhe*, the pitch contours of -*zhe* in /LH-L/ 'violator' and /LH-Ø/ 'surrounding' were re-synthesized into 6 different contours, varying the end pitch (-1 or -1.5 z-score) and the pitch-contour shape (convex or rapid-fall, straight, concave or slow-fall) that reached to the low end pitch (See Figure 5.29) because based on the result of the production test, these two features seem to best characterize the differences between the two tones. The pitch-contour shapes were schematized by controlling the midpoint pitches: midpoint+0.5 z-score (convex), midpoint (straight), and midpoint-0.5

[21] Instead of using semitone z-score as in the production study, z-score was used here. I changed the pitch presentation from z-score to semitone z-score after the perception experiments were conducted. The conversion between the z-score and the semitone z-score of this speaker is shown in Table 5.12.

z-score (concave) varying the speed of the pitch fall. 2 (words of /LH-L/ and /LH-Ø/) *2 (end pitches) *3 (shapes) = 12 stimuli of *wei-zhe* were created.

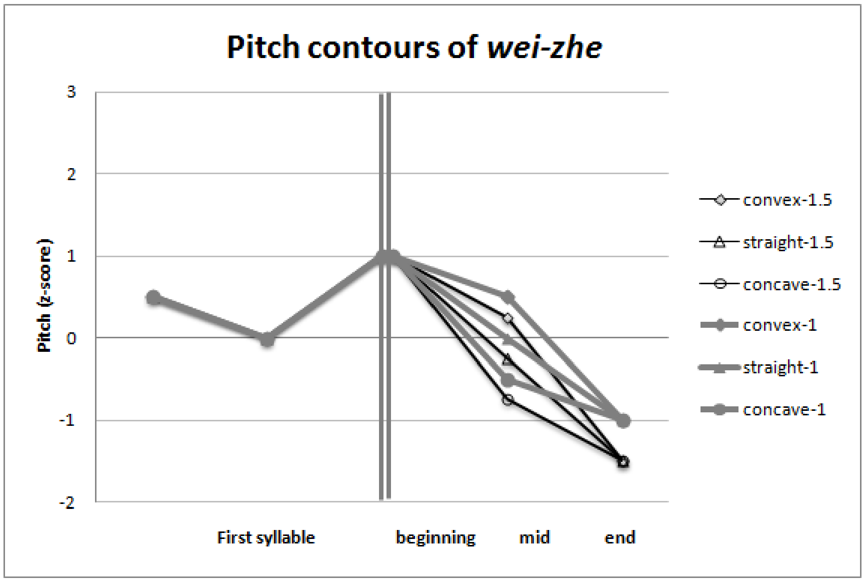


Figure 5.29 Schematized illustrations of synthesized stimuli *wei-zhe*

As mentioned above, the materials for Experiment 2 were recorded by a female TM speaker who was qualified for the production test. Her utterances of the *ji-zhe* pair and the *wei-zhe* pair, all four instances, were used to re-synthesize the stimuli. Table 5.12 shows the important z-score numbers converted to Hertz, which were used to re-synthesize the pitch. This table also includes the frequency converted to semitone z-score in order to relate to my findings in the production study. As shown in the table, the z-score number is basically similar to the semitone z-score number, except when the z-score was 2, -1.5, -2. Those manipulated were lower in semitone z-score.

Table 5.12 Conversion between z-score, frequency, and semitone z-score for the female speaker

	Pitch height								
z-score	2	1.5	1	0.5	0	-0.5	-1	-1.5	-2
Frequency (Hz)	295.9	277	258.1	239.2	220.3	201.4	182.5	163.6	144.7
Semitone z-score	1.83	1.44	1.02	0.56	0.07	-0.46	-1.05	-1.70	-2.43

12 varied *wei-zhe* tokens plus 18 varied *ji-zhe* tokens were tested in sentences. The 12+18=30 tested sentences were randomly mixed with 15 fillers produced by the same female speaker. The order of the stimuli is attached in Appendix H. The filler questions

were tonal minimal pairs that differ only in lexical tones such as *zhi-yuan* /H-LH/ ‘to support’ vs. /LH-LH/ ‘employee’. Listeners were able to answer the fillers without difficulty. The manipulated pairs and the fillers were played in the frame sentences *qing-shuo X X ba-ci* ‘please say X X eight times’.

5.2.3. Procedures

This experiment was organized in the same way as Experiment 1 and it was conducted after Experiment 1. Before this experiment, there was another practice section with two trial questions. The trial questions were produced by the same speaker who produced the stimuli for Experiment 2 in order to familiarize the listeners with her voice. The recording included a 150 millisecond (ms) beep, then 150 ms silence, and then the sentence, followed by a 300 ms silence. The sentence was then repeated and after that there was a 3-second silence for the subjects to choose the answer. Every 15 questions, there was a short break. Experiment 2 lasted about seven minutes.

5.2.4. Result

5.2.4.1. The *ji-zhe* pair

The results of the identification task with the re-synthesized *ji-zhe* are shown in Figure 5.30. The x-axis shows the three starting pitches of the *zhe* syllable, and the y-axis is the percentage of the stimuli identified as having a neutral tone. The bars of the three shades represent pitch fall, which is another variable. The result shows that both starting pitch and pitch fall seem to be factors in identifying syllables. The higher the starting pitch, the more likely it was to be identified as a neutral tone. As for the pitch fall, the pitch contours with smaller pitch falls seemed to have higher rates of being identified as a neutral tone.

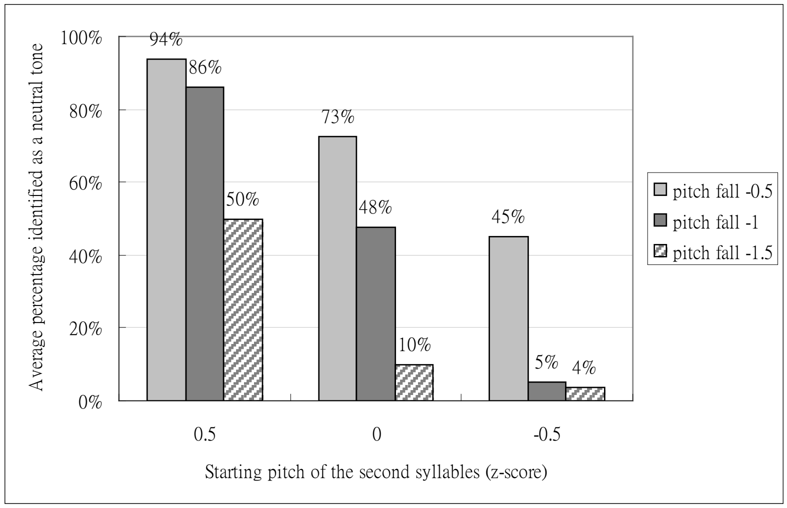


Figure 5.30 Result of Experiment 2: *ji-zhe*

However, the presentation of the bars in Figure 5.30 suggests that end pitch might be the crucial factor. Figure 5.31 reorganizes the results in terms of starting pitch and end pitch level. As shown in Figure 5.31, the end pitch levels (x-axis) affected the way the *ji-zhe* stimulus was perceived (y-axis). When the end pitch level was higher than -1 z-score, the stimuli were identified as having a neutral tone. When the end pitch level was lower than -1 z-score, the stimuli were identified as having a low tone. When the end pitch level was at -1 z-score, listeners were confused, regardless of the starting pitches or pitch falls.

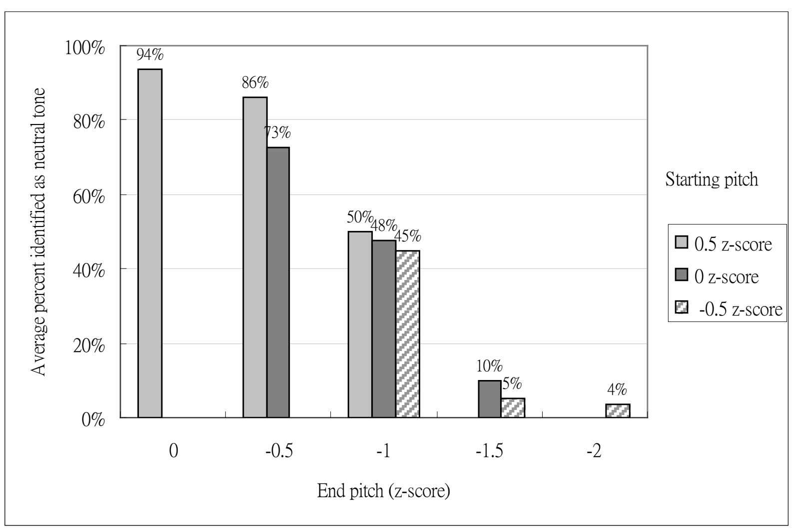


Figure 5.31 Result of *ji-zhe* by end pitch

A two-predictor logistic model was fitted to the data to test the relationship between the likelihood of being identified as a neutral tone and beginning pitch and end pitch level. Pitch fall was not included as a predictor because pitch fall is the outcome of the interaction between beginning pitch and end pitch level. The analysis showed that:

Predicted logit of (NEUTRAL TONE) = 2.95362 + 0.07938*(BEGINNING) + 3.20136*(END) + (-0.39227)*(BEGINNING:END)

As shown in the regression, the log of the odds of a stimulus being identified as a neutral tone was positively related to end pitch level ($p < .001$). However, the analysis also shows that the predictor of beginning pitch ($p = 0.917$) and the interaction ($p = 0.590$) were not statistically significant. In other words, the higher the end pitch level, the more likely it is that a stimulus would be identified as a neutral tone. For a one unit increase in end pitch level, the odds of being identified as a neutral tone increase by a factor of 24.57. On the other hand, beginning pitch and pitch fall were not determining factors in this experiment.

Aside from the end pitch level, some further characteristics of the original utterances seem to have been used in the listeners' identifications. Figure 5.32 shows the results of *ji-zhe* with the full vs. neutral tone of the un-manipulated utterances as a factor. The x-axis represents the 9 different pitch contours in the order of the end pitch level (primary) and the pitch fall (secondary). The light-shaded bars are the stimuli that were re-synthesized from the speaker's *ji-zhe* /HL-L/ 'reporter', and the dark-shaded bars are the stimuli that were re-synthesized from the speaker's *ji-zhe* /HL-Ø/ 'remembering'. With the pitch contours and the duration being equal, the vowel qualities or phonation types of the original (un-altered) utterances seemed to have an effect on the result.

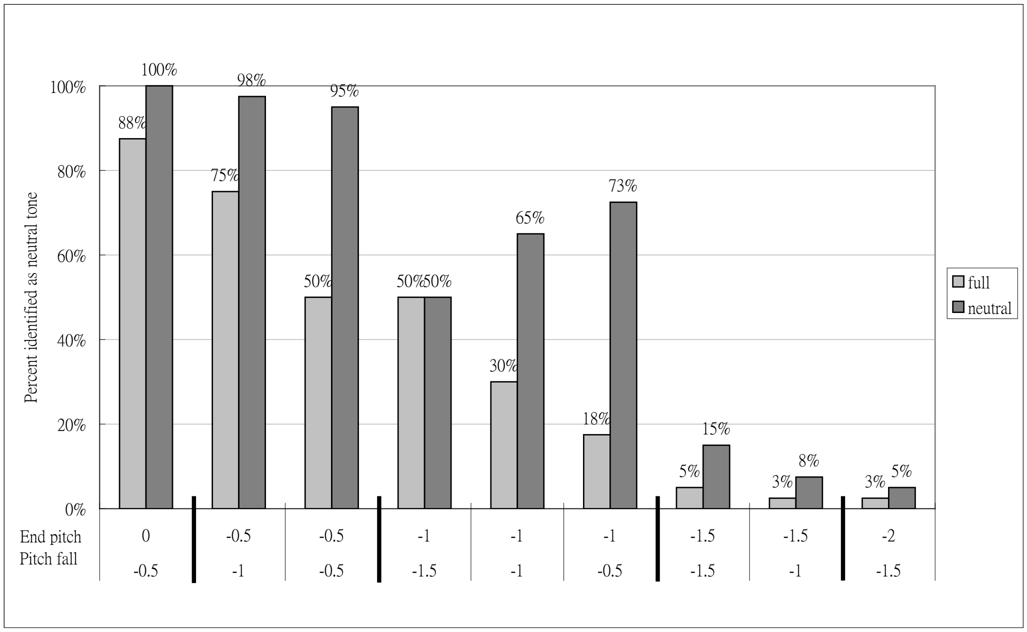


Figure 5.32 Result of *ji-zhe* by end pitch level, pitch fall, and segmental features (Full-tone original vs. Neutral-tone original)

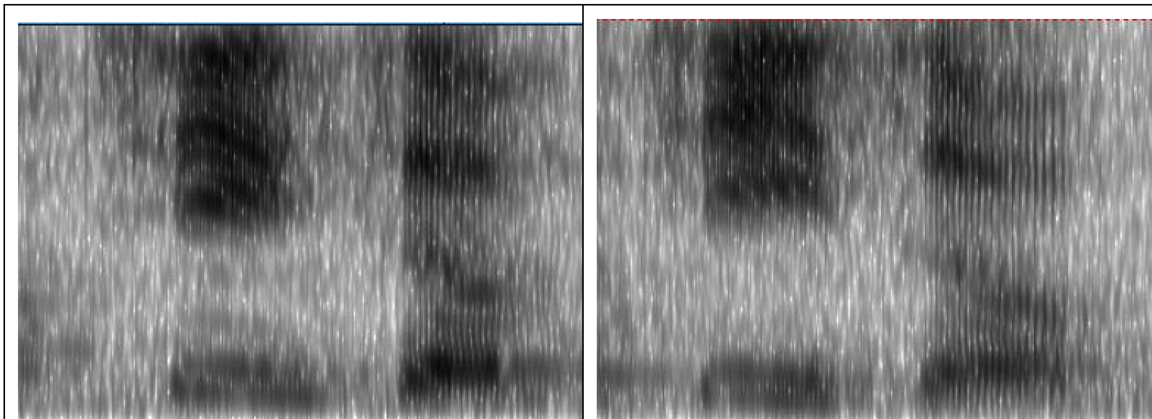
As shown in Figure 5.32, almost all the stimuli that were re-synthesized from *ji-zhe* /HL-Ø/ (neutral-tone original, dark-shaded bars) had higher percentages of being identified as a neutral tone. The effect was especially obvious when the end pitch level was -0.5 or -1 z-score. The effects of the original productions were statistically significant ($p < 0.01$) for most of the pitch contours with -0.5 and -1 z-score end pitch level (see Table 5.13 for the p-values). This shows that when the pitch cues were ambiguous, listeners used segmental features such as vowel quality or phonation type to distinguish the pair. The only exception was when the end pitch level was -1 and the pitch fall was -1.5 (steep initial pitch fall).

Furthermore, when the end pitch level was -0.5 or -1 z-score, the pitch fall seemed to affect the results too. As shown in Figure 5.32, with the end pitch level being the same, when the initial pitch fall was smaller, the listeners were more likely to notice the vowel quality/phonation differences between the original /HL-L/ and /HL-Ø/ *ji-zhe*. For example, when the end pitch level was -0.5 z-score, the difference between the original /HL-L/ and /HL-Ø/ *ji-zhe* was greater when the pitch fall was -0.5 z-score (50% vs. 95%) than when the pitch fall was -1 z-score (75% vs. 98%). Similarly, when the end pitch level was -1 z-score, the smaller pitch fall resulted in bigger differences between the original /HL-L/ and /HL-Ø/ *ji-zhe*. In other words, when the pitch contour was flatter, the listeners were more likely to notice the segmental qualities of the original (full vs. neutral) productions.

Table 5.13 Comparison between the results of original *ji-zhe* /HL-L/ and *ji-zhe* /HL-Ø/

End pitch level	0	-0.5	-0.5	-1	-1	-1	-1.5	-1.5	-2
Pitch fall	-0.5	-1	-0.5	-1.5	-1	-0.5	-1.5	-1	-1.5
Originally /HL-L/	88%	75%	50%	50%	30%	18%	5%	3%	3%
Originally /HL-Ø/	100%	98%	95%	50%	65%	73%	15%	8%	5%
Different (statistically)	no	yes	yes	no	yes	yes	no	no	no
Chi-square	3.41	6.75	18.12	0.05	8.47	22.27	1.25	0.26	0
p-value	0.07	<.01	<.01	0.82	<.01	<.01	0.26	0.61	1

The segmental qualities of the original utterances that were used by the listeners were mostly likely to be the vowel or phonation qualities of the *zhe* syllables. The vowel qualities of *zhe* in *ji-zhe* /HL-L/ and *ji-zhe* /HL-Ø/ were presented in Figure 5.27 in Section 5.2.2. It was likely that the listeners used the vowel-quality differences to help them make the decisions.



ji-zhe /HL-Ø/	*ji-zhe* /HL-L/

Figure 5.33 Spectrograms of the un-manipulated *ji-zhe* /HL-Ø/ and *ji-zhe* /HL-L/

Moreover, the listeners might have used voice quality as a perceptual cue too. Figure 5.33 shows the spectrograms of the un-manipulated *ji-zhe* /HL-Ø/ and *ji-zhe* /HL-L/. We can see that in the second half of the *zhe* syllable in *ji-zhe* /HL-L/, the striations were slightly farther apart. This is a feature of creaky voice. On the other hand, the *zhe* syllable in *ji-zhe* /HL-Ø/ did not show the creaky quality. This voicing difference might have been used by the listeners as a perceptual cue.

In summary, the end level level pitch of the second syllable in *ji-zhe* was the most important acoustic cue in identifying the tone. When the end pitch level was higher than -1 z-score, the stimulus was more likely to be identified as having a neutral tone. On the other hand, when the end pitch level was lower than -1 z-score, the stimulus was more likely to be identified as having a low tone. When the end pitch level was around -1 z-score, the listeners could not distinguish the pair based on the pitch information. As a result, the listeners utilized the voice or vowel quality, especially when the pitch was flatter.

5.2.4.2. **The *wei-zhe* pair**

The results of the perception test on re-synthesized *wei-zhe* are shown in Table 5.14. The stimuli had six different pitch contours, with three different pitch shapes and two different end pitches. Table 5.14 shows the average percentage identified as neutral tone.

Table 5.14 Average percentages identified as neutral tone for *wei-zhe*

Pitch shape / End pitch	convex	straight	concave
-1 z-score	76%	86%	**49%
-1.5 z-score	73%	**58%	19%

As shown in the table, the pitch shape and the end pitch seemed to influence how the stimuli were perceived. When the end pitch was low (-1.5 z-score) and the pitch shape was concave, the stimuli were generally identified as having a low tone (19%). With a straight pitch shape and a -1.5 z-score end pitch, or a concave pitch shape with a -1 z-score end pitch, the stimuli were confusing. Their identification rates were around the chance level statistically (58% and 49%). The three other pitch contours had higher chances of being identified as having a neutral tone (76%, 73%, and 86%). This shows that the listeners relied on both pitch shape and end pitch to distinguish the pair. When the end pitch was low enough (-1.5 z-score) and the pitch shape was concave (fall faster in the beginning), the stimuli were mostly identified as having a low tone.

Similar to the results of the *ji-zhe* pair, some characteristics of the original utterances seem to have been used in the listeners' identifications in the *wei-zhe* pair. As shown in Figure 5.34, with the pitch contours and the durations being the same, the stimuli that were re-synthesized from *wei-zhe* /LH-Ø/ (neutral-tone original, dark-shaded bars) tended to have higher percentages of identification as neutral tone compared to those that were re-synthesized from *wei-zhe* /LH-L/ (full-tone original, light-shaded bars). The only exception was concave pitch shape with -1 end pitch.

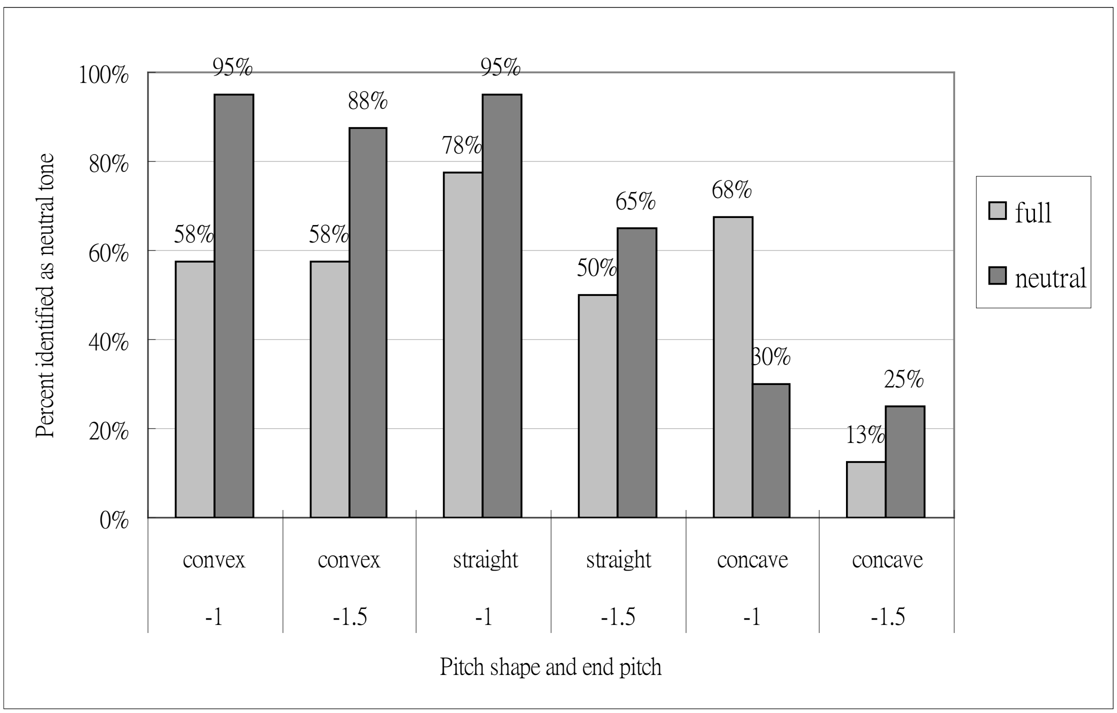


Figure 5.34 Results of *wei-zhe*

When the pitch shape was convex, regardless of the end pitches, the quality of the un-altered productions seems to influence the result. The stimuli with a convex pitch shape were generally perceived as a neutral tone when they were re-synthesized from the /LH-Ø/ *wei-zhe* (98% and 88%), but the originally /LH-L/ *wei-zhe* stimuli were ambiguous (both 58%). The effects of the un-altered productions were statistically significant [End pitch -1: χ^2= 13.53, df = 1, p < .01; End pitch -1.5: χ^2= 7.59, df = 1, p

$< .01$]. The stimuli that were re-synthesized from the /LH-L/ *wei-zhe* possibly had some characteristics of a low tone. Therefore when these stimuli had a neutral-tone-like pitch shape, they were ambiguous for the listeners.

When the pitch shape was straight, the end pitch was important to the perception of tone. The differences between the two end pitches were significant for both the stimuli that were re-synthesized from /LH-L/ [$\chi^2 = 5.41$, df=1, $p < 0.05$] and the stimuli that were re-synthesized from /LH-Ø/ [$\chi^2 = 9.45$, df=1, $p < 0.01$]. Listeners were much more confident with the stimuli that were re-synthesized from /LH-Ø/. When the end pitch was -1 z-score, both of the stimuli from /LH-L/ and /LH-Ø/ were identified as having a neutral tone (78% and 95%). However, the difference between the original /LH-L/ and /LH-Ø/ was not significant [$\chi^2 = 3.79$, df = 1, $p = 0.051$]. Similarly, with the straight pitch contour and -1.5 z-score end pitch, both of the stimuli from /LH-L/ and /LH-Ø/ were confusing (50% and 65%). The difference between the two identification rates was not significant either [$\chi^2 = 1.28$, df = 1, $p = 0.26$]. When the pitch shape was straight, the influences of the original production' segmental qualities were not as strong.

When the pitch shape was concave and the end pitch was -1.5 z-score, both of the stimuli were perceived as a low tone (13% and 25%), regardless of the original segmental quality. The result of the concave shape with the end pitch -1 z-score was rather surprising. The stimulus that was re-synthesized from /LH-L/ (68%) was perceived as a neutral tone more than the stimulus that was re-synthesized from /LH-Ø/ (30%); the differences were significant [$\chi^2 = 9.81$, df =1, $p < 0.01$]. This rather contradictory result may be due to the experiment design. This originally /LH-Ø/ *wei-zhe* stimulus was tested immediately after the *ji-zhe* stimulus with the neutral-tone identification rate of 95%. It might be that this originally /LH-Ø/ stimulus (-1 z-score, concave) was not sufficiently neutral-tone-like in comparison to the previous stimulus. Similarly, the originally /LH-L/ segment was after the *ji-zhe* stimulus with a neutral-tone identification rate of 18%. The originally /LH-L/ stimulus (-1 z-score, concave) might appear to be less low-tone-like right after a stimulus with very strong low-tone acoustic features, which might have been the pitch or vowel/voice quality.

The results above show that the listeners used some segmental characteristics to help them make the decision, especially when the pitch contours were convex (slow-falling).

The features that they might have used were likely vowel quality and voice quality because the durations and the pitch contours were controlled to be the same. The vowel-quality difference between the *zhe* syllables in *wei-zhe* /LH-L/ and *wei-zhe* /LH-Ø/ was shown in Figure 5.27 in Section 5.2.2. The low-tone *zhe* is more back than the neutral-tone *zhe*. Figure 5.35 shows the spectrograms of the un-altered *wei-zhe* /HL-Ø/ and *wei-zhe* /HL-L/. However, there were no obvious voicing quality differences shown in the spectrograms.

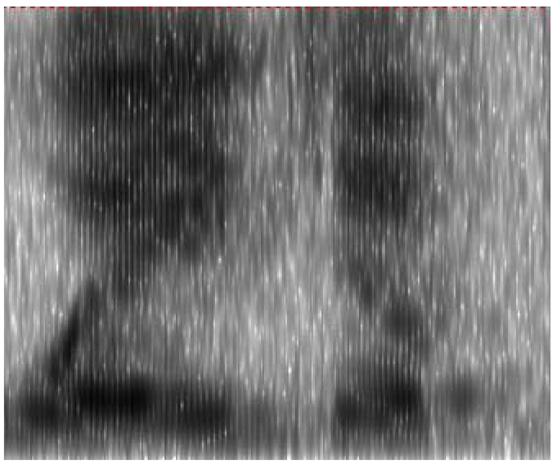
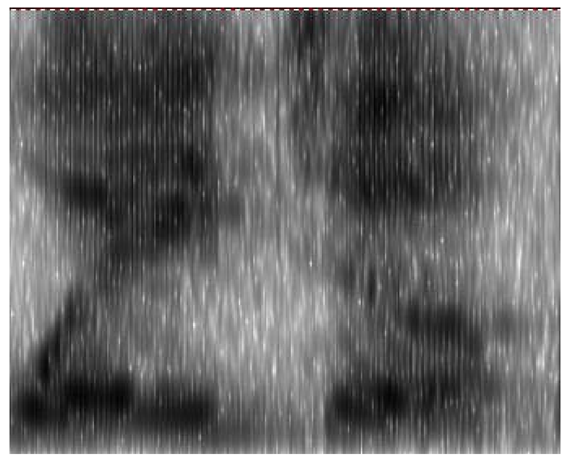
	
wei-zhe /LH-Ø/	*wei-zhe* /LH-L/

Figure 5.35 Spectrograms of the un-manipulated *wei-zhe* /HL-Ø/ and *wei-zhe* /HL-L/

However, there were voice quality differences when I examined the spectral tilts. Figure 5.36 shows spectral slices from the second half of the *zhe* rimes in the un-manipulated *wei-zhe* /HL-Ø/ and *wei-zhe* /HL-L/. From left to right, the first harmonic (H1), the second harmonic (H2), the first formant (F1), and the second formant (F2) are circled to show the spectral tilt measurement points in both spectral slices. Typically, the amplitudes of H2-H1, F1-H1, and F2-H1 are larger for creaky voice than for modal voice (See Gordon & Ladefoged (2001) for a review). My measurements show that in *wei-zhe* /LH-Ø/, H2-H1 was 0.9 dB; F1-H1 was 7.7dB; and F2-H1 was -7 dB. On the other hand, in *wei-zhe* /LH-L/, H2-H1 was 7.7dB; F1-H1 was 11.3 dB; and F2-H1 was 6.5 dB. The measurements of the spectral tilt show that the *zhe* syllable in *wei-zhe* /LH-L/ had more creaky quality than the *zhe* syllable in *wei-zhe* /LH-Ø/. It was likely that aside from the vowel quality, the listeners used this voice quality as a perceptual cue.

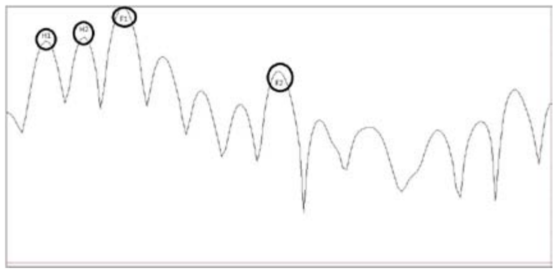
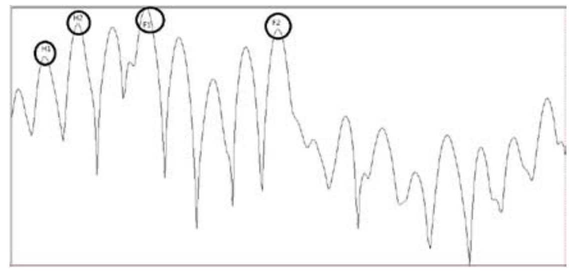
	
wei-zhe /LH-Ø/	*wei-zhe* /LH-L/

Figure 5.36 Spectral slices of the second half of the *zhe* rimes in the un-manipulated *wei-zhe* /HL-Ø/ and *wei-zhe* /HL-L/

5.2.5. Summary

In summary, the end pitch appears to be the most important acoustic cue in distinguishing the low tone from the neutral tone. As shown in both the *wei-zhe* and the *ji-zhe* experiments, when the end pitch was lower than -1 z-score, which is around -1 semitone z-score, the stimulus was more likely to be identified as having a low tone. The results from the *ji-zhe* stimuli further showed that the perceptual boundary between the low tone and the neutral tone was -1 z-score, as a stimulus was more likely to be identified as having a neutral tone when the end pitch was higher than -1 z-score. This experiment shows that the TM listeners have the knowledge that the neutral tone in TM has a low target, but that target is not as low as that of the low tone.

On the other hand, the pitch contour is also very important. The results from the *wei-zhe* stimuli showed that a concave pitch contour (falling more sharply in the beginning) is also necessary in order to be identified as a low tone. This result conforms to the finding of Fon et al. (2004). This result also suggests that TM listeners know that the target of the neutral tone is not reached as quickly as the target of the low tone.

Also, when the pitch contour was ambiguous, the listeners utilized vowel quality or phonation type as a cue. This can be seen in the results of the *ji-zhe* stimuli with -1 z-score end pitch and the *wei-zhe* stimuli with a convex contour. The result from the first perception experiment also shows similar findings (See Section 5.1.5).

This chapter shows that TM listeners do have the knowledge that the neutral tone is different from the low tone. TM listeners know the TM neutral tone has a mid-low target and the target is reached more slowly than that of the low tone. This chapter also shows that the listeners use a variety of acoustic cues including duration and vowel/phonation quality, in addition to pitch, to help them identify the tones. However, the results also show that the speakers do not always make the distinction. There is cross-speaker variation and lexical variation. The neutral tone is merging with the low tone.

CHAPTER 6. CONCLUSION

6.1. Summary

This dissertation has contributed to the study of the realization of rhythm in TM, focusing on the quality of its “unstressed” (neutral-tone) syllables. It examined the acoustic properties of the TM neutral-tone syllables, specifically their pitch behavior in different contexts; it compared the stressed (full-tone) with the unstressed (neutral-tone) syllables in TM in terms of duration and pitch; and it included perception experiments to see whether TM listeners can distinguish the TM stressed and unstressed syllables. The perception experiments further examined the perceptual cues used by TM listeners to distinguish the stressed and unstressed syllables.

Chapter One introduced my research questions. In Chapter Two, I first presented Standard Mandarin phonology, and after introducing the dialect of Taiwan Mandarin, I discussed the phonology of Taiwan Mandarin. I showed that many prescriptive neutral-tone syllables in TM are treated as their original lexical tones.

In Chapter Three, I examined the acoustic features of the neutral tone in TM in three production tests. Six male and six female TM native speakers were asked to produce 1) frequent and meaningful disyllabic words with a neutral tone, 2) productive neutral-tone syllables in novel phrases/words, and 3) meaningful sentences with 2-3 consecutive neutral-tone syllables. All three tests show that the neutral tone in Taiwan Mandarin behaves differently from the neutral tone in Standard Mandarin. In Standard Mandarin, the pitch of the neutral tone is determined by the preceding lexical tone. /H-Ø/ is [55-2], /LH-Ø/ is [35-3], /L-Ø/ is [21-4], and /HL-Ø/ is [51-1]. However, my findings for TM show that generally, unlike Standard Mandarin, there was no post-L rising, and the pitch contours of a neutral tone slowly reached to a mid-low target, which is around -1 semitone z-score. The mid-low pitch target of the neutral tone in TM was further demonstrated by the consecutive neutral-tone production test. In Standard Mandarin, the pitch of consecutive neutral tones slowly reaches to a mid target, and the influences of the preceding tones are significant even at the end of the third neutral tone. On the other hand, in this study of Taiwan Mandarin, each neutral tone in the consecutive neutral-tone productions had its own target, and the influences of the preceding tones were generally not found at the second or the third neutral tone. There were, however, some cases where the subjects produced Standard Mandarin-like pitch patterns. These instances were sporadic. They were only found in reduplications, but not found in novel occurrences. These pitch patterns were lexicalized only in some words and only by certain speakers. There were also cases showing that TM speakers would produce a lexical tone when the original tone was lexically obvious. For example, the neutral tone -*tou* 'DIM' has the same character as *tou* /LH/ 'head' and the character of the neutral tone -*men* 'PL' has a phonetic element of *men* /LH/ 'door/gate'. As a result, the neutral-tone -*tou* and some instances of the neutral-tone -*men* were found with a rising-tone-like pitch contour. Chapter Three thus established that the neutral tone of TM differs from the neutral tone of SM. The neutral-tone syllables of TM are not toneless; they have a pitch target.

Chapter Four further compared the neutral tone of TM to the full tones of TM in production. The same subjects were asked to read disyllabic words/phrases that contained corresponding syllables with the same rimes and the similar full tones. Comparison between the neutral-tone and the rising-tone corresponding syllables shows that the neutral-tone syllable -*tou* /Ø/ does not destress in Taiwan Mandarin. The prescriptive neutral-tone -*tou* /Ø/ was acoustically similar to the rising tone *tou* /LH/ in terms of pitch contour and duration. The rest of the neutral-tone syllables were compared to their low-tone corresponding syllables. The results show that the influence of the preceding tone was overcome faster in the low-tone syllables than in the neutral-tone syllables. Also, TM low tone has a lower pitch than TM neutral tone, and the low-tone syllables had a higher chance of having creaky voice than the neutral-tone syllables. The end pitches of the neutral tone ranged between -0.5 ~ -1 semitone z-score, while the end pitches of the low tone were generally around -1.5 semitone z-score. The pitch differences between the neutral tone and the low tone at the late position were mostly statistically significant, which suggests that the neutral tone in Taiwan Mandarin should be characterized as a fifth tone, with a mid-low pitch target, because it does have a pitch target and its target is different from the low tone. As for duration, the comparison shows that there was no general syllable reduction found in Taiwan Mandarin. Of all the neutral-tone syllables, only those with rimes [ə] or [ɨ] were shorter than their low-tone corresponding syllables. The rest of the neutral-tone syllables (rimes [i, u, ən]) were not statistically different from their low-tone corresponding syllables. The comparison of S2/S1 ratios also shows similar patterns. Even though some of the syllables were statistically shorter than their low-tone counterparts, the durational differences were about 20~30 ms, and the differences of the S2/S1 ratios were about 0.1 ~ 0.2, which was smaller than the S2/S1 ratio difference in SM (0.42).

Lastly, Chapter Five investigated whether the neutral-tone syllables are perceptually distinct from the full-tone syllables. 40 subjects participated in two experiments in the form of identification tasks. The first experiment used six unaltered tonal minimal pairs produced by two male and two female TM speakers as the test tokens. The results show that the perceptual distinctions made by the TM speakers between the low tone and the neutral tone were very weak. The TM listeners could only distinguish 20% of the unaltered pairs. This suggests that the neutral tone is on the verge of merging with the low tone. The second experiment used two minimal pairs with manipulated pitch contours from one TM speaker to further investigate the acoustic cues used by TM listeners to distinguish the neutral tone from the low tone. The results show that concave pitch shape and an end pitch lower than -1 semitone z-score were the perceptual cues to identify the low tone. The listeners also used duration and vowel/voice quality as cues: shorter duration was a cue for the neutral tone and creaky voice quality was a cue for the low tone.

This study found considerable variation in the neutral tone in Taiwan Mandarin. The perception test shows that some speakers produced the distinction between the neutral tone and the low tone more robustly than others. The comparison shows that some pitch contours of neutral-tone syllables after /HL/ or /L/ overlapped with the pitch contours of the low-tone syllables. Many neutral-tone syllables also exhibited the feature of creaky voice, which is a perceptual cue for the low tone. The production test even found evidence of a TM speaker treating the neutral tone as a low tone phonologically; this speaker applied Tone 3 Sandhi when attaching the neutral-tone suffixes to a low tone syllable in the novel occurrences test.

6.2. Limitations of this work

There are some limitations in this work. First, in the production test I have included six male and six female subjects. Although I tried to find a rather homogeneous group and I collected the background information of the subjects (see Section 3.1.3), considerable speaker variation appeared in my data and the variation could not be explained by the background information I collected. Other useful background information might include relative language abilities of Mandarin or Southern Min, detailed information on participants' language behaviors, and phonological indices such as the accuracy of their Mandarin retroflexes or the accuracy of their Southern Min initial /g/.

Due to the limitation of time and subjects, I included unaltered pairs by only four speakers and I manipulated only the pitch contours in the perception test. The influences of duration and voice quality were shown in the results. However, these factors were not designed as independent variables. Vowel quality and intensity also might be possible perceptual cues but I did not include them as variables. I focused on the most likely perceptual cues.

The data elicited in this study is all read speech. The tested words/phrases were either elicited in a frame sentence, or they were fitted in meaningful written sentences. Although this is the most efficient and controlled ways to elicit the data, it is not the most natural speech. It would be important to also compare the words in spontaneous speech.

In the perception test, I used the low tone to compare to the low-target neutral tone because they are perceptually the most similar and there are minimal pairs available. However, in a forced identification task, it was unclear whether the listeners actually identified the neutral-tone item or they chose the neutral-tone item because the stimuli did not sound like a low-tone item. It would be helpful if there were minimal pairs that contrast falling tone or rising tone to the low-target neutral tone to help really characterize the feature of the low-target neutral tone. Unfortunately there are not many tonal minimal pairs that contrast lexical tones with the neutral tone.

6.3. Conclusion

This study has examined several aspects of the development of the “unstressed” syllables in Taiwan Mandarin. As speakers of a more syllable-timed dialect, the TM speakers tend to assign a lexical tone to the prescriptively unstressed (neutral-tone) syllables. As shown in Chapter Two, the TM speakers simply did so by not destressing the second syllables in the disyllabic words. The results from Chapter Three and Four also show that the diminutive suffix *-tou* /Ø/ was produced with a rising tone because its character is the same as *tou* /LH/ ‘head’.

The rhythmic difference between SM and TM is also reflected in the remaining neutral-tone syllables. Most of these neutral-tone syllables are bound morphemes, and their original syllables (or lexical tones) before destressing are difficult for the speaker to trace. So the original lexical tones were not available for the TM speakers. On the other hand, the TM speakers failed to apply the SM process of tone neutralization on these syllables. Instead, they received the input of the neutral tone and analyzed the neutral tone as a fifth tone with a mid-low pitch target because, except for post-L rising, the rest of the pitch contours are all falling. As for duration, only those neutral-tone syllables with [ə] or [ɨ] rimes are shorter because they are either inherently short ([ɨ]) or lexicalized as a reduced vowel ([ə]). The rest of the neutral-tone syllables all show no signs of syllable reduction. Therefore, it seems that there are still some unstressed/reduced syllables in TM. They only have [ə] or [ɨ] rimes, but they do not show a neutral tone.

Descriptions of the phonology of Taiwan Mandarin therefore need to be modified. Many of the “unstressed” syllables should be classified as having an underlying lexical tone. For example the diminutive suffix *-tou* /Ø/ should be phonologically *tou* /LH/ and *ma-ma* /H-Ø/ ‘mother’ should be *ma-ma* /H-H/. Furthermore, the remaining “unstressed” syllables are not toneless. Instead, they should be phonologized as having a fifth tone, which has a mid-low pitch target. The innovated fifth tone does not have a pitch target that is as low as the low tone, and the speed reaching to the target is not as fast as the low tone. The segmental phonology of Taiwan Mandarin does not change because the previously “unstressed” syllables in TM do not have obvious vowel reductions.

The pitch target of the innovated fifth tone is acoustically very similar to the low tone, especially when it is after /HL/ or /L/ where the end pitch of the previous syllable is generally lower than the fifth tone after /H/ or /LH/. The pitch contours of the fifth tone after /HL/ or /L/ could reach to around -1 semitone z-score or sometimes exhibit creaky voice. As shown in the perception tests, the end pitch of -1 semitone z-score is confusing for the listeners, and creaky voice is one of the perceptual cues for the low tone. Therefore, the fifth tone can be similar to the low tone. Furthermore, the pitch contours of the fifth tone shown in the production tests are the average pitch contours. It is very likely that some pitch contours that were below the average could be as low as the low-tone pitch contours too. This similarity between the fifth tone and the low tone causes confusion for most of the TM listeners, as shown in Chapter Five. The perception test shows that even though currently the TM listeners still have the knowledge that the “neutral” tone is not as low as the low tone and the “neutral-tone” syllables are shorter than the low-tone syllable, many TM speakers do not make the distinction consistently. This suggests that the fifth tone might be merging with the low tone. We have already seen that one speaker who treated the fifth tone as the low tone phonologically. It would not be surprising if the next generation of TM speakers does not distinguish between the two tones.

This dissertation is the first study that carefully examines the acoustic quality and perceptual nature of the neutral tone in Taiwan Mandarin. I argue that Taiwan Mandarin, as a more syllable-timed dialect, possesses many “unstressed” syllables that are not really unstressed, and that speakers use different ways to re-assign tones to these “unstressed” syllables. Some neutral-tone syllables have returned to their original lexical tones, and some have acquired a new tone (mid-low), which should be characterized as a fifth tone. This study also shows that the neutral tone (the fifth tone) is on the verge of merging with the low tone. While some speakers still maintained a clear phonemic distinction between the fifth tone and the low tone, others produced outcomes that were not distinguishable by the TM listeners.

APPENDICES

Appendix A. Linguistic background questionaires

受試者語言背景

The linguistic background of the subject

年齡:

Age:

性別:

Gender:

從小成長的城市:

City you grew up:

會說的語言:	國語	台語	客家話	英語	其他:
Languages spoken	Mandarin	Southern Min	Hakka	English	Others
父母親會說的語言:	國語	台語	客家話	英語	其他:
Languages spoken by parents:	Mandarin	Southern Min	Hakka	English	Others
從小在家使用的語言:	國語	台語	客家話	英語	其他:
Languages used at home	Mandarin	Southern Min	Hakka	English	Others

您認為您的台語:	流利	普通	不大流利
You consider your Southern Min:	fluent	average	not so fluent
您認為您的國語:	非常標準	還算標準	不大標準
You consider your Mandarin:	very standard	fairly standard	not so standard

平常在什麼情況下會使用台語

When do you use Southern Min in your daily life:

在家裡	與朋友交談	學校	日常生活買東西
at home	chat with friend	at school	shopping

Appendix B. A list of tested sentences for disyllabic words/phrases

請說出家八次	請說錄音八次	請說露營八次	請說那麼八次
請說聽著八次	請說文字八次	請說讀著八次	請說君子八次
請說孽子八次	請說出嫁八次	請說難捨八次	請說佛門八次
請說禿子八次	請說施捨八次	請說嬸嬸八次	請說成本八次
請說舌頭八次	請說蛇頭八次	請說蕃薯八次	請說多麼八次
請說相思八次	請說兔子八次	請說磚頭八次	請說吃了八次
請說椅子八次	請說蚊子八次	請說課本八次	請說四個八次
請說校花八次	請說相似八次	請說熨斗八次	請說去了八次
請說鬼們八次	請說作者八次	請說酸的八次	請說北門八次
請說不捨八次	請說笑話八次	請說辣的八次	請說北斗八次
請說徒弟八次	請說鑷子八次	請說碰頭八次	請說人們八次
請說坐著八次	請說土地八次	請說傷者八次	請說他們八次
請說零錢八次	請說記著八次	請說亂扯八次	請說公投八次
請說弟弟八次	請說領錢八次	請說人格八次	請說什麼八次
請說梨子八次	請說我媽八次	請說閒扯八次	請說五個八次
請說讀者八次	請說李子八次	請說買了八次	請說不捨八次
請說支援八次	請說枕頭八次	請說評審八次	請說大門八次
請說李子八次	請說職員八次	請說發抖八次	請說八個八次
請說知道八次	請說想著八次	請說教授們八次	請說一個八次
請說我嗎八次	請說地底八次	請說基本八次	
請說實施八次	請說指導八次	請說甜的八次	
請說整頭八次	請說鬼門八次	請說桌子八次	
請說禮字八次	請說事實八次	請說家門八次	
請說衝鋒八次	請說重逢八次	請說苦的八次	
請說甲子八次	請說夾子八次	請說施捨八次	
請說冬季八次	請說動機八次	請說我們八次	
請說記者八次	請說違者八次	請說怎麼八次	
請說失去八次	請說市區八次	請說表格八次	
請說兔子八次	請說以茲八次	請說杯葛八次	
請說搬家八次	請說半價八次	請說拉扯八次	
請說圍著八次	請說吐籽八次	請說念頭八次	
請說出沒八次	請說觸摸八次	請說叔叔八次	
請說馬頭八次	請說碼頭八次	請說來了八次	

Appendix C. A list of tested sentences for novel occurrences

suffix	obsolete characters	made-up sound and meanings	given sentences
-le (-PERF)	呦	*xiu* /H/ 'to eat'	他昨天__了三碗炒飯。 *ta zuotian* __***le*** *sanwan chaofan* 'He ___ three bowls of fried rice yesterday'
	哸	*xiu* /LH/ 'to drink'	他昨天__了三杯紅茶。 *ta zuotian* __***le*** *sanbei hongcha* 'He ___ three cups of black tea yesterday'
	哂	*xiu* /L/ 'to read'	他昨天__了三本漫畫。 *ta zuotian* __***le*** *sanben manhua* 'He ___ three comic books yesterday'
	啁	*xiu* /HL/ 'to buy'	他昨天__了三隻牙刷。 *ta zuotian* __***le*** *sanzhi yashua* 'He ___ three toothbrush yesterday'
-zhe (-DUR)	抸	*qiao* /H/ 'to hold'	他手上__著一本雜誌。 *ta shoushang* __***zhe*** *yiben zazhi* 'He is ___ a magazine on his hand.'
	挠	*qiao* /LH/ 'to wear'	她頭上__著一朵紅花。 *ta toushang* __***zhe*** *yiduo honghua* 'She is ___ a red flower on her head.'
	撟	*qiao* /L/ 'to carry'	他肩上__著一個背包。 *ta jianshang* __***zhe*** *yige beibao* 'He is __ a backpack on his shoulder.'
	撨	*qiao* /HL/ 'to wear'	她腰上__著一條皮帶。 *ta yaoshang* __***zhe*** *yitiao pidai* 'She is ___ a belt on her waist.'

-men (-PL)	僡	*hui* /H/ ‘junior student’	那裡有一群__們在吃飯。 *nali you yiqun* __***men*** *zai chifan* ‘There is a group of __ eating.’
	倕	*hui* /LH/ ‘high school student’	那裡有一群__們在唱歌。 *nali you yiqun* __***men*** *zai changge* ‘There is a group of ___ singing.’
	侩	*hui* /L/ ‘farmer’	那裡有一群__們在抗議。 *nali you yiqun* __***men*** *zai kangyi* ‘There is a group of ___ protesting.’
	佹	*hui* /HL/ ‘elementary school student’	那裡有一群__們在上課。 *nali you yiqun* __***men*** *zai shangke* ‘There is a group of __ attending classes.’
-de (-GEN)	傋	*jiu* /H/ ‘junior student’	左邊那個是__的哥哥。 *zuobian nage shi* __***de*** *gege* ‘The one on the left is __’s brother.’
	俦	*jiu* /LH/ ‘old man’	右邊那台是__的車子。 *youbian natai shi* __***de*** *chezi* ‘The one on the right is ___’s car.’
	伷	*jiu* /L/ ‘adolescent girl’	旁邊那個是__的妹妹。 *pangbian nage shi* __***de*** *meimei* ‘The one on the side is __’s sister.’
	攸	*jiu* /HL/ ‘colleague’	前面那間是__的新家。 *qianmian najian shi* __***de*** *xinjia* ‘The one in front is __’s new house.’

Appendix D. A list of tested sentences for consecutive neutral-tone syllables

Set 1. X(H/LH/L/HL) + N (reduplication) + N (POSS) + Y(HL/ L)					
ta	*shuo*	***X-X-de***	***hao[L]/zhong[HL]yao***	***duo***	***le***
he	said	X-REDUP -POSS	good/important	more	ASP
'He said that X's is much better/more important.					

ta shuo ***mama de hao*** *duo le*
'He said that the mother's is much better'
ta shuo ***popo de hao*** *duo le*
'He said that the mother-in-law's is much better.'
ta shuo ***nainai de hao*** *duo le*
'He said that the grandma's is much better.'
ta shuo ***meimei de hao*** *duo le*
'He said that the sister's is much better.'

ta shuo ***mama de zhongyao*** *duo le*
'He said that the mother's is much more important.'
ta shuo ***popo de zhongyao*** *duo le*
'He said that the mother-in-law's is much more important.'
ta shuo ***nainai de zhongyao*** *duo le*
'He said that the grandma's is much more important.'
ta shuo ***meimei de zhongyao*** *duo le*
'He said that the sister's is much more important.'

他說**媽媽的**好多了
他說**婆婆的**好多了
他說**奶奶的**好多了
他說**妹妹的**好多了

他說**媽媽的**重要多了
他說**婆婆的**重要多了
他說**奶奶的**重要多了
他說**妹妹的**重要多了

Set 2. X(H/LH/L/HL) + N (reduplication) + N (PL) + N (POSS) + Y(HL/L)

ta	*shuo*	***X-X-men-de***	***hao[L]/zhong[HL]yao***	***duo le***
he	said	X-reduplication-PL-POSS	good/important	more ASP

‘He said that Xs’ are much better/more important.

ta shuo ***mama men de hao*** *duo le*

‘He said that the mothers’ are much better.’

ta shuo ***popo men de hao*** *duo le*

‘He said that the grandmas’ are much better.’

ta shuo ***nainai men de hao*** *duo le*

‘He said that the grandmas’ are much better.’

ta shuo ***meimei men de hao*** *duo le*

‘He said that the sisters’ are much better.’

ta shuo ***mama men de zhongyao*** *duo le*

‘He said that the mothers’ are much more important.’

ta shuo ***popo men de zhongyao*** *duo le*

‘He said that the grandmas’ are much more important.’

ta shuo ***nainai men de zhongyao*** *duo le*

‘He said that the grandmas’ are much more important.’

ta shuo ***meimei men de zhongyao*** *duo le*

‘He said that the sisters’ are much more important.’

他說**媽媽們的**好多了
他說**婆婆們的**好多了
他說**奶奶們的**好多了
他說**妹妹們的**好多了

他說**媽媽們的**重要多了
他說**婆婆們的**重要多了
他說**奶奶們的**重要多了
他說**妹妹們的**重要多了

Set 3. X(H/LH/L/HL) + N (PERF) + N (NML.) + Y(HL/L)
ta *xihuan* ***X-le*** ***de*** ***na[HL]****-CL* *noun* 3.S likes verb-PERF NML that noun. ‘He likes the one that ~.’ *nage* ***X-le*** ***de*** ***bi[L]****jiao* ~ that verb-PERF NML. compare adjective. ‘The one that /verb/ is more /adjective/.’
ta xihuan ***chai[H]-le de na****jian dian.* ‘He likes the shop that was torn down.’ *ta xihuan* ***tao[LH]-le de na****ge ren.* ‘He likes the person that escaped.’ *ta xihuan* ***pao[L]-le de na****ge ren.* ‘He likes the person that ran away.’ *ta xihuan* ***xiao[HL]-le de na****ge ren.* ‘He likes the one that smiled.’
nage ***chai[H]-le de bi****jiao pianyi* ‘The one that was torn down is cheaper’ *nage* ***tao[LH]-le de bi****jiao kelian* ‘The one that was escaped is more pathetic’ *nage* ***pao[L]-le de bi****jiao qingsong* ‘The one that ran away is more relaxed’ *nage* ***xiao[HL]-le de bi****jiao kuaile* ‘The one that smiled is happier’
他喜歡**拆了的**那間店 他喜歡**逃了的**那個人 他喜歡**跑了的**那個人 他喜歡**笑了的**那個人 那個**拆了的**比較便宜 那個**逃了的**比較輕鬆 那個**跑了的**比較可憐 那個**笑了的**比較快樂

Set 4. X(H/LH/L/HL) + N (DUR) + N (NML.) + Y(HL/L) *ta xihuan* ***X-zhe-de*** ***na[HL]****-ge ren* 3.S. like verb-DUR-NML. that-CL person 'He likes the person that /verb/.' *nage* ***X-zhe-de*** ***bi[L]****jiao* ~ that verb-DUR-NML. more adjective 'That /verb/-ing one is more ~.'
ta xihuan ***dun[H]-zhe-de na****ge ren.* 'He likes the person that was squatting.' *ta xihuan* ***xian[LH]-zhe de na****ge ren.* 'He likes the person that was free.' *ta xihuan* ***tang[L]-zhe de na****ge ren* 'He likes the person that was lying down.' *ta xihuan* ***zhan[HL]-zhe de na****ge ren.* 'He likes the person that was standing.'
nage ***dun[H]-zhe-de bi****jiao suibian* 'That squatting one is more casual.' *nage* ***xian[LH]-zhe-de bi****jiao landuo* 'That free one is lazier.' *nage* ***tang[L]-zhe-de bi****jiao xuruo* 'That lying down one is weaker.' *nage* ***zhan[HL]-zhe-de bi****jiao renzhen* 'That standing one is more earnest.'
他喜歡**蹲著的**那個人 他喜歡**閒著的**那個人 他喜歡**躺著的**那個人 他喜歡**站著的**那個人 那個**蹲著的**比較隨便 那個**閒著的**比較討厭 那個**躺著的**比較懶惰 那個**站著的**比較認真

Appendix E. Full-full and full-neutral tonal minimal pairs

	pinyin	IPA	full-neutral (FN)	full-full (FF)
1	*ma-tou*	[ma-tʰou]	碼頭 [L-Ø] ‘pier’	馬頭 [L-LH] ‘horse head’
2	*she-tou*	[ʂə-tʰou]	舌頭 [LH-Ø] ‘tongue’	蛇頭 [LH-LH] ‘snake head’
3	*zheN-tou*	[ʈʂən-tʰou]	枕頭 [L-Ø] ‘pillow’	整頭 [L-LH] ‘whole head’
4	*yi-zi*	[ji-tsɨ]	椅子 [L-Ø] ‘chair’	以茲 [L-H] ‘hereby’
5	*wen-zi*	[wən-tsɨ]	蚊子 [LH-Ø] ‘mosquito’	文字 [LH-HL] ‘character’
6	*li-zi*	[li-tsɨ]	李子 [L-Ø] ‘plum’	禮字 [L-HL] ‘manner’
7	*jia-zi*	[tɕja-tsɨ]	夾子 [LH-Ø] ‘clip’	甲子 [LH-L] ‘60 years’
8	*tu-zi*	[tu-tsɨ]	兔子 [HL-Ø] ‘rabbit’	吐籽 [HL-L] ‘spit the seed’
9	*nie-zi*	[nie-tsɨ]	鑷子 [HL-Ø] ‘tweezer’	孽子 [HL-L] ‘unfilial child’
10	*zuo-zhe*	[tswo-ʈʂə]	坐著 [HL-Ø] ‘sitting’	作者 [HL-L] ‘author’
11	*ji-zhe*	[tɕi-ʈʂə]	記著 [HL-Ø] ‘remembering’	記者 [HL-L] ‘reporter’
12	*wei-zhe*	[wei-ʈʂə]	圍著 [LH-Ø] ‘surrounding’	違者 [LH-L] ‘offender’
13	*du-zhe*	[tu-ʈʂə]	讀著 [LH-Ø] ‘reading’	讀者 [LH-L] ‘reader’
14	*gui-men*	[kue-mən]	鬼們 [L-Ø] ‘ghosts’	鬼門 [L-LH] ‘ghost gate’
15	*di-di*	[ti-ti]	弟弟 [HL-Ø] ‘brother’	地底 [HL-L] ‘underground’
16	*wo-ma*	[wo-ma]	我嗎 [L-Ø] ‘is it me?’	我媽 [L-H] ‘my mother’

Appendix F. Tonal minimal pairs as fillers in the perception test

	pinyin	IPA		
1	*chu-jia*	[ʈʂu-tɕia]	出家 [H-H] ‘pabbajja’	出嫁 [H-HL] ‘to marry’
2	*tu-zi*	[t^{h}u-tsɨ]	秃子 [H-Ø] ‘bold man’	兔子 [HL-Ø] ‘rabbit’
3	*xiang-si*	[ɕiaŋ-sɨ]	相思 [H-H] ‘to miss’	相似 [H-HL] ‘similar’
4	*xiao-hua*	[ɕiaw-hwa]	校花 [HL-H] ‘prom queen’	笑話 [HL-HL] ‘joke’
5	*tu-di*	[t^{h}u-ti]	徒弟 [LH-HL] ‘disciple’	土地 [L-HL] ‘land’
6	*ling-qian*	[liŋ-tɕ^{h}jən]	零錢 [LH-LH] ‘change’	領錢 [L-LH] ‘to withdraw money’
7	*li-zi*	[li-tsɨ]	梨子 [LH-Ø] ‘pear’	李子 [L-Ø] ‘plum’
8	*zhi-yuan*	[ʈʂɨ-ɰən]	支援 [H-LH] ‘support’	職員 [LH-LH] ‘employee’
9	*zhi-dao*	[ʈʂɨ-taw]	知道 [H-HL] ‘to know’	指導 [L-HL] ‘to direct’
10	*shi-shi*	[ʂɨ-ʂɨ]	實施 [LH-LH] ‘to implement’	事實 [HL-LH] ‘fact’
11	*chong-feng*	[ʈʂ^{h}oŋ-fəŋ]	衝鋒 [H-H] ‘a charge’	重逢 [LH-LH] ‘reunion’
12	*dong-ji*	[toŋ-tɕi]	冬季 [H-HL] ‘winter’	動機 [HL-H] ‘motive’
13	*shi-qu*	[ʂɨ-tɕ^{h}y]	失去 [H-HL] ‘lose’	市區 [HL-H] ‘downtown’
14	*ban-jia*	[pan-tɕja]	搬家 [H-H] ‘to move’	半價 [HL-HL] ‘half price’
15	*chu-mo*	[ʈʂ^{h}u-mwo]	出沒 [H-HL] ‘to haunt’	觸摸 [HL-H] ‘to touch’
16	*lu-yin*	[lu-jin]	錄音 [HL-H] ‘to record’	露營 [HL-LH] ‘to camp’

Appendix G. Order of the stimuli in Perception Experiment 1

Question 1-40: drawn from the first pool

1-10				21-30			
F2	椅子	*yizi*	L-Ø	F6	鬼門	*guimen*	L-LH
M3	作者	*zuozhe*	HL-L	M3	舌頭	*shetou*	LH-Ø
F2	徒弟		Filler	F6	甲子	*jiazi*	LH-L
M6	坐著	*zuozhe*	HL-Ø	M6	讀著	*duzhe*	LH-Ø
F6	失去		Filler	F6	舌頭	*shetou*	LH-Ø
M3	出嫁		Filler	M3	以茲	*yizi*	L-H
F2	支援		Filler	F6	讀者	*duzhe*	LH-L
M6	蛇頭	*shetou*	LH-LH	M3	表格		Filler
F6	以茲	*yizi*	L-H	F2	坐著	*zuozhe*	HL-Ø
M3	土地		Filler	M3	讀者	*duzhe*	LH-L
11-20				**31-40**			
F6	坐著	*zuozhe*	HL-Ø	F2	甲子	*jiazi*	LH-L
M6	觸摸		Filler	M6	鬼們	*guimen*	L-Ø
F6	錄音		Filler	F2	相思		Filler
M6	動機		Filler	M3	領錢		Filler
F6	實施		Filler	F2	鬼門	*guimen*	L-LH
M6	以茲	*yizi*	L-H	M6	夾子	*jiazi*	LH-Ø
F2	舌頭	*shetou*	LH-Ø	F2	讀著	*duzhe*	LH-Ø
M6	重逢		Filler	M6	半價		Filler
F6	知道		Filler	F2	校花		Filler
M3	甲子	*jiazi*	LH-L	M3	鬼們	*guimen*	L-Ø

Question 41-80: the second pool

41-50				61-70			
F2	梨子		Filler	F2	出家		Filler
M6	鬼門	*guimen*	L-LH	M6	甲子	*jiazi*	LH-L
F6	搬家		Filler	F6	衝鋒		Filler
M6	事實		Filler	M3	笑話		Filler
F2	鬼們	*guimen*	L-Ø	F2	作者	*zuozhe*	HL-L
M3	坐著	*zuozhe*	HL-Ø	M3	蛇頭	*shetou*	LH-LH
F2	蛇頭	*shetou*	LH-LH	F2	夾子	*jiazi*	LH-Ø
M3	相似		Filler	M3	椅子	*yizi*	L-Ø
F2	以茲	*yizi*	L-H	F2	零錢		Filler
M6	舌頭	*shetou*	LH-Ø	M3	讀著	*duzhe*	LH-Ø
51-60				**71-80**			
F6	作者	*zuozhe*	HL-L	F6	夾子	*jiazi*	LH-Ø
M3	指導		Filler	M6	錄影		Filler
F6	冬季		Filler	F6	鬼們	*guimen*	L-Ø
M6	椅子	*yizi*	L-Ø	M6	讀者	*duzhe*	LH-L
F2	讀者	*duzhe*	LH-L	F6	出沒		Filler
M6	作者	*zuozhe*	HL-L	M3	鬼門	*guimen*	L-LH
F6	蛇頭	*shetou*	LH-LH	F6	讀著	*duzhe*	LH-Ø
M6	錄音		HL-H	M6	市區		Filler
F6	椅子	*yizi*	L-Ø	F2	禿子		Filler
M3	支援		Filler	M3	夾子	*jiazi*	LH-Ø

Appendix H. Order of the stimuli in Perception Experiment 2

No.		word	Original production	Pitch manipulation			
				start	pitch fall	end pitch	shape
1	出家	Filler					
2	圍著	*weizhe*	LH-Ø			-1	straight
3	圍著	*weizhe*	LH-Ø			-1	concave
4	李子	Filler					
5	記著	jizhe	HL-L	-0.5	-1.5		
6	圍著	*weizhe*	LH-L			-1	straight
7	事實	Filler					
8	出沒	Filler					
9	記著	*jizhe*	HL-L	0.5	-1		
10	記著	*jizhe*	HL-L	0	-1.5		
11	圍著	*weizhe*	LH-Ø			-1.5	straight
12	記著	*jizhe*	HL-Ø	0.5	-1.5		
13	零錢	Filler					
14	搬家	Filler					
15	冬季	Filler					
16	記著	*jizhe*	HL-Ø	0	-1.5		
17	記著	*jizhe*	HL-Ø	0	-1		
18	土地	Filler					
19	記著	*jizhe*	HL-Ø	-0.5	-0.5		
20	記著	*jizhe*	HL-L	-0.5	-0.5		
21	圍著	*weizhe*	LH-L			-1	concave
22	記著	*jizhe*	HL-L	0	-0.5		
23	重逢	Filler					
24	記著	*jizhe*	HL-Ø	0	-0.5		
25	圍著	*weizhe*	LH-L			-1.5	straight
26	記著	*jizhe*	HL-L	0	-1		
27	圍著	*weizhe*	LH-L			-1	convex
28	禿子	Filler					
29	記著	*jizhe*	HL-Ø	0.5	-1		
30	圍著	*weizhe*	LH-Ø			-1	convex

31	支援	Filler					
32	記著	*jizhe*	HL-L	-0.5	-1		
33	記著	*jizhe*	HL-L	0.5	-0.5		
34	知道	Filler					
35	圍著	*weizhe*	LH-Ø			-1.5	concave
36	記著	Filler	HL-L	0.5	-1.5		
37	圍著	*weizhe*	LH-L			-1.5	concave
38	相似	Filler					
39	記著	*jizhe*	HL-Ø	-0.5	-1.5		
40	失去	Filler					
41	圍著	*weizhe*	LH-L			-1.5	convex
42	記著	*jizhe*	HL-Ø	-0.5	-1		
43	圍著	*weizhe*	LH-Ø			-1.5	convex
44	校花	Filler					
45	記著	*jizhe*	HL-Ø				

REFERENCES

Abercrombie, D (1967). *Elements of general phonetics.* Edinburgh: Edinburgh University Press.

Auer, P. (1993). Is a rhythm-based typology possible? A study of the role of prosody in phonological typology. *KontRI (Kontextualisierung durch Rhythmus und Intonation) Working paper (Universität Konstanz)* 21.

Benton, Matthew, Dockendorf, Liz , Jin, Wenhua, Liu, Yang & Edmondson, Jerold A. (2007). The continuum of speech rhythm: computational testing of speech rhythm of large corpora from natural Chinese and English speech. *ICPhS XVI* Saarbrücken.

Betran, A. P. (1999). Prosodic typology: on the dichotomy between stress-timed and syllable-timed languages. *Language Design* 2, 103-130.

Bybee, Joan (2001). *Phonology and Language Use.* Port Chester, NY, USA: Cambridge University Press.

Chang, Mei-yu (1996). Language use and language attitudes among Taiwanese elementary school students in native language instruction programs: a study on language maintenance, language shift, and language planning in Taiwan. *Department of Literacy, Culture, and Language Education* Bloomington: Indiana University.

Chang, Yueh-chin (1998). Taiwan Mandarin Vowels: An Acoustic Investigation. *Tsing Hua journal of Chinese Studies* 28, 255-274.

Chao, Yuen-Ren (1956). Tone, intonation, singsong, chanting, recitative tonal composition and atonal composition in Chinese. In Moris Halle (ed.), *For Roman Jakobson* 52-59. The Hague: Mouton.

Chao, Yuen-Ren (1968). *A Grammar of Spoken Chinese.* University of California: Berkeley and Los Angeles.

Chen, Chung-yu (1991). The nasal endings and retroflexed Initials in Peking Mandarin: instability and the trend of Changes. *Journal of Chinese Linguistics* 19, 139-171.

Chen, Xiaoyan (2004). On the Necessity and Consistency Principles of Qingsheng Words:Metric Research on Qingsheng Words in Modern Chinese Dictionary. *Yuyan Wenzi Yingyong* 1, 112-119.

Chen, Y & Xu, Y (2006). Production of weak elements in speech-evidence from F0 patterns of neutral tone in standard Chinese. *Phonetica* 63, 47-75.

Cheng, Robert L (1985). A comparison of Taiwanese, Taiwan Mandarin, and Peking Mandarin. *Language* 61, 352-377.

Chung, Karen (2006). Hypercorrection in Taiwan Mandarin. *Journal of Asian Pacific Communication* 16, 197-214.

Dasher, R. & Bolinger, D. (1982). On pre-accentual lengthening. *The Journal of the International Phonetic Association* 12, 58-69.

Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics* 11, 51-62.

Deng, Dan, Shi, Feng & Lu, Shinan (2008). Contrastive analysis on tone between Putonghua and Taiwan Mandarin *Chinese Journal of Acoustics* 27.

Deterding, D. (2001). The measurement of rhythm: a comparison of Singapore and British English. *Journal of Phonetics* 29, 217-230.

Duanmu, San (2000). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Duanmu, San (2007). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.

Fon, Janice & Chiang, Wen-Yu (1999). What does Chao have to say about tones? - a case study of Taiwan Mandarin. *Journal of Chinese Linguistics* 27, 15-37.

Fon, Janice, Chiang, Wen-Yu & Cheung, Hintat (2004). Production and perception of the two dipping tones (Tone 2 and Tone 3) in Taiwan Mandarin. *Journal of Chinese Linguistics* 32, 249-281.

Frota, Sónia & Vigário, Marina (2001). On the correlates of rhythmic distinctions: The European/Brazilian Portuguese case. *Probus* 13, 247-275.

Ghazali, Salem, Hamdi, Rym & Barkat, Melissa (2002). Speech rhythm variation in Arabic dialects. *Speech Prosody 2002* 331-334. Aix-en-Provence, France. http://aune.lpl.univ-aix.fr/sp2002/pdf/ghazali-hamdi-barkat.pdf.

Gordon, Matthew & Ladefoged, Peter (2001). Phonation types: a cross-linguistic overview. *Journal of Phonetics* 29, 383-406.

Grabe, E. & Low, E.L. (2002). Durational variability in speech and the rhythm class hypothesis. In N. Warner & C. Gussenhoven (eds.), *Papers in laboratory phonology 7* 515-546. Berlin: Mouron de Gruyter.

Her, One-Soon (2010). 論台灣華語的在地化 On the Indigenization of Taiwan Mandarin. *澳門語言學刊 Journal of Macau Linguistics Association* 35, 19-29.

Hombert, J.-M. (1977). Consonant types, vowel height and tone in Yoruba. *Studies in African Linguistics* 8, 173-190.

Hombert, J.-M., Ohala, J. J. & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language* 55, 37-58.

Howie, John (1976). *An Acoustic Study of Mandarin Tones and Vowels.* London: Cambridge University Press.

Hsu, Hui-chuan (2006). Revisiting tone and prominence in Chinese. *Language and Linguistics* 7, 109-137.

Hsu, Hui-ju (2005). Some Aspects of Phonological Leveling in Taiwan Mandarin. *Department of English* Taipei: National Taiwan Normal University.

Hsu, Hui-ju & Tse, John Kwock-Ping (2007). Syllable-Final Nasal Mergers in Taiwan Mandarin—Leveled but Puzzling. *Concentric: Studies in Linguistics* 33, 1-18.

Hsu, Hui-ju & Tse, John Kwock-Ping (2009). The Tonal Leveling of Taiwan Mandarin: A Study in Taipei. *Concentric: Studies in Linguistics* 35, 225-244.

Huang, Shuang-Fan (1993). *Yuyan, shehui yu zuqun yishi : Taiwan yuyan shehui xue de yanjiu [Language, society, and ethnic identity-a study of Taiwan language sociology].* Taipei: Crane Publishing Company.

Huang, Shuang-Fan (2000). Language, identity, and conflict: a Taiwanese study. *International Journal of the Sociology of Language* 143, 139-149.

Jing, Song (2002). *Xiandai hanyu qingsheng dongtai yanjiu [A study of Modern Chinese neutral tone].* Beijing: Minzu Chubanshe.

Kubler, Cornelius C (1979). Some differences between Taiwan Mandarin and 'Textbook Mandarin'. *Journal of the Chinese Language Teachers Association* 14, 27-39.

Kubler, Cornelius C (1985). The influence of Southern Min on the Mandarin of Taiwan. *Anthropological Linguistics* 27, 156-176.

Kuo, Yun-Hsuan (2005). New dialect formation: the case of Taiwanese Mandarin. University of Essex, UK PhD dissertation.

Lee, Wai-Sum & Zee, Eric (2008). Prosodic characteristics of the neutral Tone in Beijing Mandarin. *Journal of Chinese Linguistics* 36, 1-29.

Lehiste, I. (1977). Isochrony reconsidered. *Journal of Phonetics* 5, 253-263.

Li, Jennifer (2005). The preliminary study about neutral tone: dialect effect between North Official Mandarin speakers in China and Taiwan Mandarin speakers (A). *Journal of the Acoustic Society of America* 117, 2457-2457.

Li, Weimin (1981). Shi lun qingsheng he zhong yin [A preliminary discussion on stressless and stressed syllables]. *Zhongguo Yuwen* 1, 35-40.

Liao, Chao-Chih (2000). Changing dominant language use and ethnic equality in Taiwan since 1987. *International Journal of the Sociology of Language* 143, 165-182.

Liao, Chao-chih & Lii-Shih, Yu-hwei E (1993). University Undergraduates' Attitudes on

Code-Mixing and Sex Stereotypes. *Pragmatics* 3, 425-449.

Lin, Hua & Wang, Qian (2007). Mandarin rhythm: an acoustic study. *Journal of Chinese Language and Computing* 17, 127-140.

Lin, Maocan & Yan, Jingzhu (1980). Beijinghua qingsheng de shengxue xingzhi [Acoustic characteristics of neutral tone in Beijing Mandarin]. *Fangyan* 3, 166-178.

Lin, Maocan & Yan, Jingzhu (1990). Puthonghua qingsheng yu qingzhong yin [The neutral tone and stress in Mandarin]. *yuyan Jaoxue Yu Yanjiu* 3, 88-104.

Lin, Maocan & Yan, Jingzhu (1991). Tonal coarticulation patterns in quadrasyllabic words and phrases of Mandarin. *the XIIth international congress of phonetic sciences* 242-245.

Lin, Tao (1985). Tantao Beijinghua qingyin xingzhi de chubu shiyan (On neutral tone in Beijing Mandarin). In Shuangbao Hu (ed.), *Beijing yuyin shiyanlu (Working papers in experimental phonetics)* 1-26. Beijing: Beijing Daxue chubanshe (Peking University Press).

Lin, Tao & Wang, William S-Y. (1984). Shengdiao ganzhi wenti (Perception of tones). *Zhongguo Yuyan Xuebao* 2, 59-69.

Lo, Chao-chin (1991). 鬧熱滾滾:大眾傳播語的分合 The divergence and the convergence of Mass Communication. *guowen tiandi* 7, 12-15.

Low, E.L., Grabe, E. & Nolan, F. (2000). Quantitative characterisations of speech rhythm: 'syllable-timing' in Singapore English. *Language and Speech* 43, 377-401.

Mo, Ruo-ping (2000). Taiwan on the Brink of Reversing Language Shift: Its Current Development and contributory Factors. Purdue University.

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

O'Rourke, Erin (2008). Correlating speech rhythm in Spanish: evidence from two Peruvian dialects. In Joyce Bruhn de Garavito & Elena Valenzuela (eds.), *Selected Proceedings of the 10th Hispanic Linguistics Symposium* 276-287. Somerville, MA: Cascadilla Proceedings Project.

Pike, K. (1945). *The intonation of American English.* Ann Arbor: University of Michigan Press.

Ramsey, S. Robert (1987). *The languages of China.* Princeton: Princeton University Press.

Ramus, F., Nespor, M. & Mehler, J. (1999). Correlates of linguistic rhythm classes. *Cognition* 73, 265-292.

Roach, P. (1982). On the distinction between "stress-timed" and "syllable-timed" languages. In D. Crystals (ed.), *Linguistic controversies* London: Edward Arnold.

Russo, Michela & Barry, William J (2004). Interaction between segmental structure and rhythm: a look at Italian dialects and regional standard Italian. *Folia Linguistica* 38, 277-296.

Sandel, Todd (2003). Linguistic capital in Taiwan: The KMT's Mandarin language policy and its perceived impact on language practices of bilingual Mandarin and Tai-gi speakers. *Language in Society* 32, 523-551.

Sanders, Robert (2005). The Phonetic and Phonological Features of Tone 3 in Taiwan Mandarin. *二十一世紀華語文中心營運策略與教學國際研討會*: http://online.sfsu.edu/~hdomizio/LinLaoshi/899/35_The%20Phonetic%20and%20Phonological%20Features%20of%20Tone%203%20in%20Taiwan%20Mandarin.pdf.

Shen, Xiaonan (1990). Tonal coarticulation in Mandarin. *Journal of Phonetics* 18, 281-295.

Shih, Chilin & Sproat, R. (1992). Variations of the Mandarin rising tone. *the IRCS workshop on prosody in natural speech* 193-200. Philadelphia: The Institute for Research in Cognitive Science , University of Pennsylvania.

Su, Hsi-Yao (2001). Code-switching between Mandarin and Taiwanese in Three Telephone Conversations: The Negotiation of Interpersonal Relationships among Bilingual Speakers in Taiwan. *Texas Linguistic Forum* 44, 430-446.

Swihart, De-An Wu (2003). The two Mandarins: Putonghua and Guoyu. *Journal of the Chinese Language Teachers Association* 38, 103-118.

Torgerson, Richard C., Jr (2005). A comparison of Beijing and Taiwan Mandarin tone register: an acoustic analysis of three native speech styles. Birgham Young University.

Tsao, Feng-fu (1999). The language planning situation in Taiwan. *Journal of Multilingual and Multicultural Development* 20, 328-375.

Tsao, Feng-fu (2000). 臺式日語與臺灣國語--百年來在臺灣發生的兩個語言接觸實例 Taiwanized Japanese and Taiwan Mandarin--Two Case Studies of Language Contact during the Past Hundred Years in Taiwan. *Hanxue yanjiu <Chinese Studies>* 18, 273-297.

Tse, John Kwock-Ping (1983). Bilingualism in University Students in the Republic of China. *Ying Yu Yen Chiu Chi K'an [Studies in English Literature and Linguistics]* 9, 178-192.

Tse, John Kwock-Ping (1992). Production and perception of syllable final [n] and [(eng)] in Mandarin Chinese: an experiment study. *Ying Yu Yen Chiu Chi K'an / Studies in English Literature and Linguistics* 18, 143-156.

Tse, John Kwock-Ping (1998). 臺灣地區年輕人ㄓㄔㄕ與ㄗㄘㄙ真的不分嗎? . *huawen shijie <Chinese World>* 90, 1-7.

Tseng, Chin-Chin (2004). Prosodic Properties of Intonation in Two Major Varieties of Mandarin Chinese: Mainland China vs. Taiwan. *International Symposium on Tonal Aspects of Languages* Bei-jing, China.

Tseng, Hsin-I (2003). The syntax structures of contemporary Taiwanese Mandarin. National Taiwan Normal University.

Tseng, Shu-Chuan & Huang, Yun-Ru (2010). A socio-phonetic analysis of Taiwan Mandarin interview speech. *DiSS-LPSS Joint Workshop* 67-70. Tokyo, Japan.

Van den Berg, Marinus (1986). *Language planning and language use in Taiwan : a study of language choice behavior in public settings.* Taipei: Crane.

Wei, Jennifer M.Y. (2003). Codeswitching in Campaigning Discourse: The Case of Taiwanese President Chen Shui-bian. *Language and Linguistics* 4, 139-165.

Xu, Shirong (1980). *Putonghua yuyin zhishi [Phonology of Standard Chinese].* Beijing: Webzu Gauge Chubanshe.

Xu, Yi (1994). Production and perception of coarticulated tones. *Journal of the Acoustical Society of America* 95, 2240-2253.

Xu, Yi (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics* 25, 61-83.

Xu, Yi (1999). F0 peak delay: when, where, and why it occurs. In John Ohala (ed.), *International Congress of Phonetic Science* 1881-1884. San Francisco.

Xu, Yi (2001). Sources of tonal variation in connected speech. In Hana Triskova (ed.), *Tone, stress and rhythm in spoken Chinese* Journal of Chinese Linguistics Monograph Series No. 17.

Young, Russell (1988). Language maintenance and language shift in Taiwan. *Journal of Multilingual and Multicultural Development* 9, 323-338.

Zhu, Xiaonong Sean (1999). *Shanghai tonetics* (LINCOM studies in Asian linguistics, München: LINCOM Europa.