

PHONOLOGICAL IDENTITY OF THE NEUTRAL-TONE SYLLABLES IN TAIWAN MANDARIN: AN ACOUSTIC STUDY

Karen HUANG

The University of Auckland, New Zealand
k.huang@auckland.ac.nz

Abstract

Taiwan Mandarin, one of the more syllable-timed dialects of Mandarin, has fewer unstressed syllables than Standard Mandarin. Acoustic analyses show that the supposedly unstressed syllables—neutral-tone syllables—in Taiwan Mandarin behave differently from those of Standard Mandarin. Unlike Standard Mandarin, these syllables do not raise their pitch after Tone 3. They have a distinct static mid-low pitch target and the target is implemented with a stronger articulatory strength. Moreover, acoustic analyses demonstrate that not all of these “unstressed syllables” are unstressed. The phonetic evidence suggests that these neutral-tone syllables should be analyzed as unaccented rather than unstressed in Taiwan Mandarin. These unaccented syllables are only lexically marked, and their pitch is neutralized into a mid-low tone. This study sheds light on how rhythm can affect stress and accent in a lexical tone language.

Keywords: Taiwan Mandarin; neutral tone; rhythm; lexical tone; stress; neutralization

Povzetek

Tajvanska kitajščina, ena ritmično bolj zlogovnih različic kitajščine, pozna manj nepoudarjenih zlogov kot standardna kitajščina. Akustične analize kažejo, da zlogi v nevtralnem tonu, torej tisti nepoudarjeni, v tajvanski kitajščini izkazujejo drugačne lastnosti kot tisti v standardni kitajščini; za nizkim tonom ne pride do tonskega višanja. Z razliko od standardne kitajščine so ti zlogi izgovorjeni z značilnim enakomernim srednje nizkim tonom in so izgovorjeni z večjo jakostjo. Poleg tega je iz akustičnih analiz moč razbrati, da ne izkazujejo vsi nepoudarjeni zlogi akustičnih lastnosti nepoudarjenosti. Zato nekatere fonetične raziskave podajajo predlog, da bi takšne zloge v tajvanski kitajščini definirali kot nenaglašene in ne kot nepoudarjene. Nenaglašeni zlogi so opredeljeni samo pomensko, njihov ton je nevtraliziran v enakomerno srednje nizko višino. Akustične analize poleg tega razkrivajo, da niso vsi zlogi, ki naj bi bili nepoudarjeni, tudi v resnici nepoudarjeni. Fonetični dokazi nakazujejo, da bi morali v tajvanski kitajščini zloge v nevtralnem tonu obravnavati kot nenaglašene, ne pa kot nepoudarjene. Nenaglašeni zlogi so označeni le leksikalno, pri čemer se njihova tonska višina nevtralizira v srednje nizki ton. Ta študija prikazuje, kako lahko ritem vpliva na poudarek in naglas v jezikih z leksikalnim tonom.

Ključne besede: tajvanska kitajščina; nevtralni ton; ritem; leksikalni ton; poudarek; nevtralizacija



1 Introduction

Languages use accent to divide an utterance into shorter phrases. Some languages only rely on pitch to mark prominence, such as Japanese (pitch accent); some other languages, namely stress accent languages, use a range of phonetic material to mark prominence (Beckman, 1986; Ewen & Hulst, 2001). Stress accent languages differ in which acoustic correlates (and in what combinations) they use to mark prominence. Stressed syllables are associated with one or more of the following acoustic correlates: higher fundamental frequency (f_0) (Garellek & White, 2015; Gordon & Applebaum, 2010; Lieberman, 1960; Llisterri, Machuca, de la Mota, Riera, & Ríos, 2003), greater intensity (Gordon & Applebaum, 2010; Kochanski, Grabe, Coleman, & Rosner, 2005; Prieto & Ortega-Llebaria, 2006; Sluijter & Van Heuven, 1996), longer duration (Arvaniti, 2000; Crystal & House, 1988; Fry, 1955; Sluijter & Van Heuven, 1996; Williams, 1985), as well as vowel quality (De Jong & Zawaydeh, 1999; Gordon & Applebaum, 2010; Ortega-Llebaria & Prieto, 2007).

Since f_0 is one of the acoustic correlates of stress, one might wonder whether lexical stress is compatible with tone, especially the complex tone systems in which f_0 contours are utilized to contrast every syllable. The answer is yes. Lexical stress can be found in complex lexical tone languages, mostly located in East Asia. Although some complex tone languages are argued to have no lexical stress, e.g. Cantonese (Bauer & Benedict, 1997) and Southern Vietnamese (Brunelle, 2017), other complex tone languages have been documented to have lexical stress. In these cases, the contrast of stress is manifested through the reduction of unstressed syllables—reduced duration and different degrees of tone neutralization are often found. For example, in Burmese, the unstressed syllables are called “minor” syllables. A minor syllable has a shorter duration and its vowel is neutralized to [ə]. The pitch of the minor syllable has been described as “variable” (Bradley, 1982) and shows no high f_0 peak (Gruber, 2011). In Thai, it has been reported that the five lexical tones are neutralized into three tonal registers (low, mid, high) when they are unstressed (Potisuk, Gandour, & Harper, 1994, 1996). However, this analysis is still controversial (Gandour, Tumtavitikul, & Sattamnuwong, 1999; Moren & Zsiga, 2006). Similarly in Nanchang Chinese, a Gan dialect, the lexically unstressed syllables also show tone neutralization in which the five lexical tones are neutralized into two different pitch patterns due to shorter rime durations (J. Liu & Zhang, 2012).

1.1 Standard Mandarin

Standard Mandarin is perhaps the best-studied case of lexical stress in complex tone systems. The unstressed syllables are mostly lexically determined in Standard Mandarin (SM) and they are referred to as having a “neutral tone” (Chao, 1968, p. 36). An unstressed syllable reduces in SM (Chao, 1933, 1956, 1968; Luo & Wang, 1957). The unstressed vowel centralizes towards a schwa and its coda nasal is usually deleted

while the vowel is nasalized (Duanmu, 2000, p. 256; Lin & Yan, 1990). As a result, an unstressed syllable is also shorter in duration. The mean duration of unstressed syllables is about 50%-60% of the stressed ones (J. Cao, 1986; Lee & Zee, 2008; Lin & Yan, 1980). As for intensity, an unstressed syllable does not necessarily have a weakened maximum intensity compared to a stressed syllable or its preceding syllable (J. Cao, 1986; Lin & Yan, 1980), and the intensity curves seem to co-vary with the pitch contours (Lee & Zee, 2008). Furthermore, there are lenition processes found only in unstressed syllables. In connected speech, an unstressed vowel can be devoiced or deleted when the vowel is high and after a fricative, an aspirated stop, or an affricate (Duanmu 2000:257). The initial of an unstressed syllable in connected speech also often undergoes lenition (Chao, 1968, p. 38; Duanmu, 2000, pp. 255-256; S. Xu, 1980, p. 159).

Stress also affects Mandarin on the suprasegmental level. There are four lexical tones. Tone 1, Tone 2, Tone 3, and Tone 4 have high level /H/ [˥], rising /LH/ [˨˨˥], dipping (low falling) /L/ [˨˨˨], and high falling /HL/ [˨˥˨] pitch contours, respectively. The underlying tone of an unstressed syllable in SM is not realized, and its pitch is determined by the preceding tone (Chao, 1932; Luo & Wang, 1957). Therefore an unstressed syllable is considered toneless (Duanmu, 2007, pp. 242-243) and has what is normally termed neutral tone /∅/¹. The earlier descriptions of the unstressed syllables were mainly impressionistic (Chao, 1932, 1933; Kratochvil, 1968), but Chao's description (1968, p. 27) is still widely cited: the pitch of the neutral tone is *half-low* after Tone 1 /H/, *middle* after Tone 2 /LH/, *half-high* after Tone 3 /L/, and *low* after Tone 4 /HL/. The instrumental studies showed similar results, though the findings slightly varied in terms of the contour details (Cheng, 1973; Dreher & Lee, 1968; Gao, 1980; Lin & Yan, 1980; Shen, 1990). The most recent acoustic studies showed that the pitch contours of the neutral tone were [˨, ˨, ˨˨, ˨] after the four lexical tones /H, LH, L, HL/ [˥, ˨˨˥, ˨˨˨, ˨˥˨], respectively (Lee & Zee, 2008). The pitch contours of a neutral tone appear to be an extension of the preceding tone (Z. Li, 2003). However, Chen and Xu (2006) found that when there are consecutive unstressed syllables, the pitch contours slowly approach a mid-level target over the course of these syllables.

At least in Standard Mandarin, scholars disagree on how lexical stress (or lack of stress) in complex tone languages should be analyzed (L. Liu, 2002). Unlike other non-tonal languages where primary stress in a word is marked, it seems that the unstressed syllables are marked in these complex tone languages. Because of that, many analyses marked the unstressed syllables as having a short tone and specify the short duration as a feature of the underlying tone since it is common in complex tone languages to have the categories of both long and short tones (e.g. Cantonese, Thai). The problem with the short tone analysis in Standard Mandarin is that some syllables appear to be

¹ The notation /H, LH, L, HL, ∅/ are used to refer to the underlying tones following Duanmu's convention. (2007). Tone 1-4 and the neutral tone are marked as *ā, á, ǎ, à*, and *a* (without a tone mark) respectively in the standard Romanization system *Hanyu Pinyin*.

de-stressed at the post-lexical level. Aside from the lexically marked neutral-tone syllables, which comprise about 15-20% of the syllables in written texts (W. Li, 1981, p. 35), SM speakers tend to de-stress the second syllable of disyllabic or trisyllabic words in colloquial speech (Chao, 1932; Shen, 1990, pp. 38-39). Although unstressed syllables can also be stressed if the word is infrequent (Chao, 1932; Jing, 2002), overall, about one-third of all SM syllables are unstressed and toneless in SM in connected speech (Duanmu, 2000, pp. 257-258). Therefore in SM, it is better to analyze these syllables as unstressed, and treat the neutral tone in SM as a phonetic representation of an unstressed syllable.

1.2 Taiwan Mandarin

However, not all Mandarin dialects exhibit a similar pattern. Taiwan Mandarin is a dialect spoken in Taiwan that has been influenced by Southern Min. Mandarin was brought into Taiwan as a standard language when the Nationalist government lost the civil war and retreated to Taiwan in 1949. While the Mainland immigrants spoke some variety of Mandarin, 70% of the population in Taiwan spoke Southern Min as a first language and were forced to speak Mandarin as a second language (Sandel, 2003). After decades of language shift, Mandarin has become the dominant language and has developed into a fully-fledged dialect (Her, 2010). One of the features of the Taiwan Mandarin dialect is that the differences between Taiwan Mandarin (TM) stressed and unstressed syllables are not very distinct perceptually—TM is often described as being more impressionistically syllable-timed compared to SM (Kubler, 1985).

Since the syllable canons are similar across Mandarin dialects, the more syllable-timed rhythm in TM lies in the contrast between stressed and unstressed syllables. Like SM, TM also lexically marks the four lexical tones and the neutral tone. However, lexically marked neutral-tone syllables and syllable de-stressing occur less frequently in TM than in SM (Duanmu, 2007, p. 308; Kubler, 1985, p. 161; Swihart, 2003, p. 110; Tsao, 2000). Also, Duanmu (2000, pp. 266-267) noticed that, unlike in SM, vowel devoicing and deletion are not found in unstressed syllables in TM. Consonant reduction is found in TM, but it can happen in stressed syllables too. Even consonants in word initial position can reduce in TM, for example /k^hə-ʂi/: [k^hɣ-sw] ‘but’ becomes [ɣ-sw], and /zən-xəu/: [zən-xəu] ‘afterwards’ becomes [ã-əu]². This unconditioned consonant reduction may be a manifestation of the rhythmic differences between SM and TM because the reduction is not restricted to unstressed syllables—a more syllable-timed language tends to have a similar treatment on both stressed and unstressed syllables (Dauer, 1983).

Fewer neutral-tone syllables can be demonstrated in the prescriptive grammar in TM. I compared the words in the List of Neutral-Tone Words for the Standard Mandarin

² Many TM speakers often de-retroflex /ʂ/ and /z/.

Proficiency Test³ with the online Revised Mandarin Chinese Dictionary published by the Ministry of Education in Taiwan. I found that all the suffixes and the reduplicants remain prescriptively marked as having a neutral tone in TM. However, only about half of the compound words (138/272) were marked as having full-neutral disyllabic patterns in TM⁴, while the other half (134/272) were marked with lexical tones. For example, *xiàba* [ɕja^Hpa⁰] ‘chin’ in SM remains marked as /HL-Ø/ in TM, but *yuèliang* [yɛ^Hlijaŋ⁰] ‘moon’ in SM is marked as /HL-HL/ in TM. I further examined these prescriptively neutral-tone syllables in TM by trying to input the syllables/characters in the computer’s operating system⁵ assuming that the input system is designed to let TM users input these syllables in the most intuitive way. I found that 116 out of 138 (84%) of the TM prescriptive neutral-tone syllables in compound words cannot be found under the neutral-tone entry. This piece of evidence shows that only 8.1% ((138-116)/272) of the SM full-neutral compound words are treated as full-neutral in TM. The percentage of full-neutral disyllabic words might be even lower. For example, although a neutral tone needs to be typed in *luó-bo* [lwo^{LH}pwo⁰] ‘radishes’ and *yào-shi* [jao^{HL}ʂi⁰] ‘keys’, they seem to be pronounced *luó-bō* [lwo^{LH}pwo^H] and *yào-shǐ* [jao^{HL}ʂi^H] (or *yào-shí* [jao^{HL}ʂi^{LH}]) in TM respectively.

Aside from the compound words, the prescriptive neutral-tone markings on grammatical suffixes, final particles, and reduplicants also do not seem to fully reflect TM speakers’ phonological representation of these neutral-tone syllables. It seems that only some of the prescriptively marked neutral-tone syllables are recognized by TM speakers. For example, the experiential aspect suffix *-guo* [kwo⁰] and many directional complements such as *-shang* [ʂaŋ⁰] ‘up’ and *-lai* [lai⁰] ‘come’ are not reduced; instead they are pronounced with lexical tones *guò*, *shàng*, *lái* [kwo^{HL}; ʂaŋ^{HL}; lai^{LH}]. As for reduplications, only kinship terms are produced with a neutral tone. Verbal reduplication and nominal reduplication are produced with an identical tone. For example, *kànkān* [kan^{HL}-kǎ⁰] ‘to take a look’ is pronounced as [kan^{HL}-kan^{HL}]; *xīngxīng* [[ɕiŋ^H-ɕi⁰] ‘star’ is pronounced as [ɕiŋ^H-ɕiŋ^H]. When the reduplication is used to show endearment, a [L-LH] or [L-H] pitch pattern is used in TM (Hsu, 2006, pp. 119-120). For example, *bàba* [pa^{HL}-pa⁰] ‘father’ and *māma* [ma^H-ma⁰] ‘mother’ can be pronounced as *bǎbá* [pa^L-pa^{LH}] and *mǎmá* [ma^L-ma^{LH}] respectively, while *jǐjǐe* [tɕjɛ^L-tɕjɛ⁰] ‘older sister’ can be pronounced as *jǐjǐē* [tɕjɛ^L-tɕjɛ^H]. This phenomenon can also be shown in

³ *Putonghua shuiping ceshi yong qingsheng cibiao* 普通话水平测试用轻声词表 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 People’s Republic of China government.

⁴ Mandarin has trochaic disyllabic feet. Therefore unstressed syllables occur at the second syllables of a disyllabic compound words. Full-neutral here indicates that the first syllable has a full tone (one of the four lexical tone), and the second syllable has a neutral tone.

⁵ One of the most common character input methods in Taiwan is the New Phonetic Input (新注音符輸入法) in 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.

colloquial writing online. For example, sometimes 媽媽 *māma* [ma^H-ma⁰] ‘mother’ is written as 馬麻 (*mǎmá* [ma^L-ma^{LH}]), 弟弟 *dìdì* [ti^{HL}-ti⁰] ‘younger brother’ is written as 底迪 (*dǐdǐ* [ti^L-ti^{LH}]), and 寶寶 *bǎobao* [pao^L-pa⁰] ‘baby’ is written as 寶包 (*bǎobāo* [pao^L-pao^H]). It is worth noting that the [L-H] pattern is only adopted when the reduplicated syllable is /L/ (Hsu, 2006). It is very likely that the TM speakers analyze the post-L raising pitch of the reduplicants as /H/. As for final particles such as *ba* [pa⁰], *ne* [nə⁰], *la* [la⁰], *ya* [ja⁰], and *a* [a⁰], they are also marked with a neutral tone in TM prescriptively. However, their pitch values also vary depending on the intonation of the sentence (Chao, 1968; Shen, 1990).

For the remaining neutral-tone syllables in TM, some of them clearly do not possess an underlying lexical tone. For example, the second syllable of *dìdì* [ti^{HL}-ti⁰] ‘younger brother’ seems to have a neutral tone because the second *dì* [ti⁰] is not pronounced with [HL]. Many reduplicated kinship terms also show a similar pattern in which the second syllable clearly does not possess a lexical tone of the character. Some of them, mostly grammatical suffixes and final particles, do not have a lexical tone that are transparent to the present-day speakers because they were reduced and grammaticalized diachronically. For example, durative *zhe* [ʈʂə⁰] was grammaticalized from *zháo* [ʈʂao^{LH}] ‘contact’, but this origin is not obvious to the speakers.

Only a few studies have investigated the acoustic properties of these remaining neutral-tone syllables. Two preliminary studies suggest that the pitch of the TM neutral tone is not influenced by the preceding lexical tones (J. Li, 2005; Tseng, 2004). The pitch pattern of the TM neutral tone was described as having a “certain pitch target” (J. Li, 2005) or pronounced with a low entering tone (short tone) (Tseng, 2004). However, both preliminary studies were based on limited sets of data without controlling for the environment or systematically comparing them to the other lexical tones. It remains unclear how the pitch patterns of the neutral tone should be characterized in relation to the other lexical tones, and whether these syllables are in fact unstressed. If the TM neutral tone has some kind of tonal target, as both previous studies suggested, does that imply the TM neutral-tone syllable has its own tonal representation? If the TM neutral tone has an underlying pitch target that is different from the existing four lexical tones and it is not reduced, we might need to posit a new lexical tone in Taiwan Mandarin. On the other hand, it could simply be the case that the unstressed syllables in TM still reduce and undergo tone neutralization, but its pitch patterns were neutralized differently from SM. To the best of my knowledge, there has been no complete study that examines the acoustic features of the neutral tone in TM and further analyzes the phonological identity of the TM neutral tone.

This acoustic study aims to answer the following research questions: 1) What is the pitch target of the TM neutral tone? Specifically, how is its pitch contour different from the SM neutral tone, and how is the pitch contour different from the other TM lexical tones? 2) Is the TM neutral tone unstressed like in SM? Are the syllables reduced in

duration or intensity? By answering these questions, I hope to characterize the TM neutral tone syllables and provide a clearer picture of the phonological identity of these “unstressed” syllables in TM. As Taiwan Mandarin is perceived as more syllable-timed than SM, the treatment of the so-called “unstressed” syllables in TM is a perfect chance to help us understand how rhythmic differences are manifested in a tone language. It would also further our understanding of lexical stress in complex tone languages.

In this study, I will examine the neutral-tone syllables in TM. I will focus on the grammatical morphemes and the reduplicated kinship terms because unlike particles and compound words, these items are both frequent and relatively stable in pitch. Two production tests were conducted to characterize the neutral tone in TM: the first test in Section 2 investigated the consecutive neutral tones in TM in comparison with the previous SM results in order to uncover the pitch target of the neutral tone, while the second test in Section 3 compared the neutral tone with the similar Tone 3 /L/ to characterize the pitch target of the neutral tone and to observe whether the neutral-tone syllables are reduced.

2 The pitch target of the neutral-tone syllables

In Standard Mandarin, consecutive neutral tones are influenced by the preceding lexical tones and slowly approach to a mid target—unlike a lexical tone, where the influence of the preceding lexical tone is overcome quickly to implement its tonal target, the influence of the preceding lexical tones is still substantial at the end of three consecutive neutral tones (Chen & Xu, 2006). I hypothesize that unlike SM, consecutive TM neutral tones do not exhibit similar tonal behavior. I expect that when producing consecutive neutral tones, the TM speakers will approximate a pitch target in each consecutive neutral tone. Therefore the carry-over effects of the preceding lexical tones are expected to be overcome quicker than the SM neutral tones, and neutral-tone contours with different preceding tones will share similar contour shapes and pitch registers in the second and the third neutral tone.

2.1 Design

Several prescriptive neutral-tone syllables were tested. They include four reduplicants (RED)⁶ *ma*, *po*, *nai*, and *mei* [ma[∅], p^hwo[∅], nai[∅], me[∅]], the nominalizer (NMLZ)/ possessive (POSS) *de* [tə[∅]], the plural (PL) *men* [mən[∅]], the durative (DUR) *zhe* [tʃə[∅]], and the

⁶ As discussed in 1.2, reduplication in TM seldom has a neutral tone and it tends to adopt different pitch contour to show endearment. Here we have chosen kinship terms that are more likely to show a neutral tone for elicitation. However, a reduplicated /H/ in kinship terms is usually /H/ and a reduplicated /LH/ in kinship terms is usually /HL/. Although this design is not ideal, reduplication needed to be included in this experiment because it was the only way to elicit three consecutive neutral tones.

perfective (PFV) *le* [lǎ⁰]. These tested neutral-tone syllables were combined into three sets of meaningful consecutive neutral tone syllables: 1) *-X-men-de* [˜mən⁰tǎ⁰] ‘-RED-PL-POSS’, 2) *-zhe-de* [tʃǎ⁰tǎ⁰] ‘-DUR-NMLZ’, and 3) *-le-de* [lǎ⁰tǎ⁰] ‘-PFV-NMLZ’. The first set of sentences had three consecutive neutral tones and the other two sets had two. Each set was put into a meaningful sentence in order to elicit a more natural speech. An example of a sentence is given below. The full list of stimuli sentences is provided in the Appendix A.

		X			Y		
		L	N	N	HL		
<i>tā</i>	<i>xǐhuān</i>	<i>pǎo</i>	<i>-le</i>	<i>-de</i>	<i>nà</i>	<i>-ge</i>	<i>rén</i>
[t ^h ɑ ^H]	ɕi ^H hwan ^H	p ^h ɑo ^L	lǎ ⁰	tǎ ⁰	nɑ ^{HL}	kǎ ⁰	ʝən ^{LH}]
3s	like	run	-PFV	-NMLZ	that	CLF	person
‘He likes the person that ran away’							

Adopting Chen and Xu’s (2006) methodology, in order to see the influence of the preceding lexical tones, the syllable before the consecutive neutral tones (X) varied in four lexical tones. N represents the neutral-tone syllables (two or three), and the syllable following the consecutive neutral tones (Y) also varied in order to control for contextual influence. Only /L/ and /HL/ were included because /L/ starts at the lowest pitch and /HL/ starts at the highest pitch. As a result, the subjects read 4 (lexical tone of X)* 3 (sets of consecutive neutral-tone syllables)* 2 (Y) = 24 sentences.

Six male and six female TM speakers who were between 21 and 33 years old without any speaking or hearing impairment were recruited. The subjects all grew up in Taiwan and their parents all speak Southern Min as their native language. All twelve subjects speak Southern Min, Mandarin, and English. The subjects had not resided in any foreign country for more than three months in the preceding year, a fact that reduces the possibility of dialectal influences from other Mandarin varieties.

The subjects were asked to read these 24 sentences written in Chinese characters. These neutral tone combinations are semantically natural and relatively frequent, and the subjects were expected to produce them without any difficulties. There was no practice session but the subjects were encouraged to re-read a sentence if they felt they had made a mistake.

2.2 Data

The subjects were recorded with a Zoom H2 recorder in quiet offices. The speech was recorded with the sample rate of 44100 Hz and was analyzed in Praat. Only the rimes of the syllables were segmented in order 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 rimes were hand-labeled. With the aid of a Praat script, the f_0 of the rimes were measured at the midpoint of all the 1/10 intervals (5%, 15%...95% point of the rime). Measurements were then examined to adjust for pitch-halving and pitch doubling. If a tested syllable had more than 5 consecutive pitch data missing (mostly due to creaky voice), the whole sentence was deleted.

The f_0 values were first converted into semitones, and then normalized to z-score by speakers in order to reflect the relative pitch height within each speaker's pitch range—the pitch data distribute between $-2 \sim 2$ semitone z-score ($-2, -1, 0, 1, 2$), which can be interpreted as low, mid-low, mid, mid-high, and high pitch range. The pitch data presented in charts were normalized in semitone z-score. However, pitch data in semitone were used in statistical analyses.

2.3 Analysis

The data were further analyzed with linear mixed models in SPSS Statistics 23. A mixed models approach was chosen because it retained the remaining data without listwise deletion when there were missing data in repeated measures. The models first included all the tested fixed effects, random intercepts, and random slopes. The tested fixed effects were kept intact regardless of their statistical significance, but other parameters were then reduced to reach the best-fit model. Model selection was done based on the comparison of the Akaike Information criterion (AIC) score. Insignificant interactions were first dropped from the models to see if it yielded a lower AIC score; insignificant main effects were then dropped unless they contributed to a significant interaction.

In this experiment, in order to analyze the pitch contours of the consecutive neutral tones, the dependent variables were the repeated measures of pitch readings. My main focus is to see the effects of the preceding tones on the consecutive neutral tones. Therefore the fixed effects included the preceding tone X (H, LH, L, HL) as a factor, position (10 pitch points labelled 0-9 representing the 5%, 15%, 25%,..., 95% point) as a covariate, and the X–position interaction which captured the pitch changes along the repeated measured pitch points (i.e. pitch contour) by various preceding tones. The random effects included both subjects (12 speakers) and items (tested sentences) as random intercepts in order to control for the variations coming from the speakers and the tested item. By-subject random slopes on position, preceding tone X, following tone Y (HL, L), the neutral tone category (*-zhe-de*, *-le-de*) in the case of two consecutive tones, and their interactions were included; and by-item random slopes on position was also included.

Based on my hypothesis, I expect to see that at the second and the third neutral tone, the influence of the preceding tone is overcome, which means that the fixed effect of the preceding tone will not be significant; I also expect that the second and the third neutral tones will share a similar pitch contour regardless of their preceding lexical tones. Therefore the fixed effect of the X-position interaction is expected to be insignificant.

2.4 Results

2.4.1 Two consecutive neutral tones

The average pitch contours of the two consecutive neutral tones are plotted in Figure 1. Only the rimes were measured and graphed. The figure shows the eight average normalized pitch contours of the 12 TM speakers producing two sets of two consecutive neutral tones (*-zhe-de* and *-le-de*). The eight pitch contours varied the four preceding tones (syllable X) and the two following tones (syllable Y).

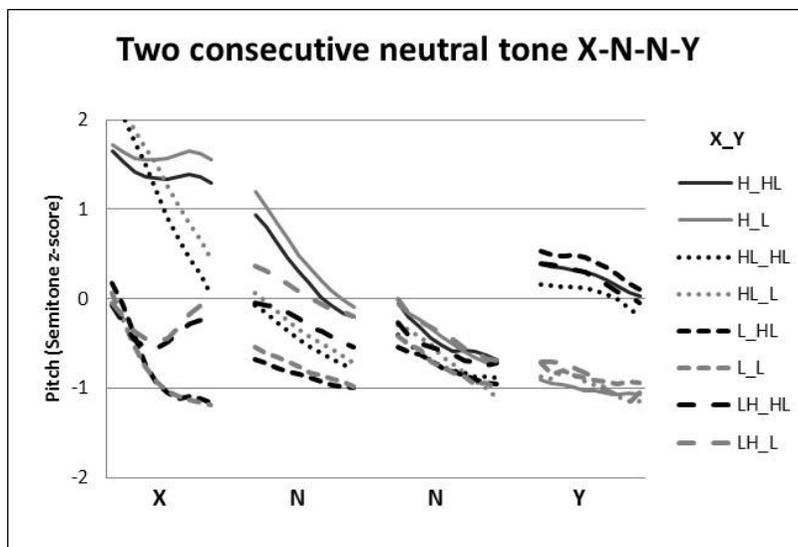


Figure 1: Results of the two consecutive neutral tones

As shown in Figure 1, the first syllable X illustrates the four different lexical tones. Then the different degrees of the falling pitch contours move from the end of the first syllable X toward the mid to mid-low range in the first neutral tone except for the one after /L/. The eight pitch contours of the second neutral tone syllables are more homogeneous compared to the first neutral tone. They generally fell from mid to mid-low pitch range.

Two sets of linear mixed model confirm the visual observation stated above. For the first neutral tone, the linear mixed model showed that there were significant effects of preceding tone [$F(3, 22.266)=23.794, p<.001$], position [$F(1, 37.524)=126.291, p<.001$], and the tone–position interaction [$F(3, 12.187)=19.923, p<.001$]. The estimates of the fixed effects, shown in Table 1, illustrate that at 5% of the first rime (baseline), the neutral tone after /H/ is 1.719 semitone higher than that after /LH/ (baseline) ($p=.001$), and the neutral tone after /L/ is 1.856 semitone lower than that after /LH/ ($p<.001$). The significant position effect confirms that the contours were not flat statistically. The first neutral tone after /LH/ (baseline) reduced 0.142 semitone per unit (10% of the rime), and the pitch contours after different lexical tones differed: the first neutral tone after /H/ reduced a further 0.148 semitone ($p<.001$), making its estimated pitch contour fall 0.290 semitone per unit; the pitch contour after /L/ fell less than that after /LH/ ($p=.039$), only 0.075 ($=-0.142+0.067$) semitone per unit. Overall, if the pitch at 5% of the rimes was higher (such as after /H/), the pitch contour fall was sharper, which suggests the convergence of the pitch contours.

Table 1: Estimates of fixed effects on the first neutral tone in two consecutive neutral tones

Parameter	Estimate (semitone)	SE	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	19.691	1.632	11.710	12.063	0.000	16.125	23.257
[X=H]	1.719	0.433	21.757	3.971	0.001	0.821	2.618
[X=HL]	-0.650	0.440	22.851	-1.476	0.154	-1.562	0.261
[X=L]	-1.856	0.433	21.723	-4.288	0.000	-2.754	-0.958
[X=LH]	0 ^b	0.000
position	-0.142	0.023	26.495	-6.146	0.000	-0.190	-0.095
[X=H] * position	-0.148	0.029	11.605	-5.178	0.000	-0.211	-0.086
[X=HL] * position	-0.029	0.030	13.147	-0.979	0.345	-0.093	0.035
[X=L] * position	0.067	0.029	11.457	2.336	0.039	0.004	0.129
[X=LH] * position	0 ^b	0.000

At the second neutral tone, the linear mixed model showed that while the effects of the preceding lexical tone were still significant on the pitch ($F(3, 38.656)=4.465, p=.009$) and position [$F(1, 30.219)=129.241, p<.000$], their interaction was not significant ($F(3, 11.677)=0.5, p=.689$). The estimates of fixed effects shown in Table 2 reveal that at 5% of the second neutral tone rime, the pitch estimate after /HL/ was 0.566 semitone lower than the pitch estimate after /LH/ ($p=.033$). The pitch fall of the second neutral tone after /LH/ was estimated to be -0.126 semitone per unit ($p<.001$), but its slope was not significantly different from those after other lexical tones. The estimates show that while the preceding tone effect was significant, the differences between different preceding tones were within 1 semitone; furthermore, these pitch contours after different lexical tones share a similar degree of pitch fall, and the

estimated degree of pitch fall was small (around -0.13 semitone per unit), suggesting a shallower pitch fall.

Table 2: Estimates of fixed effects on the second neutral tone in two consecutive neutral tones

Parameter	Estimate	SE	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	18.674	1.585	11.226	11.779	0.000	15.193	22.154
[X=H]	0.190	0.248	37.495	0.768	0.447	-0.311	0.691
[X=HL]	-0.566	0.255	39.940	-2.215	0.033	-1.082	-0.049
[X=L]	-0.517	0.247	37.447	-2.088	0.044	-1.018	-0.015
[X=LH]	0 ^b	0.000
position	-0.126	0.016	27.828	-8.006	0.000	-0.158	-0.094
[X=H] * position	-0.018	0.018	10.529	-1.044	0.320	-0.058	0.021
[X=HL] * position	-0.004	0.019	13.300	-0.220	0.829	-0.045	0.037
[X=L] * position	0.000	0.018	10.502	0.024	0.981	-0.039	0.040
[X=LH] * position	0 ^b	0.000

2.4.2 Three consecutive neutral tones

All the pitch contours of the three consecutive neutral tones were plotted and examined. Ten tokens which were obviously different from others were excluded, which will be discussed separately. The average pitch contours of the three consecutive neutral tones varying the preceding and the following tones are shown in Figure 2. The results show that the reduplicants, the first neutral tone among the three, behaved differently from other neutral-tone syllables. The reduplicants seem to have lexical tones /H/ after /H/: *māma* /H-Ø/ ‘mother’ is produced as /H-H/, suggesting that the second syllable was not produced with a neutral tone, but its lexical tone of the character. Also, the reduplicant in the word *pópo* /LH-Ø/ ‘mother-in-law’ had a higher falling pitch contour compared to the first neutral tone after /LH/ in Figure 1, and the pitch contour is similar to Tone 4 (high falling)—the high falling Tone 4 stays in the high register (0 ~ 2 semitone z-score) (Huang, 2013). Therefore, despite the reduplicants after /HL/ and /L/ falling to the mid to mid-low range as observed in two consecutive neutral tones, the first neutral tone in this set was excluded from the analysis. As for the second neutral tone, the pitch contours all fell and merged to the same pitch range (-0.5 ~ -1 semitone z-score). The pitch contours of the third neutral tone overlapped with each other and generally fell from -0.5 semitone z-score to -1 semitone z-score.

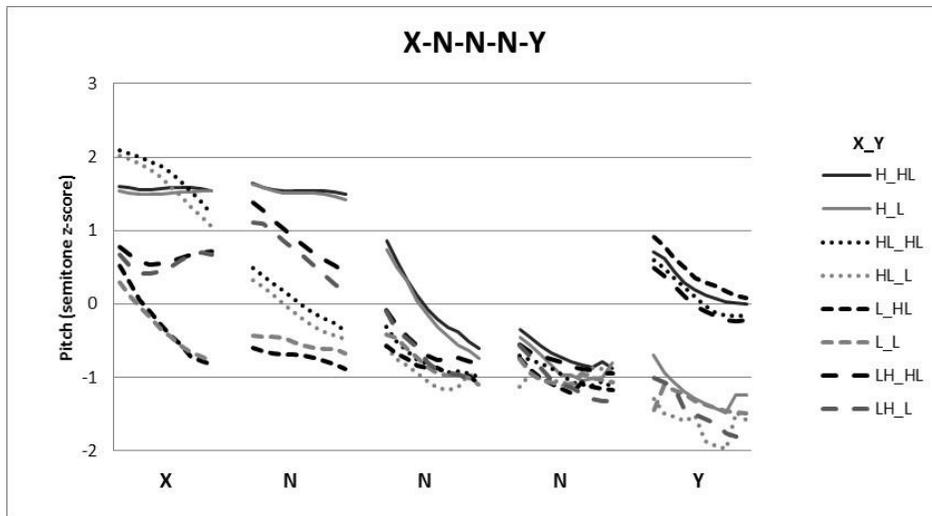


Figure 2: Results of three consecutive neutral tones: X-red-pl-poss-Y

The linear mixed model analysis shows that at 5% of the second neutral-tone rime, there were significant fixed effects of the preceding lexical tone X [$F(3, 5.694)=12.230$, $p=.007$], position [$F(1, 8.369)=67.130$, $p<.001$], as well as their interaction [$F(3, 4.244)=15.882$, $p=.009$]. The estimates of the fixed effects shown in Table 3 reveal that at 5% of the second neutral tone rime, the pitch after /H/ was 1.046 semitone higher than the pitch after /LH/ ($p=.041$), and the pitch after /L/ was 1.169 semitone lower than the pitch after /LH/ ($p=.028$). As for the pitch contour, the pitch contour after /H/ has a much sharper pitch fall (-0.285 semitone/unit) compared to the contour after /LH/ (-0.176 semitone/unit), and the difference (-0.110 semitone/unit) was statistically significant ($p=.015$). The pitch contours with varying preceding tones seem to converge as found in the first neutral tone in two consecutive neutral tones.

Table 3: Estimates of fixed effects on the second neutral tone in three consecutive neutral tones

Parameter	Estimate	SE	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	18.430	1.807	10.458	10.201	0.000	14.429	22.432
[X=H]	1.046	0.397	5.606	2.636	0.041	0.058	2.033
[X=HL]	-1.169	0.400	5.776	-2.924	0.028	-2.157	-0.182
[X=L]	-0.726	0.405	6.022	-1.795	0.123	-1.715	0.263
[X=LH]	0 ^b	0.000
position	-0.176	0.028	10.339	-6.381	0.000	-0.237	-0.115
[X=H] * position	-0.110	0.027	4.082	-4.043	0.015	-0.184	-0.035
[X=HL] * position	0.060	0.028	4.437	2.159	0.090	-0.014	0.135
[X=L] * position	0.037	0.028	4.704	1.319	0.248	-0.037	0.112
[X=LH] * position	0 ^b	0.000

Toward the third neutral tone, the effects of the preceding tone lexical tone on pitch range and pitch contour slopes had disappeared. The linear mixed model on the third neutral tone revealed that there were no significant effects from preceding tone [$F(3, 2.430)=4.509, p=.154$] or interaction [$F(3, 4.075)=2.058, p=.246$]. The position effect was also statistically insignificant [$F(1, 8.645)=2.997, p=.119$], suggesting the pitch contours were flattened. The estimates of fixed effects are shown in Table 4. In summary, the preceding lexical tone (X) affected the second neutral-tone syllables, but not the third neutral-tone syllables.

Table 4: Estimates of fixed effects on the third neutral tone in three consecutive neutral tones

Parameter	Estimate	SE	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	17.599	1.768	10.408	9.952	0.000	13.679	21.518
[X=H]	-0.121	0.374	2.497	-0.323	0.772	-1.460	1.218
[X=HL]	-1.221	0.379	2.607	-3.224	0.059	-2.537	0.094
[X=L]	-0.642	0.388	2.838	-1.654	0.202	-1.917	0.634
[X=LH]	0 ^b	0.000
position	-0.166	0.079	10.940	-2.100	0.060	-0.340	0.008
[X=H] * position	-0.003	0.052	3.942	-0.055	0.958	-0.148	0.142
[X=HL] * position	0.109	0.053	4.114	2.064	0.106	-0.036	0.253
[X=L] * position	0.057	0.053	4.205	1.083	0.337	-0.087	0.202
[X=LH] * position	0 ^b	0.000

2.5 Discussion

This experiment modifies Chen and Xu's study (2006) on the SM neutral tone and examines how consecutive neutral tones differ in TM. The first difference between the two dialects lies in the tonal variation of the reduplicants. Our results showed that not all of the reduplicants behave like other neutral tones—the reduplicant in *māma* /H-∅/ 'mother' was produced with a lexical tone /H/. Also, the reduplicant in *pópo* /LH-∅/ 'mother-in-law' had a higher falling pitch. This high falling pitch contour of the reduplicant after /LH/ is similar to the SM neutral tone after /LH/. However, because this pitch contour is only found in /LH-∅/ reduplicated kinship terms and it is not observed in other /LH-∅/ words, this is likely to be a sporadic variation. The TM speakers likely lexicalized the SM pitch contour of /LH-∅/ as /LH-HL/ in those kinship terms. If the TM speakers had adopted the SM neutral tone of high falling contour after /LH/, we would expect it to show up in other neutral tone syllables after /LH/ as well. Therefore this production test was unable to elicit three consecutive neutral tones in the context of four different lexical tones.

In addition, several tokens were excluded due to the tonal variations of the reduplicants. Six out of the ten excluded pitch contours also show possible SM influences on these reduplicated kinship terms. First, one male and one female subject

produced the reduplicated /H/ *ma* with a falling pitch contour instead of a high level pitch when X was /H/ and Y was /L/. This pitch contour of high-falling was similar to the pitch pattern of /H-Ø/ in SM. However, both of the speakers only had this pitch contour in X-RED-PL-POSS before /L/, but not before /H/, which suggests that the pitch pattern was sporadic, not consistent. In addition, six tokens were removed because two other male subjects produced the word *nǎi-nai* /L-Ø/ ‘grandmother’ with a [L-H] pitch contour, and a female subject consistently produced the words *nǎi-nai* /L-Ø/ ‘grandmother’ and *pó-po* /LH-Ø/ ‘mother-in-law’ with a [L-LH] pitch contour. As discussed earlier in 1.2, it is common in colloquial TM to adopt [L-H] or [L-LH] pitch contours on reduplicated kinship terms or nicknames when the reduplicated syllables are /L/ or /LH/, in order to show endearment (Hsu, 2006). The tonal variation in the reduplicated kinship terms in the data was likely due to dialect mixtures. The speakers likely borrowed a certain lexicalized pitch pattern rather than borrowing the neutral tone from a different dialect.

The main difference between TM and SM neutral tones, which is the main focus of this experiment, lies in the pitch contour of these consecutive neutral tones. Figure 3 compares the results of three consecutive neutral tones before /L/ produced by the TM subjects in (a) with the ones produced by two SM speakers in (b). Two SM speakers were recruited to complete the same task as the TM speakers. The results from the two SM speakers show similar pitch patterns as reported by Chen and Xu (2006). Therefore it is presented for comparison, despite the small sample size.

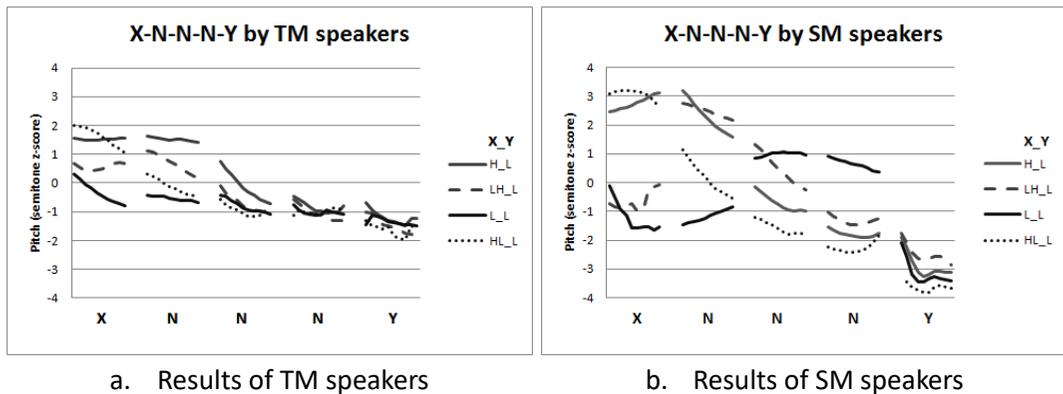


Figure 3: Comparison of the results of X-red-pl-poss-Y by TM and SM speakers

Comparing Figure 3(a) and (b), the pitch range of TM is smaller than that of SM, as reported by previous acoustic studies (Fon & Chiang, 1999; Torgerson, 2005). The pitches of TM range between 2 ~ -2 semitone z-score, while the pitches of SM range between 3 ~ -3.5 semitone z-score. More importantly, the pitch contours of TM and SM differ drastically. In TM, aside from the reduplicated /H/ and the reduplicated /LH/, all neutral tones have slight falling pitch contours reaching to -0.5 ~ -1 semitone z-score.

In contrast, as found in Chen and Xu (2006), all four SM pitch contours are distinct. The SM neutral tones after /H/ gradually move from high to the mid-low pitch range; the neutral tones after /LH/ complete the rising at the first neutral tone, and then lower to the mid-low range in the following neutral tones; the neutral tones after /L/ gradually rise to the high register until the second neutral tone, and then slowly lower to mid pitch range; the neutral tones after /HL/ have a fast pitch drop in the first neutral tone, and remain at mid-low pitches in the following neutral tones. The four pitch contours only start to converge to mid-level at the end of the third neutral tone in SM.

The influences of the preceding lexical tone differed in SM and TM. The pitch range of the neutral tones in SM remains 3 ~ 3.5 semitone z-score even in the third neutral tone, while the pitch range of the TM neutral tones reduced over time. In SM, the neutral tones are heavily influenced by the preceding lexical tones, and the influence of the preceding lexical tone is significant even at the end of the third neutral tone syllables (Chen & Xu, 2006, p. 61). Furthermore, aside from the obvious different pitch patterns of the post-L neutral tones contour, the influences of the preceding /H, LH, L/ were still obvious at the third neutral tone in SM. On the other hand, the four pitch contours in TM immediately started to converge in the first neutral tones. The results of the first neutral tone in 2.4.1 and the second neutral tones in 0 show that if the pitch was higher due to the influence of the preceding lexical tone, the pitch fall was also estimated to be larger. The pitch contour started to converge and reach toward a mid to mid-low pitch. Although the effect preceding tone was still significant at the second neutral tone in two consecutive neutral tones, the estimated effect was small, and the pitch contours in the context of varying preceding tones share a similar slope. The third neutral tone in three consecutive neutral tones further shows that the influences of preceding tone on pitch were no longer significant, and these pitch contours share a similar flattened slope as well. Overall, compared to the SM neutral tone, the pitch target of the TM neutral tone is implemented in a faster manner. In SM, the mid pitch target claimed by Chen and Xu was unable to be implemented at the end of the third neutral tone. However, the results show that the TM pitch target was approximated immediately at the first neutral tone, and the preceding lexical tone effects had disappeared at the third neutral tone.

Unlike SM, the TM neutral tone seems to have a static pitch target in the mid-low to low pitch range. This experiment shows that in TM, each consecutive neutral tone aims to reach to a mid-low pitch range. The pitch contours started to flatten in the following neutral tones, which indicates that the pitch contour had been approximating to the pitch target. Although the estimated beginning pitches of these neutral tone rimes were slightly higher than the end point of the preceding neutral tone, the differences were minimum (less than 1 semitone). These differences are very likely due to the pitch raising effect of their preceding voiceless onset /t/ (Hombert, 1977; Hombert, Ohala, & Ewan, 1979; Ohala, 1972), rather than targeting for a falling pitch contour.

It should be noted that, although the TM neutral tone has a static pitch target that is approximated in a faster manner compared to the SM neutral tone, the articulation strength to approach this target was still weaker than a pitch target of a lexical tone. The previous studies on carry-over effects of preceding lexical tones showed that the influence of the end pitch of the preceding lexical tone lasted until the end of following /H/ and /LH/ in SM, but not until the end of following /HL/ and /L/ (Y. Xu, 1997); and the carry-over effects of the preceding tones are significant only at the end of the /LH/ tone in TM (Huang, 2013).

To sum up, the results show that the consecutive TM neutral tone has a mid-low or low pitch target, and the target was approached faster than the SM neutral tone, which suggests that it is more similar to a lexical tone. However, the results from this experiment alone are unable to characterize the TM neutral tone. It is unclear how these neutral-tone syllables fit into TM phonology, and if so, how they should be characterized. The next experiment compares the TM neutral-tone syllables with low-tone syllables, with the aim to investigate these questions.

3 Neutral-tone syllables vs. low-tone syllables

This experiment aims to find out whether the TM neutral tone should be treated as a lexical tone or a neutralized tonal pattern. Although the previous experiment showed that the TM neutral tones are more lexical-tone-like compared to SM, it is also possible to analyze these TM neutral-tone syllables as unstressed, and their pitches were neutralized into a mid-low tone. Furthermore, if these neutral-tone syllables had a lexical tone, how would it be distinct from the other lexical tones? In order to investigate 1) whether these neutral-tone syllables are unstressed, and 2) whether these neutral-tone syllables are distinct from the other lexical tones, this experiment compares the pitch, duration, and intensity of these TM neutral-tone syllables with the same acoustic properties of the low-tone (Tone 3) syllables. Out of the four lexical tones, Tone 3 was chosen because it is the shortest tone in connected speech in Taiwan Mandarin (Deng, Shi, & Lu, 2008; Shi & Deng, 2006). Moreover, the pitch contour of Tone 3 (low-falling) is the most similar to the TM neutral-tone syllables.

If these neutral-tone syllables were to have a weaker prominence, they would have a reduced duration (or even a reduced intensity) compared to the shortest lexical tone. In that case, it is best to analyse them as unstressed syllables with tone being neutralized into a mid-low pitch. However, if the neutral-tone syllables were not reduced in length or intensity, there is no acoustic evidence to support the analysis to treat these neutral-tone syllables as unstressed. They should be analysed as having a lexical tone. In that case, if these neutral-tone syllables had a distinct pitch target from the low tone, they should be analysed as having a fifth tone.

3.1 Design

The same subjects in the first experiment were asked to read a list of disyllabic words/phrases that contain the neutral-tone syllables and their corresponding low-tone syllables with the same segments or rimes. The tested (near) minimal pairs were both the second syllables of disyllabic words with varying preceding lexical tones.

Table 5: Tested corresponding control syllables with the same segment

Pair	words	Disyllabic words for elicitation			
		[H-Ø] [H-L]	[LH-Ø] [LH-L]	[L-Ø]	[HL-Ø] [HL-L]
-zi	<i>zi</i> [tʂi ^Ø]	<i>zhuō-zi</i> [tʂwo ^H -tʂi ^Ø] 'table'	<i>jiá-zi</i> [tɕja ^{LH} -tʂi ^Ø] 'clip'	<i>yǐ-zi</i> [i ^L -tʂi ^Ø] 'chair'	<i>niè-zi</i> [ɲje ^{HL} -tʂi ^Ø] 'tweezer'
	<i>zi</i> [tʂi ^L]	<i>jūn-zǐ</i> [tɕyn ^H -tʂi ^L] 'gentleman'	<i>jiǎ-zǐ</i> ⁷ [tɕja ^{LH} -tʂi ^L] '60 years'	-	<i>niè-zǐ</i> [ɲje ^{HL} -tʂi ^L] 'unfilial child'
-zhe	<i>zhe</i> [tʃə ^Ø]	<i>tīng-zhe</i> [t ^h iŋ ^H -tʃə ^Ø] 'listening'	<i>dú-zhe</i> [tu ^{LH} -tʃə ^Ø] 'reading'	<i>xiǎng-zhe</i> [ɕjaŋ ^L -tʃə ^Ø] 'thinking'	<i>zuò-zhe</i> [tʂwo ^{HL} -tʃə ^Ø] 'sitting'
	<i>zhe</i> [tʃɤ ^L]	<i>shāng-zhě</i> [ʃaŋ ^H -tʃɤ ^L] 'injured person'	<i>dú-zhě</i> [tu ^{LH} -tʃɤ ^L] 'reader'	-	<i>zuò-zhě</i> [tʂwo ^{HL} -tʃɤ ^L] 'author'
-en	<i>men</i> [mən ^Ø]	<i>tā-men</i> [t ^h a ^H -mən ^Ø] '3PL'	<i>rén-men</i> [zən ^{LH} -mən ^Ø] 'people'	<i>wǒ-men</i> [uo ^L -mən ^Ø] '1PL'	<i>jiào-shòu-men</i> [tɕjao ^{HL} -ʃo ^{HL} - mən ^Ø] 'professors'
	<i>ben</i> [pən ^L]	<i>jī-běn</i> [tɕi ^H -pən ^L] 'basic'	<i>chéng-běn</i> [tʃ ^h əŋ ^{LH} -pən ^L] 'cost'	-	<i>kè-běn</i> [k ^h ɤ ^{HL} -pən ^L] 'textbook'
-Cə	<i>le</i> [lə ^Ø]	<i>chī-le</i> [tʃi ^H -lə ^Ø] 'eaten'	<i>lái-le</i> [lai ^H -lə ^Ø] 'come'	<i>mǎi-le</i> [mai ^L -lə ^Ø] 'bought'	<i>qù-le</i> [tɕ ^h y ^{HL} -lə ^Ø] 'gone'
	<i>de</i> [tə ^Ø]	<i>suān-de</i> [swan ^H -tə ^Ø] 'sour ones'	<i>tián-de</i> [t ^h jan ^{LH} -tə ^Ø] 'sweet ones'	<i>kǔ-de</i> [k ^h u ^L -tə ^Ø] 'bitter ones'	<i>là-de</i> [la ^{HL} -tə ^Ø] 'spicy ones'
	<i>ge</i> [kə ^Ø]	<i>bā-ge</i> [pa ^H -kə ^Ø] 'eight'	<i>yí-ge</i> [i ^{LH} -kə ^Ø] 'one'	<i>wǔ-ge</i> [u ^L -kə ^Ø] 'five'	<i>sì-ge</i> [si ^{HL} -kə ^Ø] 'four'

⁷ *jiǎ-zǐ* /tɕja^L-tʂi^L/ was used to elicit the [LH-L] tonal pattern with *zi* in order to create a minimal pair with *jiá-zi* /tɕja^{LH}-tʂi^Ø/ because /L-L/ is phonetically the same as [LH-L] in TM (Myers & Tsay, 2003).

Pair	words	Disyllabic words for elicitation			
		[H-Ø] [H-L]	[LH-Ø] [LH-L]	[L-Ø]	[HL-Ø] [HL-L]
	<i>che</i> [tʃ ^h ɿ ^l]	<i>lā-chě</i>	<i>xián-chě</i>	<i>luàn-chě</i>	
	<i>she</i> [ʃɿ ^l]	[lɑ ^h -tʃ ^h ɿ ^l] 'pull and drag'	[ɕjan ^{LH} -tʃ ^h ɿ ^l] 'chitchat'	[lwan ^{HL} -tʃ ^h ɿ ^l] 'chitchat'	
		<i>shī-shě</i> [ʃi ^H -ʃɿ ^l] 'to give'	<i>nán-shě</i> [nan ^{LH} -ʃɿ ^l] 'hard to let go'	<i>bù-shě</i> [pu ^{HL} -ʃɿ ^l] 'can't let go'	
<i>RED</i>	<i>shu</i> [ʃu ^Ø]		<i>shú-shu</i> [ʃu ^{LH} -ʃu ^Ø]	<i>shěn-shen</i> [ʃən ^L -ʃən ^Ø]	
	<i>shen</i> [ʃən ^Ø]			<i>dì-dì</i> [ti ^{HL} -ti ^Ø]	
	<i>dì</i> [ti ^Ø]		'uncle'	'aunt'	
			'brother'		
	<i>shǔ</i> [ʃu ^l]	-	<i>fān-shǔ</i> [fan ^H -ʃu ^l]	<i>píng-shěn</i> [p ^h in ^{LH} -ʃən ^l]	
	<i>shěn</i> [ʃən ^l]			<i>dì-dǐ</i> [ti ^{HL} -ti ^l]	
	<i>dǐ</i> [ti ^l]		'sweet potato'	'judges'	
				'underground'	

Table 5 lists all the tested words used to elicit the syllables. The neutral-tone *-zì* [tsi^Ø] 'DIM' was compared with *-zǐ* [tsi^l], and *-zhe* [tʃə^Ø] 'DUR' has a corresponding low-tone syllable *-zhě* [tʃɿ^l] 'person'. The neutral-tone *-men* [mən^Ø] 'PL' was compared with *-běn* [pən^l] 'origin' because *měn* does not exist in Mandarin due to a phonological gap. For the same reason, three neutral-tone syllables with an /ə/ rime (*le* [lə^Ø] 'PFV', *de* [tə^Ø] 'POSS; NMLZ', and *ge* [kə^Ø] 'CLF') were compared with *shě* [ʃɿ^l] and *chě* [tʃ^hɿ^l].⁸ Phonetically the rimes of *chě* and *shě* are described as [ɿ], while the rimes of the neutral-tone *de*, *le*, *ge* are described as [ə] because they are unstressed and reduced. However, even in SM, the quality difference between a stressed /ə/: [ɿ] and an unstressed /ə/: [ə] is not obvious (M. Lin & Yan, 1990). As for the reduplicated syllables, they were compared to their corresponding syllables that have a similar pitch and the same segments respectively: *shu* [ʃu^Ø], *shen* [ʃən^Ø], *dì* [ti^Ø] in the kinship terms were compared to *shǔ* [ʃu^l], *shěn* [ʃən^l], and *dǐ* [ti^l]. The two kinship terms *shú-shu* [ʃu^{LH}-ʃu^Ø] and *shěn-shen* [ʃən^L-ʃən^Ø] are sometimes applied with a /L-LH/ or /L-H/ pitch contour to show endearment in colloquial TM. However, the adaptation of the pitch contours is generally less common in read speech. Therefore these words were still chosen to compare with the low tone counterparts. The reduplicated high tone syllables were not included because the reduplicated *mā* kept the high tone as mentioned in the previous experiment.

The tested neutral-tone syllables were elicited in four different disyllabic words with varying preceding tones, but the tested low-tone syllables were tested only in the combination of /H-L/, /LH-L/, and /HL-L/ due to Tone 3 Sandhi—/L/ becoming [LH] when it is after another /L/. Therefore /L-L/ was not elicited because it is phonetically the same as [LH-L] in TM (See Myers and Tsay (2003) for a review). All the disyllabic

⁸ *ge* /Ø/ does have a low-tone corresponding syllable *ge* /L/. However, this syllable is very limited in its occurrences in words. As a result, it was not included as a corresponding syllable.

words were carried in the frame sentence *qǐngshuō X X bācì* [tɕʰiŋ^L ʃwo^H X X pɑ^H tsɿ^{HL}] ‘Please say XX eight times’. All the sentences were written in Chinese characters. The 45 tested sentences were randomly mixed with the 73 filler sentences. The subjects were asked to read every sentence just once, but they could repeat a sentence if they felt they made a mistake. The whole procedure, including both experiments, took about 15 minutes.

3.2 Data analysis

The pitch data of this experiment was processed as described in 2.2. In addition, the duration and the average intensity were also extracted. In order to compare the pitch (contour) between /L/ and /∅/, two sets of linear mixed models were fitted to the data. The first set analyzed the pitch contour differences between the low tone and the neutral tone at the first half and the second half of the rime. For the first half of the rime, the dependent variable is the repeated measured pitch in semitone at 5%, 15%, 25%, 35%, 45%; and for the second half of the rime, it is at 55%, 65%, 75%, 85%, 95%. Fixed effects for the two models included the tested tone (L, ∅) as a factor, position (the 5 pitch points coded in 0, 1, 2, 3, 4) as a covariate, and their interaction. The factor tested tone can capture their pitch differences at the starting point, and the tone–position interaction can capture the differences of the contour movement. In order to account for the speaker and item variances, the random effects included both intercepts of the subjects and tested items, as well as by-subject random slopes for tested tone, preceding tone (H, LH, L, HL), tested pair (*zhe*, *zi*, Cə, *en*, RED) and their interactions.

The second set of linear mixed model further investigated the effects of preceding tones and testing pair between /L/ and /∅/ at the 75% of the second rime. The dependent variable was the pitch in semitone at 75% of the second rime. Fixed effects included preceding tone to examine the carry-over effects, tested pair to examine pair variations, and tested tone (L, ∅) to examine whether they are distinct. All their interactions were included as fixed effects as well. As for random effects, it included subjects and items as intercepts and their random slopes for all the fixed effects. If these TM neutral-tone syllables had a distinct lexical tone, I expect them to have a distinct pitch contour from the low tone regardless of its preceding tone and tested pair. Their pitch target should also be implemented in a similar manner compared to the low tone.

As for the duration, the rimes were also chosen to be measured instead of syllables because languages generally do not count the onset in their calculation of syllable weight; only the rimes are relevant to syllable weight. As for the intensity, the average intensity of the rimes were also extracted with a Praat script. Durational and intensity comparison were carried out with linear mixed models with pair segments (*-en*, *zi*, *-zhe*, Cə, and *-u*, *-en*, *-i* in reduplication pairs) and interaction between tone and pair segment as fixed effects. As for random effects, random intercepts for subjects and tested words

and their random slopes for all the fixed effects were included. If the neutral-tone syllables were unstressed, I expect the durations of the TM neutral-tone syllables to be significantly shorter than the low-tone syllables regardless of the pairs. I do not expect the TM neutral-tone syllables to have a lower intensity as reduced intensity was not found in SM unstressed syllables either (J. Cao, 1986; Lin & Yan, 1980).

3.3 Results

3.3.1 Pitch contours

The average normalized pitch contours of the low tone and the neutral tone with varying preceding lexical tones are plotted in Figure 4. The neutral tone had four pitch contours with varying preceding tones, represented by the solid lines, but the low tone syllable had only three average pitch contours, represented by the dotted lines, with preceding [H, LH, HL] because [L-L] does not exist in Mandarin due to Tone 3 Sandhi.

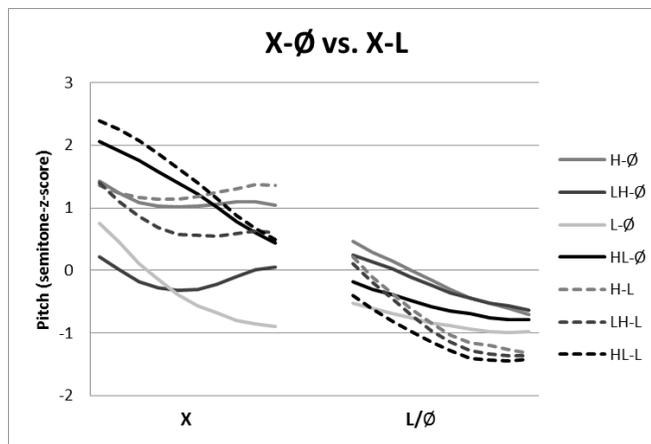


Figure 4 Comparisons of tested neutral-tone syllables with the Tone 3 syllables by the same rimes

As shown in Figure 4, the tested neutral tone contours reach to a mid to mid-low pitch range (0.5 ~ -1 semitone z-score) toward the second half of the second rime, while the low tone contours overcome the influences of the preceding tones in a faster manner and reach to the mid-low to low pitch range (-1 ~ -1.5 semitone z-score). The first sets of linear mixed model was conducted to examine the contour movement. The best-fit linear mixed model⁹ for the first half of the rime revealed significant fixed effects of position [$F(1, 11.869)=150.233, p<.001$] and tone–position interaction [$F(1, 1037.321)=139.452, p<.001$], but not for the tone [$F(1, 25.218=.025, p=.875)$]. The

⁹ The best fit model included tested tone, position, and their interaction as fixed effects, and the random effects included intercepts of subjects and items, as well as by-subject random slopes of preceding tone, tone, pair, position, and pair–position interaction.

estimates of the fixed effects, as shown in Table 6, suggests that at the 5% of the rime (baseline), the neutral tone pitch estimate was not different from the low tone one. However, the estimated pitch change for the neutral tone is -0.246 semitone per unit (10% of the rime), which is smaller than the estimated pitch change for the low tone -0.562 (= -0.246 - 0.317) semitone, and the difference was statistically significant ($p < .001$). The pitch fall of the low tone in the first half of the rimes was more substantial (2.248 semitone from 5% to 45%). On the other hand, at the second half of the rime, the best-fit model¹⁰ revealed that there was significant fixed effects of both tone [$F(1, 15.510) = 23.264, p < .001$] and position [$F(1, 967.208) = 68.610, p < .001$], but the tone–position interaction [$F(1, 966.724) = .849, p = .357$] was not significant anymore. The estimates of fixed effects in the right column of Table 6 shows that at the 55% of the rime (baseline), the low tone is 1.41 semitone lower than the neutral tone ($p < .001$)—the pitch differences between the /L/ and /Ø/ were quite substantial at the middle part of the rimes. The pitch change estimate for the neutral tone were -0.117 semitone per 10% rime, which is not significantly different from the pitch change estimate for the low tone -0.094 (= -0.117 + 0.023) semitone. This suggests that both /L/ and /Ø/ have a flattened pitch fall at the second half of the rimes. The estimates of fixed effects are shown in Figure 5 to visually present what the modeled pitch contours looks like.

Table 6: Estimates of fixed effects on the first half and the second half of the rime

Parameter	First half of the rime (5%, 15%, 25%, 35%, 45%)					Second half of the rime (55%, 65%, 75%, 85%, 95%)				
	Estimate	SE	df	t	Sig.	Estimate	SE	df	t	Sig.
Intercept	19.205	1.709	11.351	11.236	.000	18.073	1.595	11.201	11.333	.000
[Tone=L]	.063	.393	25.218	.159	.875	-1.414	.293	15.510	-4.823	.000
[Tone=N]	0	0	.	.	.	0	0	.	.	.
position	-.246	.033	12.635	-7.352	.000	-.117	.014	955.279	-8.266	.000
[Tone=L] * position	-.317	.027	1037.321	-11.809	.000	.023	.025	966.724	.921	.357
[Tone=N] * position	0	0	.	.	.	0	0	.	.	.

¹⁰ Its fixed effects included tone, position, and their interaction; its random effects included intercepts of subjects and items, as well as by-subject random slopes of preceding tone, tone, and pair.

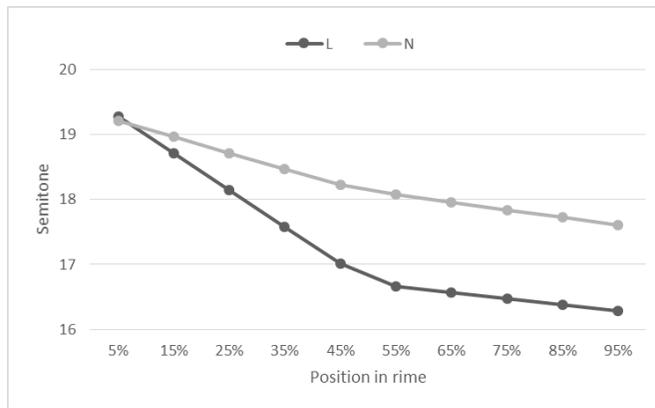


Figure 5: Estimated pitch contours of the rimes of /L/ and /∅/

The second set of the linear mixed model was carried out to investigate the effects of preceding tones and testing pair between /L/ and /∅/ at the 75% of the second rime. The best fit model included preceding tone, tested pair, tested tone, and all the interactions as fixed effects. The random effects included an intercept for subjects and by-subject random slopes for the effects of tone and the interaction between tone and preceding tone. This model revealed that at 75% of the second rime, the pitch was significantly influenced by preceding tone [$F(3, 46.962)=9.190, p<.001$] and tested tone [$F(1, 13.356)=51.545, p<.001$]. The pair and all the two-way interactions were not statistically significant, but there were a significant three-way interaction between tone, preceding tone, and pair [$F(7, 259.041)=2.929, p=.006$]. The pitch of the neutral tone at this position was significantly higher than the low tone, and the estimated mean difference between /L/ and /∅/ was 1.422 semitone ($p<.001$), similar to the estimated difference at the 55% point in the previous model. To better understand the fixed effects, the estimated marginal means with varying preceding tones and tested tones were calculated as shown in Table 7. Post-hoc pairwise comparisons using Bonferroni were carried out to 1) compare /L/ and /∅/ after different preceding tones, and 2) compare the preceding tone effect on /L/ and /∅/. The first comparison shows that /L/ and /∅/ are different after /H, HL, LH/, the estimated mean differences are 1.746, 1.427, 1.712 semitone (all $p<.000$). The second comparison indicates that at this position, the pitch of the neutral tone was influenced by preceding tone—when after /L/, it is lower than after /H, LH, HL/, the estimated mean differences are 0.906 ($p=.001$), 1.014 ($p<.001$), and 0.341 ($p=.883$) semitone respectively. On the other hand, the pitches of the low tone were quite similar regardless of its preceding tone, their differences were not statistically significant.

Table 7: Estimated marginal means by preceding tone and tested tone

Prec. Tone	/L/			/Ø/		
	Mean	SE	df	Mean	SE	df
H	16.468	1.595	11.404	18.214 ^b	1.589	11.226
HL	16.223	1.598	11.482	17.649	1.591	11.272
L	.	.	.	17.308	1.590	11.252
LH	16.610	1.593	11.321	18.322	1.589	11.216

To further investigate the significant interaction between tone, preceding tone, and pair, post-hoc pairwise comparisons using Bonferroni were carried out, and the results are shown in Table 8. This Table only compares /L/ and /Ø/ when the preceding tone was /H, HL, LH/ because [L-L] do not exist due to Tone 3 Sandhi. The results show that most of the differences between /L/ and /Ø/ after different preceding tones were significant except for the pitch differences in the *en* pair and the *zhe* pair when the preceding tone was /HL/.

Table 8: Pairwise comparisons between the pitch of /L/ and /Ø/ at the 75% of the rime by pair and preceding tone

Prec.T	/H/				/HL/				/LH/			
	Mean	SE	df	Sig.	Mean	SE	df	Sig.	Mean	SE	df	Sig.
Ce	1.070*	.359	65.073	.004	1.808*	.346	58.915	.000	1.242*	.327	53.517	.000
en	1.370*	.583	230.121	.020	.520	.467	151.883	.267	2.082*	.423	122.268	.000
RED	. ^d	.	.	.	1.522*	.478	160.868	.002	1.705*	.429	127.149	.000
zhe	1.923*	.477	159.292	.000	.820	.694	269.626	.238	2.115*	.561	220.741	.000
zi	2.721*	.510	185.017	.000	2.463*	.672	262.548	.000	1.415*	.475	161.586	.003

In addition to the pitch data, it is worth noting that the pitch contours of many tested syllables were excluded when more than five consecutive f₀ readings in their second rime were not available due to their creaky voice quality, which resulted from the lowering of the pitch. The numbers of syllables excluded due to creaky voice are shown in Table 9. Patterning with the pitch data, the low-tone syllables generally have a higher percentage of creaky voice (42.6%) than the neutral-tone syllables (21.0%). Furthermore, the preceding tone also plays a role as well—the neutral tone after /HL, L/ was found with more creaky-voice syllables (31.0% and 26.2%) than the neutral tone after /LH, H/ (11.9% and 13.9%).

Table 9: Proportions of syllables excluded due to creaky voice

	Cə		zhe		-en		-zi		red		Total	
	/L/	/∅/	/L/	/∅/	/L/	/∅/	/L/	/∅/	/L/	/∅/	/L/	/∅/
H	11/24	5/36	5/12	1/12	7/12	1/12	7/12	3/12	3/12		45.8%	13.9%
HL	8/24	10/36	8/12	4/12	5/12	3/12	9/12	8/12	4/12	1/12	47.2%	31.0%
L	-	9/36	-	4/12	-	4/12	-	5/12	-	0/12	-	26.2%
LH	7/24	4/36	6/12	2/12	3/12	1/12	5/12	3/12	4/12	0/12	34.7%	11.9%
Total	36.1%	19.4%	52.8%	22.9%	41.7%	18.8%	58.3%	39.6%	30.6%	2.8%	42.6%	21.0%

A model of Generalized Estimating Equations (GEE) was further fitted to analyze the repeatedly measure binary outcomes. The best fit model selected by the lowest QICC indicated significant effects of tested tone (L or ∅), preceding tone, pair (all $p < .001$), as well as interaction between tone and preceding tone ($p = .002$). The details of the parameter estimates are attached in Appendix B. These estimates show that the difference between /L/ and /∅/ was statistically significant ($B = 1.41$, Wald $\chi^2 = 10.462$, $p = .001$), which indicates that compared to /∅/, having a /L/ increase the likelihood of having creaky voice by 4.068 times ($\exp(B) = 4.068$). Furthermore, the effects of preceding tone and tone–preceding tone interaction show that for the neutral tone, the likelihood to be creaky is influenced by the preceding tone. For example, the neutral-tone syllables after /HL/ and /L/ were more likely to have a creaky voice compared to those after /LH/, the likelihood increased by 3.481 and 2.723 times ($p = .006$, $p = .002$) respectively. However, the effect of preceding tone was not significant for the low-tone syllables.

In summary, the neutral-tone syllables are less likely to have a creaky voice, and their pitch contours were generally higher than the pitch contours of the low tone. Also, the pitch height of the neutral tone was influenced by the preceding lexical tones. When the preceding lexical tone had a low offset (/L, HL/), the pitch contours of the neutral tone were closer to the low-tone pitch contours. The number of syllables with creaky voice (and thus excluded) also shows similar trends—a neutral-tone syllable is more likely to be creaky after /L/ and /HL/.

3.3.2 Durations

Table 10 shows the average durations of the rime of the first syllable (S1) and the second syllable (S2) of the tested disyllabic words in milliseconds (ms). The S2/S1 ratios provide a reference for the rhythm of the disyllabic words, but no further comparison will be made because the S1 rimes were not controlled in this experiment. Two S2/S1 ratios are presented in case there was speed variation: slower speech would have more influence on the (average S2)/(average S1) while the average S2/S1 ratios capture the rhythmic proportion of the disyllabic words.

Table 10 Durations and the statistical analyses of the tested pairs

Pair	Tested syllables	S1 rime duration (ms)		S2 rime duration (ms)		(avg. S2)/ (avg. S1)	Avg. S2/S1 ratio Mean
		Mean	SD	Mean	SD		
<i>-en</i>	<i>men</i> [mən ⁰]	117.3	33.1	147.3	33.1	$p=.719$	1.256
	<i>běn</i> [pən ¹]	115.3	44.8	144.5	44.8		1.253
<i>zi</i>	<i>zi</i> [tsi ⁰]	144.1	22.6	56.5	22.6	$p=.009$	0.403
	<i>zǐ</i> [tsi ¹]	157.7	28.6	78.0	28.6		0.495
<i>zhe</i>	<i>zhe</i> [tʃə ⁰]	140.7	31.2	80.0	31.2	$p<.001$	0.569
	<i>zhě</i> [tʃə ¹]	147.9	30.4	111.7	30.4		0.755
<i>Cə</i>	<i>Cə /ø/</i>	130.2	27.6	82.7	27.6	$p<.001$	0.635
	<i>Cə /L/</i>	146.0	23.8	111.7	23.8		0.765
RED	<i>shu</i> [ʃu ⁰]	114.1	20.3	91.7	20.3	$p=.760$	0.804
	<i>shǔ</i> [ʃu ¹]	170.6	27.1	87.3	27.1		0.512
	<i>shen</i> [ʃən ⁰]	139.2	28.2	146.1	28.2	$p=.740$	1.050
	<i>shěn</i> [ʃən ¹]	145.3	39.3	150.9	39.3		1.038
	<i>di</i> [ti ⁰]	99.6	33.9	96.4	33.9	$p=.691$	0.968
	<i>dǐ</i> [ti ¹]	116.8	20.4	102.2	20.4		0.875

A statistical analysis on durations shows that the best-fit model included pair, as well as interaction between tone and pair as fixed effects. For random effects, it included a random intercept for subjects and tested words, and a by-subject random slope for pair. The analysis reveals that the durations of each pair are different [$F(6, 40.858)=32.822, p<.001$] as each rime has inherently different length, but there is also an significant interaction between the pair and the tone [$F(7, 32)=8.825, p<.001$]. As shown in Table 10, the durational differences between /L/ and /ø/ within each pair were only statistically significant among the *zi*, *zhe*, and *Cə* pairs, but not for the *-en* pairs and the reduplication pairs. The model estimates of fixed effects (shown in Appendix C) illustrate that for the *Cə* pair, /L/ is 29.3 ms longer than /ø/ ($t(32)=6.071, p<.000$); for the *zhe* pair, /L/ is 37.1 ms longer ($t(32)=4.089, p<.001$); and for the *zi* pair, /L/ is 21.5ms longer ($t(32)=2.774, p=.009$). To sum up, the durational evidence suggests that neutral-tone *zi*, *zhe*, *de*, *le*, *ge* are reduced in duration, but there is no evidence suggesting that the neutral-tone *men* and the reduplicants are reduced in duration.

The comparison of the average intensities of the rimes of the two tones showed that there was no signs of intensity reduction of the neutral-tone syllables in TM, which is not surprising because intensity differences were not even found in SM. The average intensities of the neutral-tone syllables were all slightly higher than their low-tone counterparts, suggesting that the intensity co-varies with the pitch contours as observed by Lee and Zee (2008). However, the differences were not statistically significant.

3.4 Discussion

This experiment compares the pitch contours, durations and intensity of the neutral-tone syllables with the low-tone syllables of the same rimes. The results show that overall the pitch contours of the neutral tone are higher than the low tone. The analysis of the pitch contours show that both /L/ and /Ø/ had a steeper pitch fall in the first half of the rime and a shallower pitch fall in the second half of the rime. According to Y. Xu and Wang (2001), the pitch target of a lexical tone is approximated in an asymptotic manner—the f_0 first rapidly move to the pitch target, then slow down gradually over time reaching to a steady state. The asymptotic pitch contours for both /L/ and /Ø/ suggests that both tones approached to their static pitch targets respectively. The statistical analysis further show that the two have a distinct pitch target. At both the 55% and the 75% of the rime, the low tone was about 1.4 semitone lower than the pitch target of the neutral tone, and their differences were statistically significant. Considering the normalized pitch contours shown in Figure 2, the static pitch target of the neutral tone should be characterized as mid-low because by the third neutral tone, the pitch targets were around -1 semitone z-score, the mid-low range of a speaker's pitch, just slightly higher than the following low tone.

One might wonder whether the higher pitch contour of the neutral tone is a result of the pitch “undershoot” because the duration result shows that some of them are shorter than their low-tone counterparts. With a shorter duration, the f_0 movement may approach to the pitch target with a dynamic pitch contour rather than an asymptotic one, and it may not reach the pitch target either (Y. Xu, 2005, p. 229). However, the comparison between each pair after different lexical tone (Table 8) shows that those pairs with significant durational differences (*zi*, *zhe* and *Cə*) did not necessarily have larger estimated pitch differences. Therefore the higher pitch contours of the neutral tone should not be analyzed as a consequence of a shorter duration.

As shown in Figure 4, the end pitches of the neutral tone syllables were generally above -1 semitone z-score, while the low tone syllables mostly end below -1 semitone z-score. A previous perception test (Huang, 2011) also suggested that the end pitch of -1 semitone z-score is the perceptual boundary. When a pitch contour fell to the lower register (-2 ~ 0 semitone z-score), the stimulus was more likely to be identified as a low tone if the end pitch was lower than -1 semitone z-score; and the stimulus was more likely to be identified as a neutral tone if the end pitch was higher than -1 semitone z-score. The results of the perception test showed that the acoustic difference observed here is perceptually distinct to the TM listeners. Furthermore, our acoustic results showed that compared to the neutral tone, the low-tone syllables were more likely to be creaky. The low tone in Mandarin has been observed to co-occur with creaky phonation (Chao, 1968; Davison, 1991; Fon, Chiang, & Cheung, 2004; Zhu, 2012) as creaky voice is a side-effect of low f_0 (Kuang, 2013). Mandarin listeners were also

reported to utilize the creaky quality as a secondary cue to aid the low tone identification (Belotel-Grenié & Grenié, 1997; R. X. Yang, 2011), including contrasting with the neutral tone (Huang, 2011). Our results along with the perception evidence suggest that the low tone and the neutral tone have phonologically distinct pitch contours.

A more detailed factorial analysis was carried out at the 75% of the rimes. It shows that the pitch differences between /L/ and /Ø/ are most significant except for two pairs after /HL/. Also, the influence of the preceding tone was still significant on the neutral tone at this point, as suggested in Section 2, but not on the low tone. However, the pairwise comparison between /L/ and /Ø/ shows that the differences after each preceding tone were still significant. The examination of the proportion of creaky syllables also shows a similar pattern. Therefore the lower pitch contours of the neutral tone after /HL/ and /L/ should be treated as phonetically-driven rather than phonologically-targeted, and their contrast with the low tone were still maintained.

Moreover, aside from the phonetic evidence above, the phonological process of Tone 3 Sandhi also suggests that the neutral tone is distinct from the low tone (Tone 3). Tone 3 Sandhi in Mandarin requires that Tone 3 becomes Tone 2 when it is before another Tone 3. However, when a Tone 3 syllable is before a neutral-tone syllable, Tone 3 Sandhi was not triggered. This is another piece of evidence to show that the neutral tone is not the same as Tone 3.

The duration results indicate that not all of the neutral-tone syllables are reduced. The durations of the neutral-tone *zi*, *zhe* and *Cə* were all shorter than their low-tone corresponding syllables, suggesting that these neutral-tone syllables were shortened. However, the durations of *men* [mən^Ø] and the neutral-tone reduplicants were similar to their low-tone corresponding syllables. Even though these neutral-tone syllables in TM were shorter than their low-tone corresponding syllables, the durational differences were small compared to SM. Specifically, *zi* [tsi^Ø] was 28% shorter than *zǐ* [tsi^l] (56.5 vs. 78.0 ms), and the durations of *Cə* /Ø/ and *zhe* [tʃə^Ø] (about 80 ms) were 27% shorter than the durations of *Cə* /L/ and *zhě* [tʃə^l] (about 110 ms). In contrast, in a study comparing the tonal minimal pair of full-full and full-neutral in SM, Lin and Yan (1990) found that the average duration of the neutral-tone S2 rime was 53% shorter than the average duration of the Tone 4 /HL/ S2 rime (100 ms vs. 214 ms). In terms of S2/S1 ratios, TM X-*zi* [tsi^Ø], X-*zhe* [tʃə^Ø], and X-*Cə* /Ø/ were smaller than their low tone corresponding pairs, but the differences were generally around 0.1~0.2. On the other hand, the SM data showed their S2/S1 ratio of full-full was 0.89 and the S2/S1 ratio of full-neutral was 0.47. The difference between these two S2/S1 ratios was 0.42. Although Lin and Yan used a different lexical tone to compare, the results are still important for us because when a syllable is in a sentence, /HL/ is the shortest tone in SM while /L/ is the shortest tone in TM (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 split of the results in durations, with the *zi*, *zhe* and *Cə* pairs on the one hand and the *-men* and reduplicants pairs on the other, might be related to their frequency and lexical class. The neutral-tone syllables with shorter durations are all functional morphemes with high frequency, which are more susceptible to phonological reduction compared to content words which carry semantic meaning (Aylett & Turk, 2004; Bybee, Perkins, & Pagliuca, 1994; Selkirk, 1996; Z. Yang, Ramanarayanan, Byrd, & Narayanan, 2013). Specifically, *zhe* [tʃə⁰] ‘DUR’, *de* [tə⁰] ‘POSS, NMLZ’, and *le* [lə⁰] ‘PFV’ are productive verbal and nominal suffixes, *ge* [kə⁰] ‘CLF’ is the most productive generic classifier, and the diminutive suffix *zi* [tsi⁰] is a very frequent fossilized morpheme, which does not carry a semantic meaning anymore. All of these morphemes are also very frequent in Mandarin speech (See Cai and Brysbaert (2010) for their frequency). On the other hand, the plural suffix *-men* [mən⁰] is not obligatory and also limited in its distribution (Norman, 1988, p. 159; Ramsey, 1987, p. 64). As there is no grammatical number in Chinese, the plural meaning carried by the suffix *-men* [mən⁰] is not redundant. Therefore it is more similar to content words. As for the reduplications, as discussed in 1.2, the verbal reduplication in TM simply reduplicates the lexical tone, and many of the reduplicated kinship terms and nicknames adopt a /L-H/ or /L-LH/ tone regardless of their lexical tone (Hsu, 2006). Therefore it is not surprising that the tested items (all kinship terms) did not have a reduced second syllable.

Although there are multiple factors that could influence the duration of a syllable, such as focus, inherent segment lengths of the rime, or its lexical tone, these factors were controlled in this experiment. The rimes of the compared syllables were the same and were put in an identical frame sentence, and the syllables that were used for comparison with the neutral tone possess the shortest tone in TM: a low tone. If the plural suffix and the reduplicants were unstressed, we would have found the shortening of their duration as observed in the rest of the tested neutral-tone syllables (*zi*, *zhe* and *Cə*). The diverse results suggest that syllable reduction does not apply to all the neutral-tone syllables in TM. The TM neutral-tone syllables all share a similar tonal target, but no acoustic evidence supports the claim that all of the neutral-tone syllables are reduced (unstressed).

4 General discussion

The results from both experiments demonstrate that the TM neutral tone differs from the SM neutral tone in the following ways: 1) The two dialects exhibit different pitch patterns on their neutral tone. As demonstrated in the first experiment, the TM neutral-tone syllables do not show post-L rising, instead they have a static pitch target in the mid-low to low range, and the comparison with the low tone in the second

experiment confirms that the TM neutral tone has a static mid-low pitch target; 2) Compared to the SM neutral tone, the TM neutral tone shows a much stronger articulation strength. The mid-low pitch target of the TM neutral tone was approached in a faster manner compared to that of the SM neutral tone. As shown in the first experiment, the carry-over effect had disappeared in TM at the third consecutive neutral tone, but it remained strong in SM; 3) While the tonal neutralization is tightly connected to stress in SM—a syllable is toneless because it is unstressed—it does not seem to be the case for TM. As mentioned in 1.1, the SM “neutral-tone” syllables were in fact unstressed syllables because the reduction and the pitch patterns were found in all the unstressed syllables, whether they are lexically-marked as unstressed/neutral tone or not. On the other hand, the TM neutral tone does not seem to be tightly connected to stress in TM. It is less common and strictly lexically determined. The presence of the pitch pattern is not found in the de-stressed syllables—it is only found in those syllables which are specified as having a neutral tone underlyingly at the lexical level. Moreover, the durational comparison in Experiment 2 shows that not all of the neutral-tone syllables are reduced. Therefore, it is hard to argue for a correlation between stress and the process of tonal neutralization in Taiwan Mandarin.

Based on the acoustic evidence listed above, I propose that the neutral-tone syllables in TM should be analyzed as unaccented, and the lexical tone in these syllables is neutralized into a mid-low tone. The mid-low pitch target can be interpreted as an unspecified position with a reduced pitch articulation, similar to [ə] in the vowel space. The term *unaccented* is used to show that pitch is the only acoustic cue to mark prominence, such as in pitch-accent languages like Japanese, as opposed to stress-accent language that might utilize other acoustic cues such as durations and intensity to mark prominence (van der Hulst, 2010). By analyzing these syllables as unaccented, we can capture the fact that durational cues are not consistent among the TM neutral-tone syllables, only some of them—mostly frequent function words—are prone to reduction. Therefore, describing the TM “neutral tone” syllables as “unaccented” would be a better analysis than describing all of them as “unstressed”. There have been cases of other Mandarin dialects being documented as having unaccented syllables. For example, in the Xinjiang Mandarin dialects Barköl Mandarin and Ürümqi Mandarin, their neutral-tone syllables are analyzed as being unaccented because they undergo tone loss, but are not reduced, and the pitches of these unaccented syllables are conditioned by the preceding lexical tone (D. Cao, 1988; Wei, 2011). Similarly, the TM neutral-tone syllables also undergo tone loss but they are not reduced. The difference is that the TM unaccented syllables are neutralized into a mid-low tone.

There are three alternative ways to analyse the TM neutral tone. In the following discussion I will illustrate why they are inadequate, and thus the TM neutral tone should be analysed as unaccented.

First, the TM neutral-tone syllables can be analysed as having a mid-low lexical tone, which I term the fifth tone. It is because they have a distinct pitch target which differs from the other four lexical tones, and the mid-low pitch target cannot be predicted by the stress pattern and has to be specified underlyingly. Although the TM neutral tone is more lexical-tone like compared to the SM neutral tone, it should be noted that its articulatory strength is still weaker than lexical tones in the similar pitch range. As shown in Experiment 1, although the pitch contours start to converge immediately in the first neutral tone, the carry-over effects were not overcome until the third neutral tone. Although carry-over effects can be still significant on lexical tones, they are usually observed on lexical tones with higher offsets such as /H/ and /LH/ (Huang, 2013; Y. Xu, 1997). This piece of evidence suggests that these syllables do not seem to possess a lexical tone either. If these neutral-tone syllables possess a mid-low lexical tone, their pitch target should be implemented in a faster manner. Furthermore, an extra mid-low level tone will be hard to fit in the tonal structure of /H, LH, L, HL/, or the newly proposed tonal structure of Taiwan Mandarin /H, M, L, HM/ (Huang, 2017).

Second, one might analyze these TM neutral-tone syllables the same as those in SM—“these syllables were neutralized because they are unstressed”, and the difference between the two dialects is that in TM, the unstressed syllables are neutralized into a mid-low tone. This analysis would be similar to one of the analyses of Standard Thai unstressed syllables: lexical tones are said to be neutralized into a mid tone (Abramson, 1962). Nonetheless, this analysis needs to be rejected because our results show that not all the neutral-tone syllables are unstressed—some do not show a reduced duration. If the neutralized tone is a result of being unstressed, shorter durations are expected to be found in all the neutral-tone syllables.

Another possibility, suggested by some linguists in China (See L. Liu (2002)), is to analyze these neutral-tone syllables as a result of tone sandhi, i.e., syllables that undergo this tone sandhi change their tones to an unspecified tone with a mid-low pitch target. The problem with this analysis is determining the condition of this tone sandhi process. A tone sandhi process might be triggered by adjacent tones, e.g. Mandarin Tone 3 Sandhi; or by its position in a prosodic domain, e.g. Southern Min tone sandhi. However, this tone sandhi process is neither conditioned by its adjacent tones, nor by its position in a prosodic domain. Although they all locate at the right edge of a foot (and possibly other prosodic units), other lexical tones might occur in these positions as well. Therefore, it is not possible to predict the condition of this tone sandhi. In order to adopt this analysis, these linguists argue that the trigger of this tone sandhi is lexically determined, which makes no generalization about this process. This analysis is like stating that all the [d] in English are actually a result of a voicing of /t/, and this voicing process is lexically determined—of course no linguist actually claims this, instead the logical analysis is that /d/ is specified underlyingly. Similarly, rather than specifying the condition of the tone sandhi lexically, one could simply specify the unaccented syllables underlyingly.

The synchronic analysis of unaccented syllables is likely to be one of the manifestations of how the Mandarin neutral-tone syllables were adopted in Taiwan. As discussed in 1.2, the rhythmic features are manifested in the following ways: first of all, TM has less full-neutral disyllabic words. The evidence from the dictionary, computer input system, and online typing conventions suggests that the majority of the prescriptive neutral-tone syllables in compound words in SM are produced with lexical tones. Secondly, there is evidence suggesting that the post-L rising in the SM neutral tone was reanalyzed by the TM speakers as having a /H/ tone. When a Tone 3 is reduplicated in nicknames or kinship terms, a /L-H/ tonal pattern is often used to show endearment (Hsu, 2006). TM speakers likely produce the post-L rising without a reduced duration, and thus reanalyzed the post-L “neutral-tone syllables” as /H/. In comparison, the SM speakers do not treat the post-L rising as a /H/ because the post-L rising is always accompanied by a shorter duration, which serves as an additional perceptual cue. Therefore, the SM listeners identify the high pitch after /L/ as a neutral-tone, not a /H/ tone. This study demonstrates another manifestation of the TM rhythmic features on tones. It shows that, for the rest of the remaining prescriptive neutral tone syllables, at least for those that were examined in this study, the TM speakers likely reanalyzed the SM neutralized tone patterns into a mid-low tone with a weak articulation strength instead of adopting one of the existing lexical tones. This is likely because the SM neutral-tone syllables all have falling pitch contours reaching to the low register except for when they occur after /L/. When the TM speakers did not reduce the duration for these neutral-tone syllables, the SM neutralized pitch patterns were likely treated as a specific pitch contour, rather than a result of destressing and subsequent tone loss. Consequently, these falling pitch contours were generalized by the TM speakers as an unaccented syllable with a mid-low pitch target.

It is worth mentioning that these unaccented syllables are rather peripheral and might be moving towards merging with Tone 3. Acoustic data show their voice quality and pitch contours are sometimes similar to Tone 3, especially when they follow a falling tone or a low tone. Evidence in writing also suggests that the TM speakers are aware of this phonetic similarity. For example, recently some young TM online users use the character 惹 *rě* /zə^L/ ‘to provoke’ as a substitute for 了 *le* /lə⁰/ ‘PFV’ when posting comments online. The character 惹 *rě* /zə^L/ can be used as a phonogram because the *lě* /lə^L/ syllable does not exist in Mandarin and many TM speakers merged /z/ with /l/; I have also observed that online users write 鵝紫 *ézi* /ɣ^{LH}-tsi^L/ ‘goose-purple’ as a substitute for 兒子 *érzi* /ɣ^{LH}-tsi⁰/ ‘son’. These pieces of evidence suggest that TM speakers are aware of the similarity between Tone 3 and the unaccented syllables, and in fact are willing to use Tone 3 to mark them if given the chance to create their own writing.

Furthermore, phonologically some TM speakers seem to treat the neutral-tone syllables as a Tone 3. I conducted a “wug test” on the four productive neutral-tone syllables perfective *le*, durative *zhe*, possessive *de*, and plural *men*. Twelve subjects

were given obsolete characters 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 each made-up character. Interestingly, one subject consistently applied Tone 3 Sandhi on the novel Tone 3 syllable before the neutral-tone syllable, suggesting that he treated the neutral tone as Tone 3 phonologically. This shows that possibly for some TM speakers, the low pitch of the neutral tone has led them to categorize the neutral-tone syllables as a low-tone syllable. Although Tone 3 Sandhi is not triggered in disyllabic L-Ø words because the pitch patterns are lexicalized, in a completely novel situation these neutral-tone syllables are treated like a low tone. Sociolinguistic research on a larger scale needs to be carried out to investigate whether this is a developing tendency, and, if so, how it affects speakers' production.

Unfortunately, this study was not able to examine all the neutral-tone syllables. The ones that were not examined include: 1) those disyllabic compound words that are prescriptively marked as having a neutral tone in computer input systems, 2) final particles, and 3) more nominal reduplications such as other kinship terms and verbal reduplications. However, despite a limited sample, the tested neutral-tone syllables in this study were among those words/syllables that were most frequently used in Mandarin. The emerging pitch contours of the unaccented syllables found in this study should be representative considering that they comprise a large portion of neutral-tone input/output in speech. Furthermore, the split durational results mean that whether or not those untested neutral-tone syllables reduced will not affect the conclusion of this study. Nevertheless, it will still be important to examine more neutral-tone syllables in future studies to investigate how neutral-tone syllables developed, and how they differ between different word classes and frequency. More work on different lexical-tone languages or other Mandarin dialects will also be useful in observing the interaction between tone and stress.

5 Conclusion

This study shows that unlike Standard Mandarin, the prescriptive neutral-tone syllables in TM either possess one of the four lexical tones, or they have a static mid-low pitch target that is different from other lexical tones. Furthermore, not all of these neutral-tone syllables are reduced. In other words, the tonal neutralization is not a result of the de-stressing. I propose that the TM "neutral-tone" should therefore be analyzed as unaccented with a mid-low pitch target. While stress can affect the pitch patterns of lexical tones through different ways of tone neutralization, this study demonstrates that when the contrast between stressed and unstressed syllables is less pronounced, the unstressed syllables cannot be "reversed" back to their original lexical tones, and an analysis of unaccented syllables thus needs to be proposed.

References

- Abramson, A. S. (1962). *The vowels and tones of standard Thai: Acoustical measurements and experiments* (Vol. 20). Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics.
- Arvaniti, A. (2000). The phonetics of stress in Greek. *Journal of Greek Linguistics*, 1(1), 9-39.
- Aylett, M., & Turk, A. (2004). The smooth signal redundancy hypothesis: A functional explanation for relationships between redundancy, prosodic prominence, and duration in spontaneous speech. *Language and Speech*, 47(1), 31-56.
- Bauer, R. S., & Benedict, P. K. (1997). *Modern Cantonese phonology*. Berlin; New York: Mouton de Gruyter.
- Beckman, M. (1986). *Stress and non-stress accent*. Dordrecht: Foris.
- Belotel-Grenié, A., & Grenié, M. (1997). Types de phonation et tons en chinois standard. *Cahiers de linguistique - Asie orientale*, 26(2), 249-279.
- Bradley, D. (1982). Register in Burmese. *Pacific Linguistics. Series A. Occasional Papers*(62), 117.
- Brunelle, M. (2017). Stress and phrasal prominence in tone languages: The case of Southern Vietnamese. *Journal of the International Phonetic Association*, 1-38.
- Bybee, J., Perkins, R., & Pagliuca, W. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. Chicago: University of Chicago Press.
- Cai, Q., & Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLoS one*, 5(6), e10729.
- Cao, D. (1988). Balikun hua de qingsheng ci 巴里坤话的轻音词 [Unstressed syllables in Barköl Mandarin]. *Xinjiang daxue xuebao 新疆大学学报 [Journal of Xinjiang University]*, 3, 91-102.
- Cao, J. (1986). Putonghua qingsheng yinjie texing fenxi 普通话轻声音节特性分析 [An acoustic analysis of the neutral-tone syllables in Mandarin]. *Yingyong shengxue 应用声学 [Applied Acoustics]*, 5(4), 1-6.
- Chao, Y.-r. (1932). A preliminary study of English intonation (with American variants) and its Chinese equivalent *The Tsai Yuan P'ei anniversary volume (supplementary Vol. 1 of the Bulletin of the Institute of History and Philology)* (pp. 105-156).
- Chao, Y.-r. (1933). Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology*, 4(3), 121-134.
- Chao, Y.-R. (1956). Tone, intonation, singsong, chanting, recitative tonal composition and atonal composition in Chinese. In M. Halle (Ed.), *For Roman Jakobson* (pp. 52-59). The Hague: Mouton.
- Chao, Y.-R. (1968). *A Grammar of Spoken Chinese*: University of California: Berkeley and Los Angeles.
- Chen, Y., & Xu, Y. (2006). Production of weak elements in speech-evidence from F0 patterns of neutral tone in standard Chinese. *Phonetica*, 63(1), 47-75.
- Cheng, C.-c. (1973). *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.
- Crystal, T. H., & House, A. S. (1988). Segmental durations in connected-speech signals: Syllabic stress. *The Journal of the Acoustical Society of America*, 83(4), 1574-1585.

- Dauer, R. M. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11, 51-62.
- Davison, D. S. (1991). An acoustic study of so-called creaky voice in Tianjin Mandarin. *UCLA Working Papers in Phonetics*, 78, 50-57.
- De Jong, K., & Zawaydeh, B. A. (1999). Stress, duration, and intonation in Arabic word-level prosody. *Journal of Phonetics*, 27(1), 3-22.
- Deng, D., Shi, F., & Lu, S. (2008). Contrastive analysis on tone between Putonghua and Taiwan Mandarin. *Chinese Journal of Acoustics*, 27(2)
- Dreher, J. J., & Lee, P.-c. e. (1968). Instrumental investigation of single and paired Mandarin tonemes. *Monumenta serica*, 343-373.
- Duanmu, S. (2000). *The Phonology of Standard Chinese*. Oxford: Oxford University Press.
- Duanmu, S. (2007). *The Phonology of Standard Chinese* (2nd ed.). Oxford: Oxford University Press.
- Ewen, C. J., & Hulst, H. v. d. (2001). *The phonological structure of words*. Cambridge: Cambridge University Press.
- Fon, J., & Chiang, W.-Y. (1999). What does Chao have to say about tones? - a case study of Taiwan Mandarin. *Journal of Chinese Linguistics*, 27(1), 15-37.
- Fon, J., Chiang, W.-Y., & Cheung, H. (2004). Production and perception of the two dipping tones (Tone 2 and Tone 3) in Taiwan Mandarin. *Journal of Chinese Linguistics*, 32(2), 249-281.
- Fry, D. B. (1955). Duration and intensity as physical correlates of linguistic stress. *The Journal of the Acoustical Society of America*, 27(4), 765-768.
- Gandour, J., Tumtavitikul, A., & Sattthamnuwong, N. (1999). Effects of speaking rate on Thai tones. *Phonetica*, 56(3-4), 123-134.
- Gao, Y. (1980). Beijinghua de qingsheng wenti 北京话的轻声问题 [Issues of the neutral tone in Beijing Mandarin]. *Yuyan Jiaoxue Yu Yanjiu 语言教学与研究 [Language teaching and research]*, 2, 82-98.
- Garellek, M., & White, J. (2015). Phonetics of Tongan stress. *Journal of the International Phonetic Association*, 45(01), 13-34.
- Gordon, M., & Applebaum, A. (2010). Acoustic correlates of stress in Turkish Kabardian. *Journal of the International Phonetic Association*, 40(1), 35-58.
- Gruber, J. F. (2011). *An articulatory, acoustic, and auditory study of Burmese tone*. (PhD), Georgetown University, Washington D.C.
- Her, O.-S. (2010). Lun Taiwan Huayu de zaidihua [On the indigenization of Taiwan Mandarin] [On the indigenization of Taiwan Mandarin]. *Journal of Macau Linguistics Association*, 35(1), 19-29.
- Hombert, J.-M. (1977). Consonant types, vowel height and tone in Yoruba. *Studies in African Linguistics*, 8(2), 173-190.
- Hombert, J.-M., Ohala, J. J., & Ewan, W. G. (1979). Phonetic explanations for the development of tones. *Language*, 55, 37-58.
- Howie, J. (1976). *An Acoustic Study of Mandarin Tones and Vowels*. London: Cambridge University Press.

- Hsu, H.-c. (2006). Revisiting tone and prominence in Chinese. *Language and Linguistics*, 7(1), 109-137.
- Huang, K. (2011). On the perception of the neutral tone in Taiwan Mandarin. *Proceedings of the 17th International Congress of Phonetic Sciences (ICPhS XVII)* (pp. 910-913).
- Huang, K. (2013). *Revisiting the four tones in Taiwan Mandarin*. Paper presented at the Proceedings of the International Conference on Phonetics of the Languages in China (ICPLC-2013), Hong Kong.
- Huang, K. (2017). From pitch contour variation to tone change: An analysis of the phonological representations of the tones in Taiwan Mandarin. *International Journal of Chinese Linguistics*, 4(2), 273-307. <https://doi.org/10.1075/ijchl.16016.hua>
- Jing, S. (2002). *Xiandai hanyu qingsheng dongtai yanjiu 现代汉语轻声动态研究 [A dynamic study of neutral tone in Modern Chinese]* Beijing: Minzu Chubanshe.
- Kochanski, G., Grabe, E., Coleman, J., & Rosner, B. (2005). Loudness predicts prominence: Fundamental frequency lends little. *The Journal of the Acoustical Society of America*, 118(2), 1038-1054.
- Kratochvil, P. (1968). *The Chinese language today*. London: Hutchinson University Library.
- Kuang, J. (2013). *Phonation in Tonal Contrasts*. University of California, Los Angeles.
- Kubler, C. C. (1985). The influence of Southern Min on the Mandarin of Taiwan. *Anthropological Linguistics*, 27(2), 156-176.
- Lee, W.-S., & Zee, E. (2008). Prosodic characteristics of the neutral Tone in Beijing Mandarin. *Journal of Chinese Linguistics*, 36(1), 1-29.
- Li, J. (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(4), 2457-2457.
- Li, W. (1981). Shi lun qingsheng he zhong yin 试论轻声和重音 [A preliminary discussion on stressless and stressed syllables]. *Zhongguo Yuwen 中国语文 [Studies of the Chinese Language]*, 1, 35-40.
- Li, Z. (2003). *The phonetics and phonology of tone mapping in a constraint-based approach*. (PhD), Massachusetts Institute of Technology.
- Lieberman, P. (1960). Some acoustic correlates of word stress in American English. *The Journal of the Acoustical Society of America*, 32(4), 451-454.
- Lin, M., & Yan, J. (1980). Beijinghua qingsheng de shengxue xingzhi 北京话轻声的声学性质 [The acoustic features of the neutral tone in Beijing dialect] [Acoustic characteristics of neutral tone in Beijing Mandarin]. *Fangyan 方言 [Dialect]*, 3, 166-178.
- Lin, M., & Yan, J. (1990). Putonghua qingsheng yu qingzhong yin 普通话轻声与轻重音 [The neutral tone and stress in Mandarin]. *Yuyan Jiaoxue Yu Yanjiu 语言教学与研究 [Language teaching and research]*, 3, 88-104.
- Liu, J., & Zhang, J. (2012). A production and perception study on tonal neutralization in Nanchang Chinese. *The Journal of the Acoustical Society of America*, 132(3), 2004-2004.
- Liu, L. (2002). Ershi shiji Hanyu qingsheng yanjiu zongshu [An overview of researches on Chinese neutral tone in the 20th century]. *Yuwen Yanjiu [Language Studies]*, 3, 43-47.

- Llisterri, J., Machuca, M., de la Mota, C., Riera, M., & Ríos, A. (2003). *The perception of lexical stress in Spanish*. Paper presented at the Proceedings of the 15th International Congress of Phonetic Sciences. Barcelona.
- Luo, C., & Wang, j. (1957). *Putong yuyinxue gangyao 普通語音學綱要 [Outline of general phonetics]* (Revised ed.). Beijing: Beijing: Kexue Chubanshe (Reprinted and partially revised in 2002 by Beijing: Shangwu Yinshuguan).
- Moren, B., & Zsiga, E. (2006). The lexical and post-lexical phonology of Thai tones. *Natural Language and Linguistic Theory*, 24(1), 113-178.
- Myers, J., & Tsay, J. (2003). Investigating the phonetics of Mandarin tone sandhi. *Taiwan Journal of Linguistics*, 1(1), 29-68.
- Norman, J. (1988). *Chinese*. Cambridge: Cambridge University Press.
- Ohala, J. J. (1972). How is pitch lowered? *Journal of the Acoustical Society of America*, 52, 124.
- Ortega-Llebaria, M., & Prieto, P. (2007). Disentangling stress from accent in Spanish: production patterns of the stress contrast in deaccented syllables. In P. Prieto, J. Mascaro & M. J. Sole (Eds.), *Segmental and prosodic issues in romance phonology*. Amsterdam, Philadelphia: John Benjamins.
- Potisuk, S., Gandour, J., & Harper, M. (1994). F0 correlates of stress in Thai. *Linguistics of the Tibeto-Burman area*, 17(2), 1-27.
- Potisuk, S., Gandour, J., & Harper, M. (1996). Acoustic correlates of stress in Thai. *Phonetica*, 53(4), 200-220.
- Prieto, P., & Ortega-Llebaria, M. (2006). *Stress and Accent in Catalan and Spanish: Patterns of duration, vowel quality, overall intensity, and spectral balance*. Paper presented at the Proceedings of Speech Prosody.
- Ramsey, S. R. (1987). *The languages of China*. Princeton: Princeton University Press.
- Sandel, T. (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.
- Selkirk, E. (1996). The prosodic structure of function words. In J. Beckmen, L. Walsh Dickey & S. Urbanczyk (Eds.), *Papers in optimality theory. University of Massachusetts occasional papers in linguistics. 18*: Graduate linguistic student association.
- Shen, X. (1990). *The prosody of Mandarin Chinese*. Berkeley: University of California Press.
- Shi, F., & Deng, D. (2006). Putonghua yu Taiwan guoyu de yuyin duibi [Phonetic contrasts between Standard Mandarin and Taiwan Mandarin]. In D.-a. Ho, H. S. Cheung, W. Pan & F. Wu (Eds.), *Shan gao shui chang: Ding Bangxin xian sheng qi zhi shou qing lun wen ji [Linguistic studies in Chinese and neighboring languages : festschrift in honor of professor Pang-hsin Ting on his 70th birthday]* (pp. 373-391). Taiwan: Institute of Linguistics, Academia Sinica.
- Sluijter, A. M., & Van Heuven, V. J. (1996). Spectral balance as an acoustic correlate of linguistic stress. *The Journal of the Acoustical Society of America*, 100(4), 2471-2485.
- Swihart, D.-A. W. (2003). The two Mandarins: Putonghua and Guoyu. *Journal of the Chinese Language Teachers Association*, 38(3), 103-118.
- Torgerson, R. C., Jr. (2005). *A comparison of Beijing and Taiwan Mandarin tone register: an acoustic analysis of three native speech styles*. (MA), Birgham Young University.

- Tsao, F.-f. (2000). Taishi riyu yu Taiwn guoyu - bainian lai zai Taiwan fasheng de liangge yuyan jiechu shili 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 漢學研究 [Chinese Studies]*, 18, 273-297.
- Tseng, C.-C. (2004, March 28-31). *Prosodic Properties of Intonation in Two Major Varieties of Mandarin Chinese: Mainland China vs. Taiwan*. Paper presented at the International Symposium on Tonal Aspects of Languages With Emphasis on Tone Languages, Beijing, China.
- van der Hulst, H. (2010). Word accent: Terms, typologies and theories. In H. v. d. Hulst, R. Goedemans & E. v. Zanten (Eds.), *A Survey of Word Accentual Patterns in the Languages of the World* (pp. 3-54). Berlin/New York: Walter de Gruyter.
- Wei, Y. (2011). Wulumuqi hanyu fangyan qingsheng de yuyin xingzhi ji youxuanlun fenxi 乌鲁木齐汉语方言轻声的语音性质及优选论分析 [An acoustic and OT analysis of the neutral tone in Ürümqi Mandarin]. *Xinan nongye daxue xuebao (shehui kexue ban) 西南农业大学学报(社会科学版)*, [Journal of Southwest Agricultural University(Social Sciences Edition)], 1, 115-118.
- Williams, B. (1985). Pitch and duration in Welsh stress perception: The implications for intonation. *Journal of Phonetics*, 13, 381-406.
- Xu, S. (1980). *Putonghua yuyin zhishi 普通话语音知识 [Phonology of Standard Mandarin]*. Beijing: Wenzhi Gaige Chubanshe.
- Xu, Y. (1997). Contextual tonal variations in Mandarin. *Journal of Phonetics*, 25, 61-83.
- Xu, Y. (1999). *F0 peak delay: when, where, and why it occurs*. Paper presented at the International Congress of Phonetic Science, San Francisco.
- Xu, Y. (2005). Speech melody as articulatorily implemented communicative functions. *Speech Communication*, 46(3), 220-251.
- Xu, Y., & Wang, Q. E. (2001). Pitch targets and their realization: Evidence from Mandarin Chinese. *Speech Communication*, 33, 319-337.
- Yang, R. X. (2011). The Phonation factor in the categorical perception of Mandarin tones. *Proceedings of the 17th International Congress of Phonetic Sciences*, 2204-2207.
- Yang, Z., Ramanarayanan, V., Byrd, D., & Narayanan, S. (2013). *The effect of word frequency and lexical class on articulatory-acoustic coupling*. Paper presented at the INTERSPEECH.
- Zhu, X. (2012). Multi registers and four levels: A new tonal model. *Journal of Chinese Linguistics*, 40(1), 1-18.

Appendix A: Sentences used to elicit consecutive neutral tones

X-N-N-N-Y	X-REDUP.-PL-POSS-Y	<p><i>tā shuō X~X-men-de hǎo/zhòngyào duō le</i> [tʰaʰ ʃwoʰ mən tə hao˨/tʃoŋʰljaoʰl twoʰ lə] 3S say X~RED-PL-POSS good/important more ASP ‘He said that Xs’ are much better/more important.’</p>
Y: /L/	<p><i>tā shuō mā~ma-men-de hǎo duō le</i> 他說媽媽們的好多了 [tʰaʰ ʃwoʰ maʰma˨mən˨tə˨ hao˨ twoʰ lə] ‘He said that the mothers’ are much better.’</p> <p><i>tā shuō pó~po-men-de hǎo duō le</i> 他說婆婆們的好多了 [tʰaʰ ʃwoʰ pʰwo˨phwo˨mən˨tə˨ hao˨ twoʰ lə] ‘He said that the mother-in-laws’ are much better.’</p> <p><i>tā shuō nǎi~nai-men-de hǎo duō le</i> 他說奶奶們的好多了 [tʰaʰ ʃwoʰ nai˨nai˨mən˨tə˨ hao˨ twoʰ lə] ‘He said that the grandmas’ are much better.’</p> <p><i>tā shuō mèi~mei men de hǎo duō le</i> 他說妹妹們的好多了 [tʰaʰ ʃwoʰ mɛ˨mɛ˨mən˨tə˨ hao˨ twoʰ lə] ‘He said that the sisters’ are much better.’</p>	
Y: /HL/	<p><i>tā shuō mā~ma-men-de zhòngyào duō le</i> 他說媽媽們的重要多了 [tʰaʰ ʃwoʰ maʰmaʰmən˨tə˨ tʃoŋʰljaoʰl twoʰ lə] ‘He said that the mothers’ are much more important.’</p> <p><i>tā shuō pó~po-men-de zhòngyào duō le</i> 他說婆婆們的重要多了 [tʰaʰ ʃwoʰ pʰwo˨phwo˨mən˨tə˨ tʃoŋʰljaoʰl twoʰ lə] ‘He said that the mother-in-laws’ are much more important.’</p> <p><i>tā shuō nǎi~nai-men-de zhòngyào duō le</i> 他說奶奶們的重要多了 [tʰaʰ ʃwoʰ nai˨nai˨mən˨tə˨ tʃoŋʰljaoʰl twoʰ lə] ‘He said that the grandmas’ are much more important.’</p> <p><i>tā shuō mèi~mei-men-de zhòngyào duō le</i> 他說妹妹們的重要多了 [tʰaʰ ʃwoʰ mɛ˨mɛ˨mən˨tə˨ tʃoŋʰljaoʰl twoʰ lə] ‘He said that the sisters’ are much more important.’</p>	
X-N-N-Y	X- PFV -NOM.-Y	<p><i>tā xǐhuān X-le-de nà-[CLF] [noun]</i> [tʰaʰ ɕi˨hwanʰ lə˨tə˨ naʰl] 3S likes verb- PFV-NMLZ that-CLF noun. ‘He likes the one that ~.’</p> <p><i>nàge X-le-de bǐjiào [adjective]</i> [naʰlkə˨ lə˨tə˨ pi˨tɕjaoʰl] that verb- PFV-NMLZ compare adjective. ‘The one that [verb] is more [adjective].’</p>
Y: /HL/	<p><i>tā xǐhuān chāi-le de nàjiān diàn.</i> 他喜歡拆了的那間店 [tʰaʰ ɕi˨hwanʰ tʃaɪ˨lə˨tə˨ naʰlɕjanʰ tjanʰl] ‘He likes the shop that was torn down.’</p> <p><i>tā xǐhuān táo-le de nàge rén.</i> 他喜歡逃了的那個人 [tʰaʰ ɕi˨hwanʰ tʰao˨lɛ˨tə˨ naʰlkə˨ zənʰl] ‘He likes the person that escaped.’</p> <p><i>tā xǐhuān pǎo-le de nàge rén.</i> 他喜歡跑了的那個人 [tʰaʰ ɕi˨hwanʰ pʰao˨lɛ˨tə˨ naʰlkə˨ zənʰl] ‘He likes the person that ran away.’</p>	

	<p><i>tā xǐhuān xiǎo-le de nàge rén.</i> 他喜歡笑了的那個人 [tʰaʰ ɕi˥hwan˥ ɕjao˥lɿə̯tə̯ na˥kə̯ zən˥] ‘He likes the one that smiled.’</p>
Y:/L/	<p><i>nàge chāi-le de bǐjiào piányi</i> 那個拆了的比較便宜 [na˥kə̯ tʃai˥lɿə̯tə̯ pi˥tɕjao˥lɿ pʰjan˥lɿ] ‘The one that was torn down is cheaper’</p> <p><i>nàge táo-le de bǐjiào qǐngsōng</i> 那個逃了的比較輕鬆 [na˥kə̯ tʰao˥lɿə̯tə̯ pi˥tɕjao˥lɿ tɕiŋ˥son˥] ‘The one that was escaped is more relaxed’</p> <p><i>nàge pǎo-le de bǐjiào kělián</i> 那個跑了的比較可憐 [na˥kə̯ pʰao˥lɿə̯tə̯ pi˥tɕjao˥lɿ kʰə̯lɿjan˥] ‘The one that ran away is more pathetic’</p> <p><i>nàge xiào-le de bǐjiào kuàilè</i> 那個笑了的比較快樂 [na˥kə̯ ɕjao˥lɿə̯tə̯ pi˥tɕjao˥lɿ kwaɪ˥lɿ] ‘The one that smiled is happier’</p>
X-N-N-Y	<p>X-DUR-NMLZ-Y</p> <p><i>tā xǐhuān X-zhe-de nà-ge rén</i> [tʰaʰ ɕi˥hwan˥ tʃə̯tə̯ na˥] 3S like verb-DUR-NMLZ that-CLF person ‘He likes the person who is [verb]-ing.’</p> <p><i>nàge X-zhe-de bǐjiào [adjective]</i> [na˥kə̯ tʃə̯tə̯ pi˥tɕjao˥lɿ] that verb-DUR-NMLZ more adjective ‘That [verb]-ing one is more [adjective].’</p>
Y:/HL/	<p><i>tā xǐhuān dūn-zhe-de nàge rén.</i> 他喜歡蹲著的那個人 [tʰaʰ ɕi˥hwan˥ twən˥tʃə̯tə̯ na˥kə̯ zən˥] ‘He likes the person that was squatting.’</p> <p><i>tā xǐhuān xián-zhe-de nàge rén.</i> 他喜歡閒著的那個人 [tʰaʰ ɕi˥hwan˥ ɕjan˥tʃə̯tə̯ na˥kə̯ zən˥] ‘He likes the person that was free.’</p> <p><i>tā xǐhuān tāng-zhe-de nàge rén.</i> 他喜歡躺著的那個人 [tʰaʰ ɕi˥hwan˥ tʰaŋ˥tʃə̯tə̯ na˥kə̯ zən˥] ‘He likes the person that was lying down.’</p> <p><i>tā xǐhuān zhàn-zhe-de nàge rén.</i> 他喜歡站著的那個人 [tʰaʰ ɕi˥hwan˥ tʃan˥tʃə̯tə̯ na˥kə̯ zən˥] ‘He likes the person that was standing.’</p>
Y:/L/	<p><i>Nàge dūn-zhe-de bǐjiào suíbiàn</i> 那個蹲著的比較隨便 [na˥kə̯ twən˥tʃə̯tə̯ pi˥tɕjao˥lɿ swɛ˥lɿpjan˥] ‘That squatting one is more casual.’</p> <p><i>Nàge xián-zhe-de bǐjiào tǎoyàn</i> 那個閒著的比較討厭 [na˥kə̯ ɕjan˥tʃə̯tə̯ pi˥tɕjao˥lɿ tʰao˥jan˥] ‘That free one is more annoying.’</p> <p><i>Nàge tāng-zhe-de bǐjiào lǎnduò</i> 那個躺著的比較懶惰 [na˥kə̯ tʰaŋ˥tʃə̯tə̯ pi˥tɕjao˥lɿ lan˥two˥] ‘That lying down one is lazier.’</p> <p><i>Nàge zhàn-zhe-de bǐjiào rènzhēn</i> 那個站著的比較認真 [na˥kə̯ tʃan˥tʃə̯tə̯ pi˥tɕjao˥lɿ zən˥tʃən˥] ‘That standing one is more earnest.’</p>

Appendix B: Parameter Estimates on the occurrences of creaky syllable

Parameter	B	Std. Error	Hypothesis Test			Exp(B)	95% Wald Confidence Interval for Exp(B)	
			Wald χ^2	df	Sig.		Lower	Upper
(Intercept)	-1.754	.4656	14.194	1	.000	.173	.069	.431
[Tone=L]	1.403	.4338	10.462	1	.001	4.068	1.738	9.519
[Tone=N]	0 ^a	1	.	.
[PrecT=H]	.070	.4389	.025	1	.874	1.072	.454	2.534
[PrecT=HL]	1.247	.4549	7.516	1	.006	3.481	1.427	8.489
[PrecT=L]	1.002	.3238	9.571	1	.002	2.723	1.444	5.137
[PrecT=LH]	0 ^a	1	.	.
[Tone=L] * [PrecT=H]	.425	.5627	.570	1	.450	1.530	.508	4.608
[Tone=L] * [PrecT=HL]	-.693	.6220	1.241	1	.265	.500	.148	1.693
[Tone=L] * [PrecT=LH]	0 ^a	1	.	.
[Tone=N] * [PrecT=H]	0 ^a	1	.	.
[Tone=N] * [PrecT=HL]	0 ^a	1	.	.
[Tone=N] * [PrecT=L]	0 ^a	1	.	.
[Tone=N] * [PrecT=LH]	0 ^a	1	.	.
[pair=e]	-.446	.1862	5.726	1	.017	.640	.445	.923
[pair=en]	-.362	.2344	2.381	1	.123	.697	.440	1.103
[pair=i]	.552	.2864	3.710	1	.054	1.736	.990	3.043
[pair=redup]	-1.237	.2967	17.383	1	.000	.290	.162	.519
[pair=zhe]	0 ^a	1	.	.
(Scale)	1							

Appendix C: Estimates of fixed effects of S2 rime duration

Parameter	Estimate	Std. Error	df	t	Sig.	95% Confidence Interval	
						Lower Bound	Upper Bound
Intercept	.056511	.007863	32.298	7.187	.000	.040502	.072521
[pair=Ce]	.025736	.006740	43.748	3.818	.000	.012150	.039322
[pair=en]	.090802	.007915	41.981	11.473	.000	.074829	.106774
[pair=i]	.039883	.011836	36.967	3.370	.002	.015900	.063867
[pair=shen]	.089551	.011836	36.967	7.566	.000	.065567	.113534
[pair=u]	.035204	.011836	36.967	2.974	.005	.011220	.059187
[pair=zhe]	.023497	.007915	41.981	2.969	.005	.007525	.039470
[pair=zi]	0 ^b	0
[tone=L] * [pair=Ce]	.029344	.004833	32	6.071	.000	.019499	.039189
[tone=L] * [pair=en]	-.002816	.007762	32	-.363	.719	-.018626	.012994
[tone=L] * [pair=i]	.005769	.014372	32	.401	.691	-.023506	.035044
[tone=L] * [pair=shen]	.004819	.014372	32	.335	.740	-.024456	.034094
[tone=L] * [pair=u]	-.004434	.014372	32	-.309	.760	-.033709	.024841
[tone=L] * [pair=zhe]	.031737	.007762	32	4.089	.000	.015927	.047548
[tone=L] * [pair=zi]	.021532	.007762	32	2.774	.009	.005722	.037343
[tone=N] * [pair=Ce]	0 ^b	0
[tone=N] * [pair=en]	0 ^b	0
[tone=N] * [pair=i]	0 ^b	0
[tone=N] * [pair=shen]	0 ^b	0
[tone=N] * [pair=u]	0 ^b	0
[tone=N] * [pair=zhe]	0 ^b	0
[tone=N] * [pair=zi]	0 ^b	0