

Phonetic observations on stress and tones in Mandarin Chinese

Corinne B. Moore

1. Introduction

Phonetic research on stress has shown that fundamental frequency and duration are effective acoustic correlates of stress in many languages (e.g., Fry, 1955; see Lehiste, 1970; and Cooper and Sorensen, 1981). Thus, we expect stressed syllables to have higher F0 and longer durations than unstressed syllables, other factors being equal. In addition to these effects, stress has been shown to affect formant values in vowels. For example, formants in unstressed vowels may centralize, becoming more schwa-like (Lindblom, 1963). Therefore, stress, or rather lack of stress, may sometimes alter the acoustic characteristics of phonemes to the point of neutralization. While substantial data exist documenting the effects of stress on segments and syllables, there has been less investigation of stress in lexical tones. In particular, there is little published data examining the effectiveness of F0 and duration as correlates of stress for tones. Moreover, it is not clear how stress-altered F0 patterns may affect tones, which critically depend upon F0 for phonemic contrasts. For instance, it is conceivable that changes in F0 due to lack of stress may neutralize the distinctive contour of a tone, similar to the case of centralized formants in unstressed vowels.

To examine the effects of stress on tones, the current study reports data on stress in Mandarin Chinese. In Mandarin, four tones serve to distinguish all lexical items except a small number of atonic syllables. Syllables with tones may have primary, secondary and lower degrees of stress, while atonic syllables are always unstressed. This study presents evidence in two areas. First, the investigation provides data on the acoustic correlates of stress in Mandarin tones. Second, the study considers whether F0 contours of unstressed tones reduce or neutralize completely, resulting in tonelessness. References to unstressed syllables in the literature typically assume that their tones are lost (Yip, 1980; Tseng, 1981; Coster and Kratochvil, 1984; Shen, 1988; Duanmu, 1990; Davison, 1991), though specific acoustic data are needed to determine whether these syllables neutralize completely.

Previous experimental work on Mandarin suggests that F0 and duration are influenced by stress, though these studies have not specifically reported on unstressed tones independent of intonational effects. Nor has there been a study which compares F0 patterns on unstressed tones with those on atonic syllables. Coster and Kratochvil (1984), in an acoustic study of stressed and unstressed Mandarin tones in running speech, found

that F0, duration and amplitude are effective correlates of stress. With regard to unstressed syllables, however, the authors state that "the contour properties of individual syllables become neutralized in stress groups" (p. 122), where a stress group may include both atonic syllables and unstressed syllables with lexical tones. Therefore, F0 patterns in atonic syllables and in unstressed syllables with tones are treated as equivalent, concealing any evidence which could be used to determine the extent of neutralization on unstressed tones. Moreover, attributing acoustic variations to stress is impossible in spontaneous speech, because of the confounding influence of intonational F0 contours and syntactic context. In another study on Mandarin using elicited sentences of different intonation types, Shen (1988) observes that unstressed syllables have shorter durations than stressed syllables, though little data are given. Another acoustic study by Tseng (1981) examines tones excised from free conversation. In this study, Tseng consolidates two types of unstressed syllables: one type consists of atonic syllables, the other type includes tonic syllables whose F0 pattern "does not necessarily correspond completely to that morpheme's lexical tone...and thus [becomes] a phonetically neutral tone" (p. 94, brackets mine). This consolidation of both atonic and tonic syllables suggests that unstressed tones are neutralized, though no detailed acoustic data on the F0 patterns of these unstressed syllables are supplied. Tseng does provide a comparison of average durations of 100 stressed tones with 100 unstressed tones, showing that the average duration for the stressed tones is longer than the average duration for the unstressed tones. Tseng's use of spontaneous speech, however, inhibits conclusions we may draw about stress effects. Another study with limited phonetic data on Mandarin tones is Shih (1988), who analyzes disyllabic words in different intonational contours. She observes that F0 values are affected by prominence, where prominence refers to intonational focus. In particular, Shih notes that "prominence is reflected by expanding pitch range" (p. 93) and longer duration. Shih's study, in addition to the ones mentioned above, observes that F0 and duration are correlates of stress in Mandarin tones, though none of the studies determines whether neutralization occurs in weakly stressed tones.

Experiments to date have not explored whether a metrically weak tone retains its canonical F0 contour or takes on the neutral tone pattern. Furthermore, the studies discussed earlier examine tones without controlling for the effects of intonation. The current study investigates these issues by implementing two strategies. First of all, since the category, "unstressed" has encompassed both tonic and atonic syllables, this study examines each category independently to determine whether tonal neutralization occurs in tonic syllables, yielding the same F0 contours as atonic syllables. Second, the experiment

uses non-spontaneous speech in order to reveal the effects of stress apart from the influence of intonation. Three acoustic parameters are investigated: duration of stressed versus unstressed stimuli, F0 range of stressed versus unstressed tones, and F0 contours of unstressed tones. The first two parameters, duration and F0 range, are reported to be stress correlates for tones, while the third area, F0 contour, forms a basis for comparison between unstressed tones. Results and discussion of these data are provided after an introduction to Mandarin tones in the following section.

1.1 Mandarin Tones

Mandarin Chinese has four underlying contrastive lexical tones. For example, the minimal pairs below show how different tones on the syllable *ma* change the meaning of the word.

(1)

Tone 1: *ma* 'mother' Tone 2: *ma* 'hemp' Tone 3: *ma* 'horse' Tone 4: *ma* 'scold'

The fundamental frequency contours of each tone are distinct from each other, and are represented descriptively as follows:

(2)

Tone 1: high level (55)

Tone 2: mid-high rising (35)

Tone 3: low-falling rising (213)

Tone 4: high-falling (51)

The notation above provides references to the numerical system which makes use of a five point scale where 1 is low and 5 is high. These are the tone shapes which surface when pronounced in isolation. When stressed, all tones appear with the above contours except Tone 3, which has a low-falling variant in non-final positions. The descriptions in (2) represent canonical F0 patterns for the tones. In order to show how these tones appear in actual speech, F0 contours of the four stressed tones, spoken by one of the subjects in this study, have been plotted in Figure 1 below.

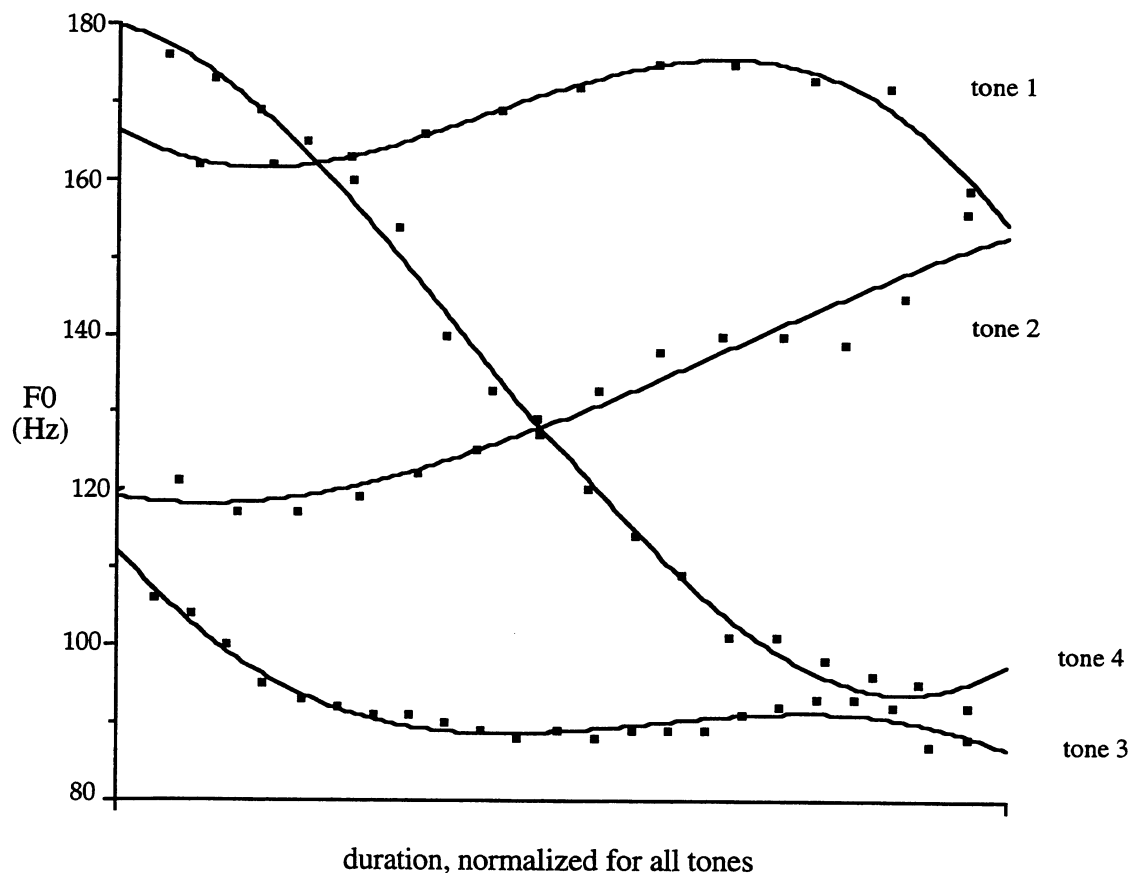


Figure 1: stressed (citation) tones for one speaker

The data points in Figure 1 represent average F0 values for eight stimuli, measured in steps of 10 ms. A polynomial was fitted to the data points to generate the contours for each tone. These F0 contours illustrate how the tones correspond to the descriptions in (2): Tone 1 remains high, while Tone 4 begins high but falls to a low point; Tone 2 begins lower than Tones 1 and 4, but achieves a high F0 by the end of the tone; Tone 3 occupies the lowest area of the range. Note that Tone 3 exhibits the low-falling variant, since the tone was produced in a carrier phrase.

1.2 Atonic Syllables: The Neutral Tone (Tone 0)

In addition to the lexical tones there is another set of F0 contours which correspond to atonic syllables. These F0 patterns refer to a broad range of tone patterns which do not

correspond to the other four lexical tones. The canonical F0 pattern occurs on atonic and stressless syllables, a small group of lexical items which serve as grammatical markers or suffixes, and are never stressed (Chao, 1968). These F0 contours on atonic syllables vary predictably according to the preceding tone, though the contours do not serve to distinguish lexical items. Rather, these well-documented F0 contours are captured by the terms the "neutral tone", or Tone 0. By definition, then, the presence of the neutral tone indicates tonelessness and lack of stress.

The fundamental frequency patterns associated with the neutral tone have been documented by several previous acoustic studies (Tseng, 1981; Gao, 1980; Cheng, 1973; Kratochvil, 1968; Zadoenko, 1958; Qi, 1956; Chao, 1933, 1948, 1956, 1968). When produced in isolation, the neutral tone is low and flat (Tseng, 1981), though in regular speech "it does not have a distinct fundamental frequency pattern of its own" (p. 54). Chao (1968, 1956, 1948) proposes that the pitch of the neutral tone depends upon the tone of the preceding syllable. The diagrams below, as first presented by Chao (1948:27), show how the pitch¹ of the neutral tone changes depending on which tone precedes it.

(3)

after Tone 1 .l
 after Tone 2 .l
 after Tone 3 ·l
 after Tone 4 .l

The schema above developed by Chao shows that after Tone 1, the pitch of the neutral tone is fairly low, compared to the pitch after Tone 2 which is somewhat higher; after Tone 3 the neutral tone is at about the mid-point of the range; and after Tone 4, the pitch of the neutral tone is typically very low. Additional acoustic data for the neutral tone will be presented in a later section, and these data will be compared to results reported by previous studies.

For now, we may consider the neutral tone to correspond to toneless and stressless syllables, in contrast with the lexical tones, which appear on stressed syllables. Therefore, it is assumed that an unstressed tone which has become neutralized will exhibit an F0 contour associated with the neutral tone, not one of the lexical tones. This assumption allows us to compare F0 contours on metrically weak syllables, which tend to lose their tones, with the F0 on unstressed atonic syllables. If an unstressed tone is neutralized, the resulting F0 contour should resemble the neutral tone. If, on the other hand, an unstressed

¹The term "pitch" is used here because Chao's description is an impressionistic account, not based upon F0 measurements.

tone is reduced, rather than completely neutralized, the acoustic pattern should resemble the F0 pattern of the lexical tone, though reduction may alter the tone to some degree. The current phonetic experiment was designed to test this hypothesis, as well as to study acoustic correlates of stress for Mandarin tones.

2. Phonetic Experiment

2.1 Methods

2.1.1 Subjects

The subjects were two male native speakers of Beijing Mandarin, both approximately aged 35, without known speech or hearing disorders. The subjects were both born and raised in Beijing, and had been in the United States for two years.

2.1.2 Stimuli

The subjects each produced three repetitions from a list of prepared sentences, written in Chinese characters. Each sentence in the list included a target syllable, embedded in a carrier phrase. The target syllables carried one of the four Mandarin tones or Tone 0, described earlier.

In choosing the target syllables, consideration was given to both tone and word function. As stated above, the target words included examples from each of the four tones plus Tone 0. In addition, syllables which serve a grammatical function were chosen, since their tones tend to neutralize (Li and Thompson, 1981). Finally, the position of the syllable within the phrase was considered, since phrase groups in Mandarin have predictable stress placement; that is, the position of the syllable in the phrase will determine whether the syllable will receive primary, secondary or tertiary stress. Tseng (1981), Shen (1988) and Davison (1991) state that syllables in the middle of a trisyllabic phrase are most vulnerable to destressing. Based upon these criteria, the syllables which appear best suited to serve as stimuli are the classifiers.

Classifiers are morphemes which are obligatorily placed between a number and a noun. According to Norman (1988), classifiers do indeed have underlying tones. However, Shih (1986) treats classifiers as clitics, a type of syllable which is predicted to lose its underlying tone (Yip, 1980; Tseng, 1981; Shen 1988). The classifier attaches phonologically to the preceding syllable in a cliticization process which produces the neutral tone (Yip, 1980).

Therefore, the status of classifiers as clitics would predict that their tones neutralize and exhibit the neutral tone pattern.

Different nouns use different classifiers, according to the meaning of the noun, though there is a "default" measure word, "ge" [kɤ]. Listed below are the four classifiers used as target syllables, and examples of how they are used in Mandarin. (the notation T = Tone)

(4)

Classifier		Example
		T1 T1 T1
zhi [dʒi]	Tone 1	yi zhi mao 'one cat'
		T1 T2 T1
he [hɤ]	Tone 2	yi he yan 'one (box of) cigarettes'
		T1 T3 T2
wan [ʋan]	Tone 3	yi wan you 'one (bowl of) oil'
		T1 T4 T4
ge [kɤ]	Tone 4	yi ge cai 'one dish'

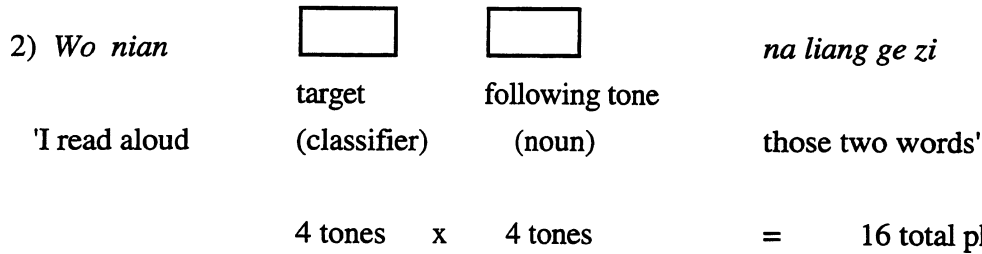
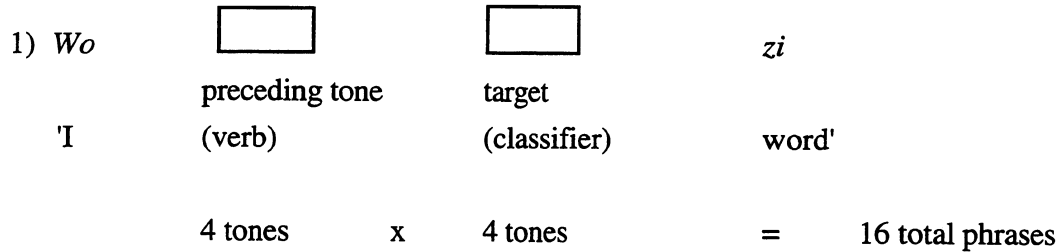
The four classifiers above have a grammatical function and also appear in the middle of a trisyllabic phrase. These two characteristics, therefore, would be expected to produce the neutralization of underlying tones in unstressed contexts.

While classifiers are suitable representatives of underlyingly specified tones, the lexically toneless syllable, *de* [tɤ], was chosen to represent atonic syllables. This syllable is the possessive particle in Mandarin. It never appears stressed, and its tone pattern is dictated by the preceding tone (see the discussion of the neutral tone in the preceding section). The particle is placed between the noun and its possessor, forming a trisyllabic phrase, as in *ta de mao* 'his cat'.

The stimuli for Tones 1–4 and Tone 0 were produced in carrier phrases designed with two goals in mind. First, the carrier phrases needed to induce both stressed and unstressed target syllables, in order to assess the impact of stress on tones. Second, the carrier phrases needed to place each target syllable in varying tonal environments, to see whether the preceding tone affects the lexical tone in the same way that it does a neutral tone.

Carrier phrases for the stressed environments were designed to induce primary stress. Therefore, the trisyllabic phrase format could not be used. Instead, targets were placed in a carrier phrase which varied either the preceding or the following tone. Schematized diagrams of the stressed carrier phrases are as follows:

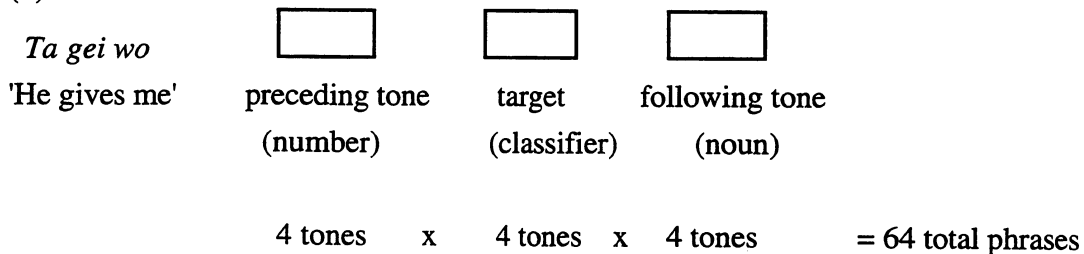
(5)



Note that phrase 1) above varied the tone of the preceding syllable, while phrase 2) varied the tone of the following syllable.

For unstressed contexts, target words were placed in the middle of a trisyllabic phrase, since it is commonly assumed that this configuration induces the neutral tone (Tseng, 1981; Shen, 1988; Davison, 1991), while the first and last syllables in such a phrase receive normal stress. Thus, the following carrier sentence was used to measure unstressed syllables:

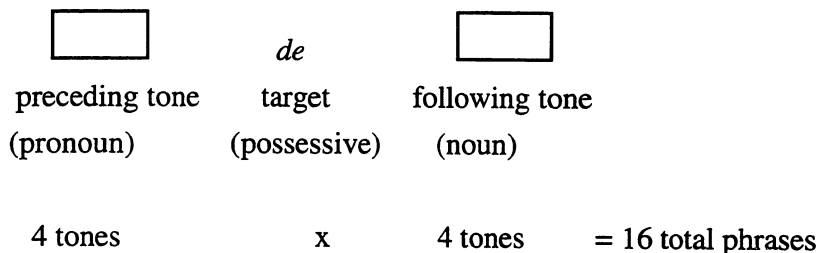
(6)



The numbers, which are the syllables preceding the test word, vary in tone. The following noun also varied in tone, and also had to be appropriate for the measure word.

Finally, the carrier sentence for the possessive *de* is shown below, again allowing for both preceding and following tones to vary:

(7)



The carrier sentences above allowed the target syllable to vary in tone as well as the surrounding syllables, and also differed in whether the targets were stressed or unstressed. Segmental composition across carrier phrases was not completely controlled, since classifiers occur with a restricted set of lexical items. However, within the target syllables, Tones 2, 4 and 0 contain the same vowel, [ɤ]. Tone 3 has a low-back vowel, [a], and Tone 1 contains a high central retroflex unrounded vowel (Howie, 1976). All initial consonants are voiceless, while the Tone 3 target has an initial on-glide.²

F0 contour data presented in this paper will focus on the effects from the preceding tone, as that factor has been principally used in the literature to describe the neutral tone. The fundamental frequency of a neutralized lexical tone may be influenced by the preceding tone as well. Thus, data including information about the preceding tone supplies another comparison criterion, in addition to the F0 contour on the stimulus, by which to verify

²Segmental context varied, but results show that canonical tone patterns were maintained. Howie (1976) also found that tone patterns were not altered by segmental context to the extent that the F0 contours were confused with other lexical tones. Differences in F0 values which might occur would result from differences in consonant voicing and vowel height. With regards to differences in consonant voicing, the Tone 1, 2 and 4 stimuli all have a voiceless initial consonant, and no final consonant. In terms of vowel height, the vowels in these stimuli are not substantially different in height: the Tone 1 vowel is high central, while the Tones 2 and 4 stimuli have a high-mid back vowel. The major differences are with the Tone 3 syllable, which has a low back vowel, an initial glide and a nasal final consonant. The Tone 3 initial glide was treated as a vowel here. In Howie's data, stimuli with an initial glide and a final nasal had onsets and offsets nearly identical to vowel-initial data. The low vowel would be expected to exhibit lower F0 values than the vowels of the other stimuli (Bo and Zhang, 1986; Lehiste, 1970), though the average Tone 3 onset results will show that the lowest F0 point is still below what differences in intrinsic pitch would predict. In any case, comparisons between stressed and unstressed tokens, a major focus of this study, did not vary in segmental context. Thus, there is little evidence that the varying segmental context affects other comparisons between tones in this study, though the inconsistent contexts limit the kind of cross-tonal comparisons that can be made.

whether the neutral tone and neutralized lexical tones yield the same fundamental frequency patterns.

The four tones of the target syllables and the four tones on the preceding and following syllables resulted in 64 (4x4x4) sentences for unstressed lexical tones. Target neutral tone syllables were also tested with the preceding and surrounding tones varied, adding 4x4=16 sentences. Note that carrier sentences for the stressed syllables were different from the unstressed syllables. This is because stressed syllables cannot be in the middle of a trisyllabic phrase, but do occur in disyllabic phrases. Consequently, only tones within the same phrase of the target syllables were varied. That is, for each target syllable, only eight instances for each tone occurred, since the sentence varied either the preceding or the following tone (target tone + preceding four tones = 4, plus target tone + following four tones = 4). Stressed contexts, then, comprised 32 sentences altogether (4 target tones x 4 following tones = 16, plus 4 target tones x 4 preceding tones = 16). Each speaker thus produced 112 sentences for each repetition. The entire corpus included 448 sentences (112 sentences x 2 repetitions x 2 speakers).

2.1.3 Recording

The two speakers were recorded in a sound-proof booth, using a cardioid microphone (Electrovoice, model RE 20) and a high-quality cassette recorder (Marantz, model PMD 222).

The sentences were pseudo-randomized so that only the syllable preceding the test word remained the same in any block. Each speaker read the sentences three times at a normal speaking rate. The last two of the repetitions were chosen for analysis.

2.1.4 Analysis

Recordings were digitized on a SUN 3/160 computer. The stimuli were sampled at 10 kHz with a low-pass filter setting of 5 kHz. The data were stored as files to be processed by the software package WAVES+.

Durations of the target words were measured from the waveform. Because of the nature of running speech, many words showed no distinct boundary from one word to the next. In these cases, visual and audio information provided cues to word boundaries. In addition, a wide-band spectrogram displayed simultaneously with the waveform further refined the task of delineating word boundaries. Duration was taken both including and excluding the initial consonant. Measurements include the burst of the initial consonant, if visible and/or audible. The onset of F1 was taken to be the beginning of the vowel, while

the offset of F2 was considered to be the end. In the case of the Tone 3 syllable, [ʊan], three measurements were taken: the entire syllable, the vowel plus the final consonant, and just the nuclear vowel. For this syllable, the point of the final nasal offset was determined by examination of the waveform and spectrogram.

Fundamental frequency patterns were measured for target syllables in each sentence by using the WAVES+ pitch-tracking program. Pitch tracks were generated using a step size of 5 ms. Points of measurement include beginning, ending, peak and valley F0 values for the target syllables. A measurement roughly corresponding to the midpoint of the voiced part of the syllable was also taken. Beginning (initial) F0 values were taken starting from the vowel for Tones 1, 2, and 4, since they begin with voiceless consonants, and starting from the onset glide in the Tone 3 stimuli. Ending values were taken at vowel offset in the Tone 1, 2 and 4 stimuli, where vowel offset was determined in the waveform to be the point where the vowel no longer showed higher formant patterns. In ambiguous cases, spectrograms were also used to examine the point of vowel offset, determined to coincide with F2 offset. Tonal offset in the Tone 3 stimuli was determined to occur at the end of the final nasal. Peak and valley F0 values were taken at the highest and lowest F0 points of the stimulus, respectively. Measurements of F0 peak and valley, beginning and ending values are represented in Figure 2 below:

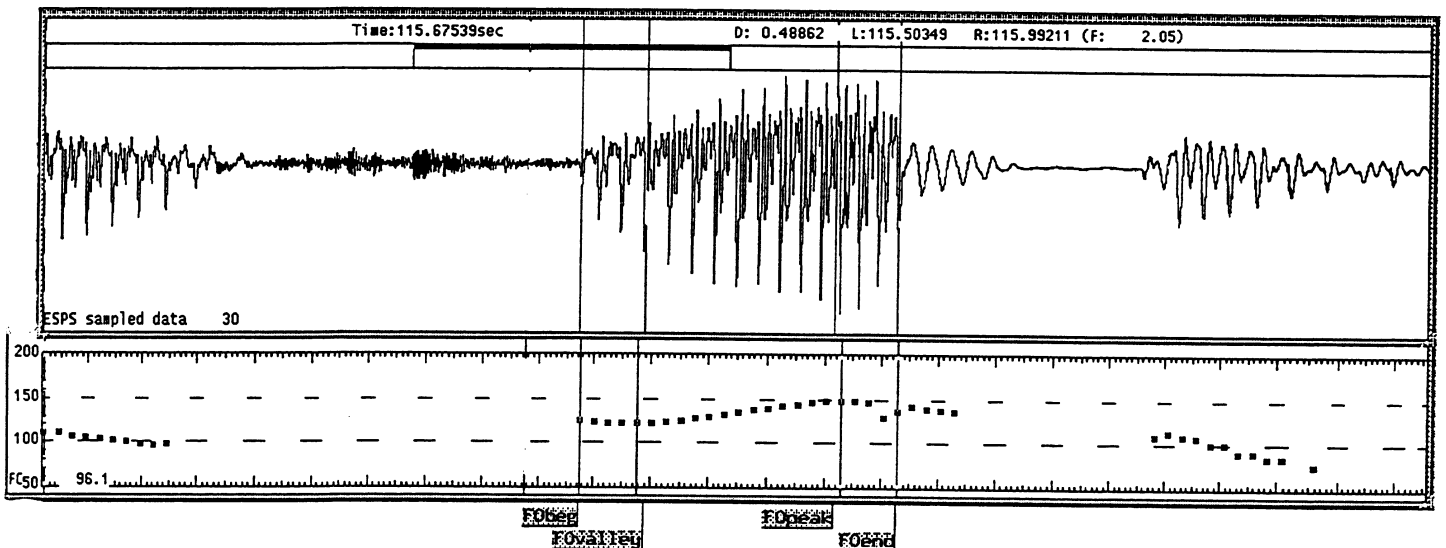


Figure 2: Waveform and pitch track for the Tone 2 stimulus *he*. Points on the pitch track represent steps of 5 ms. Measurement locations for F0 peak, valley, beginning (F0beg) and ending (F0end) points are indicated by the labeled cursors.

The waveform and pitch track in Figure 2 show the Tone 2 classifier, *he*, with peak, valley, beginning and ending F0 values marked by the cursors, labeled accordingly. In some cases, data points on the pitch tracks showed a sharp fall or rise at the onset of a word which were considered to be artifacts and so not included. Measuring one period of the vowel in the corresponding part of the waveform served as a verification of the F0 values for these ambiguous points.

3. Results

The presentation of results which follow covers two major areas: duration and fundamental frequency. With respect to duration, this study discusses comparisons between stressed and unstressed values, and among the tones themselves. Duration measurements for both the entire syllable and only the tone (as realized on the vowel, or voiced portion of the syllable) are included. This reflects the fact that although tone will only be realized phonetically on the voiced portion of the syllable, its prosodic domain is properly the entire syllable. In fact, duration results indicate that in addition to the rime, stress affects duration of the initial consonant as well. Comparisons between tones will be limited, however, because of the differences in segmental content across tones. Finally, a comparison is made between the neutral tone and unstressed tone durations.

The results for fundamental frequency focus on two factors: 1) F0 range, since that has been observed to be a stress correlate in Chinese tones, and 2) F0 contours in various tonal contexts. Examining F0 contours allows us to make comparisons between the neutral tone and the other target tones.

All values are averaged over the two speakers, since both subjects show the same pattern in results for the criteria examined here. The discussion is divided into duration, F0 range, and F0 contour sections.

3.1 Results of Duration Measurements for Stressed and Unstressed Stimuli

Tables 1 and 2 below give measurements for duration of stressed and unstressed stimuli. Table 1 reports duration data for syllables. Table 2 reports durations of the tones only, measured from the onset of the vowel to its offset, as detailed in section 2.1.4.

Table 1: Duration (in ms) of target syllables in stressed vs. unstressed environments. An asterisk indicates statistical significance at the .05 level.

	<u>stressed</u>	<u>unstressed</u>	<u>Difference</u>
Tone 1 [dzi]	177	144	33*
Tone 2 [hy]	286	213	73*
Tone 3 [ʊan]	228	205	23*
Tone 4 [ky]	194	93	101*
Tone 0 [ty]	N/A	81	-.**

**Tone 0 is never stressed

Table 2: Duration (in ms) of tones in stressed vs. unstressed environments. An asterisk indicates statistical significance at the .05 level.

	<u>stressed</u>	<u>unstressed</u>	<u>Difference</u>
Tone 1	108	86	22*
Tone 2	148	111	37*
Tone 3[an]/[a]	225/129	195/125	30*/4(ns)
Tone 4	161	73	88*
Tone 0	N/A	66	-

3.2 Discussion

Table 1 gives duration measurements (in ms) for stressed and unstressed syllables, including the atonic syllable. Table 2 reports duration measurements for tones only, not including the voiceless initial consonants. Both tables show clearly that duration is affected by stress in Mandarin. Table 1 indicates that there is an average 58 ms duration difference between stressed and unstressed syllables. Tone 1, for example, reduces in duration by 33 ms when unstressed. Likewise, Tone 4 has a duration of 194ms in stressed contexts, but drops to 93ms when unstressed. Paired, two-tailed t-test results show that the differences in duration between stressed and unstressed syllables (based upon the figures in Table 1) are all statistically significant: for Tone 1, $[t(94) = -5.18, p < .001]$; for Tone 2, $[t(94) = -8.88, p < .001]$; for Tone 3, $[t(93) = -2.29, p < .03]$; for Tone 4, $[t(93) = -17.85, p < .001]$. The measurements in Table 2 show similar results in that all tones show statistically different durations in stressed versus unstressed stimuli: for Tone 1, $[t(94) = 4.89, p < .001]$; for Tone 2, $[t(94) = 7.70, p < .001]$; for Tone 3 [an] durations, $[t(93) = 4.62, p < .001]$; and for Tone 4, $[t(94) = 15.61, p < .001]$. The Tone 3 comparison of the vowel only durations was not significant. These results corroborate the previous studies by Shen

(1988), Coster and Kratochvil (1984), and Tseng (1981), who also showed that unstressed tones are shorter in duration than stressed syllables.

In addition to the comparisons within instances of the same tone, the data also indicate duration differences between tones. Inherent duration differences between the four stressed lexical tones have been studied before (Dreher and Lee, 1966; Kratochvil, 1971; Ting, 1971; Howie, 1976; Nordenhake and Svantesson, 1983).³ The results of these studies do not agree completely, though the first four all report Tone 3 as being the longest, and Tone 4 as being the shortest (sentence-medially in the Nordenhake and Svantesson study). All except Kratochvil show Tones 2 and 3 as the longest, and Tones 1 and 4 as the shortest. The data in Tables 1 and 2 show the same duration relationships between the tones as reported by these other studies. Comparing figures in Table 2, unstressed Tone 3 has the greatest duration, whether comparing [an] with 195ms or [a] with 125ms. Although Tone 4 is greater in duration when stressed, it loses the most duration when unstressed. Unstressed tones, then, generally show the same relative durations as reported in previous duration studies. Tseng's (1981) conclusions from spontaneous speech were that inherent durations are lost (p. 170). While the current study is not designed to test inherent durations, the data here suggest that relative durations are maintained in unstressed environments.

The other interesting duration comparison to be made is between the atonic syllable, Tone 0, and the underlyingly specified tones. Tables 1 and 2 show that the neutral tone is shorter in duration than Tones 1-4 in unstressed contexts. Comparing tone duration values from Table 2, these differences are all significant, even between Tone 0 and Tone 4: t-test results for Tone 1 versus Tone 0, $[t(63) = 6.65, p < .001]$; for Tone 2 and Tone 0, $[t(63) = 16.25, p < .001]$; for Tone 3 [an] versus Tone 0, $[t(125) = 32.62, p < .001]$; and for Tone 4 and Tone 0, $[t(63) = 2.21, p < .03]$. Thus, Tones 1-4 are distinct from the atonic syllable with respect to duration.

A final observation concerns stress effects in initial consonants. The data show that durations of initial consonants in the stressed stimuli are shorter than in unstressed stimuli. This effect is evident in data for Tones 1, 2 and 4, which have initial consonants. For example, the duration of the stressed Tone 1 initial consonant [dz] is 69 ms, but 58ms when unstressed. In stressed Tone 2, the initial consonant [h] is 138ms, compared to 102ms when unstressed. Similarly, the initial [k] of the Tone 4 stimuli are 33ms when

³Durations were measured over the vowel for all studies listed here.

stressed, but 20ms in unstressed tokens. This evidence supports the theory that stress effects extend beyond the rime.

In summary, there are three important conclusions based upon the duration data. First, duration is an effective correlate of stress. This is shown by the fact that unstressed tones and syllables are significantly shorter than stressed tones and syllables. Second, relative duration differences are generally maintained in the unstressed tones, not neutralized to the duration of the atonic syllable. Third, Tones 1-4 are distinct from the atonic syllable with respect to duration. These findings demonstrate that unstressed tones are not equivalent in phonetic surface realization. Differences appear to be due to the presence or absence of underlying tones in the target words. Syllables with Tones 1-4 have an underlying tone, and thus maintain a unique duration component, whereas syllables with no underlying specification do not. Therefore, phonetic differences between these two categories of tones may occur.

3.3 Results of Fundamental Frequency Range Changes in Stressed and Unstressed Tones

Section 1 notes that in addition to duration, stress is also considered to be reflected in F0 range changes (Tseng 1981; Coster and Kratochvil 1984; Shih, 1988). Specifically, it is expected that stressed syllables extend their F0 range by making their highs higher and their lows lower (Chao, 1968; Dow, 1972; Shen, 1988; Shih, 1988). Fundamental frequency range results for stressed and unstressed tones from the current study are presented in Table 3 below.

Table 3: F0 range (in Hz) in stressed and unstressed contexts, measured according to F0 Peak, Valley and Range (Peak minus Valley) for each tone.

Values given for Peak_{stressed} (P_S), Peak_{unstressed} (P_U), Valley_{stressed} (V_S), Valley_{unstressed} (V_U) and Range_{stressed} (R_S), Range_{unstressed} (R_U).

An asterisk indicates statistical significance at the .05 level (see text).

<u>Tone</u>	<u>P_S</u>	<u>P_U</u>	<u>P_S-P_U</u>	<u>V_S</u>	<u>V_U</u>	<u>V_S-V_U</u>	<u>R_S</u>	<u>R_U</u>	<u>R_S-R_U</u>
Tone 1	136	121	15*	118	110	8	17	11	6*
Tone 2	113	104	9*	96	95	1	17	9	8*
Tone 3	106	112	-6	80	82	-2	25	29	-4
Tone 4	138	109	29*	83	88	-5	55	21	34*
	<u>Peak</u>	<u>Valley</u>	<u>Range</u>						
Tone 0	111	94	17						

3.4 Discussion of F0 Range Results

The fundamental frequency range data in Table 3 report F0 values for stressed and unstressed tones, represented by Peak and Valley measurements for each tone. Range values were calculated by taking the difference between Valley and Peak values. An asterisk indicates statistical significance for the difference between stressed and unstressed values. The table shows how the range of tones in stressed contexts is larger than in unstressed contexts. The change in range is significant for all tones except Tone 3. Note particularly the results for Tone 4, which show a decline in range of 34 Hz. Rather than the three parameters Peak, Valley and Range all showing significant differences between stressed and unstressed values, however, unpaired t-test results indicate that only Peak and Range differences are significant: Tone 1 Peak [$t(94) = -2.18, p < .04$], Tone 1 Range [$t(94) = -3.56, p < .001$]; Tone 2 Peak [$t(94) = -1.98, p < .05$], Tone 2 Range [$t(94) = -4.36, p < .001$]; Tone 4 Peak [$t(92) = -4.89, p < .001$], Tone 4 Range [$t(92) = -8.15, p < .001$]. The fact that only Peak and Range values differ significantly means that changes in the F0 range due to stress effects occur as a result of changes in the Peak value alone. Changes in the Valley are not significant, and thus do not contribute to the expansion or narrowing of the range. This result contradicts the commonly held view that F0 range expands at both ends. However, some work on intonation has shown the same results; that is, that Lows do not appear to expand to the extent that F0 Highs do (Liberman and Pierrehumbert, 1984). In fact, Tone 2, which is phonologically Low-High (upper register) but mid-rising phonetically, also shows non-significant expansion at the low F0 end. Thus, it may not be just the phonetic Low which does not contribute to range expansion, but a phonological Low.

The F0 range of Tone 3 for these data does not change significantly. The reasons for this are still unclear and require further study. Ting (1971), who analyzed tone changes in various sentence positions, also notes that the range of Tone 3 generally shows very little change. However, since Tone 3 is phonologically a Low tone, this result may also be explained by the hypothesis that Lows do not shift to the same degree as Highs.

In summary, since Tones 1, 2 and 4 show reduced F0 range in unstressed contexts, these results support the claim that F0 range is wider in stressed contexts, and narrower when unstressed. The data further show that these Range differences are due to changes in the Peak. The results here agree with other studies which show that F0 range is an effective correlate of stress.

Fundamental frequency range can only be used in comparing stressed vs. unstressed instances of the same tone, however. In order to discuss the question of whether

unstressed tones "lose" their inherent F0 contours, more data are needed to show the F0 patterns of tones. These data are discussed in the next section.

3.5 Fundamental frequency patterns for target syllables in various tonal contexts

Data reported in preceding sections have shown that unstressed lexical tones are distinct from stressless atonic syllables in terms of duration, and also that lack of stress contracts F0 range due to Peak lowering. Now we examine fundamental frequency patterns of the stimuli, which crucially determine whether unstressed tones "lose" their characteristic contours, becoming neutralized. If the unstressed lexical tones retain their canonical F0 contours, we may conclude that tonal neutralization has not occurred. If, on the other hand, the F0 patterns on the unstressed syllables are neutralized, they may be compared to the F0 patterns on atonic syllables, which are collectively referred to as the neutral tone. In addition, unstressed tones may be compared to the neutral tone with respect to the degree of influence exerted by a preceding tone. Since the F0 pattern of the neutral tone is considered to be dependent upon the preceding tone (see section 1.2 on the neutral tone), that same influence may be evident in F0 contours on unstressed tones.

Figures 3-7 below show F0 patterns for each unstressed tone, from the offset of the preceding syllable to the end of the target word. The points plotted along the horizontal axis include the pretone F0 Peak or Valley, the beginning, middle and endpoint of the tone, measured as described in the Methods section. The vertical axis shows the F0 level in hertz. Note that the contours are analyzed as a function of the preceding tone, not the following tone. Thus, the ending points represent average F0 values when the tone is followed by all four lexical tones. The pretone, however, only represents the one tone coded as per the legend. The pre-tone endpoint corresponds to the F0 peak or valley of that tone according to the citation form.

3.6 F0 contour results

The following results focus primarily on two characteristics of the data: beginning and ending F0 values, and F0 contour. Discussions of the data follow each figure.

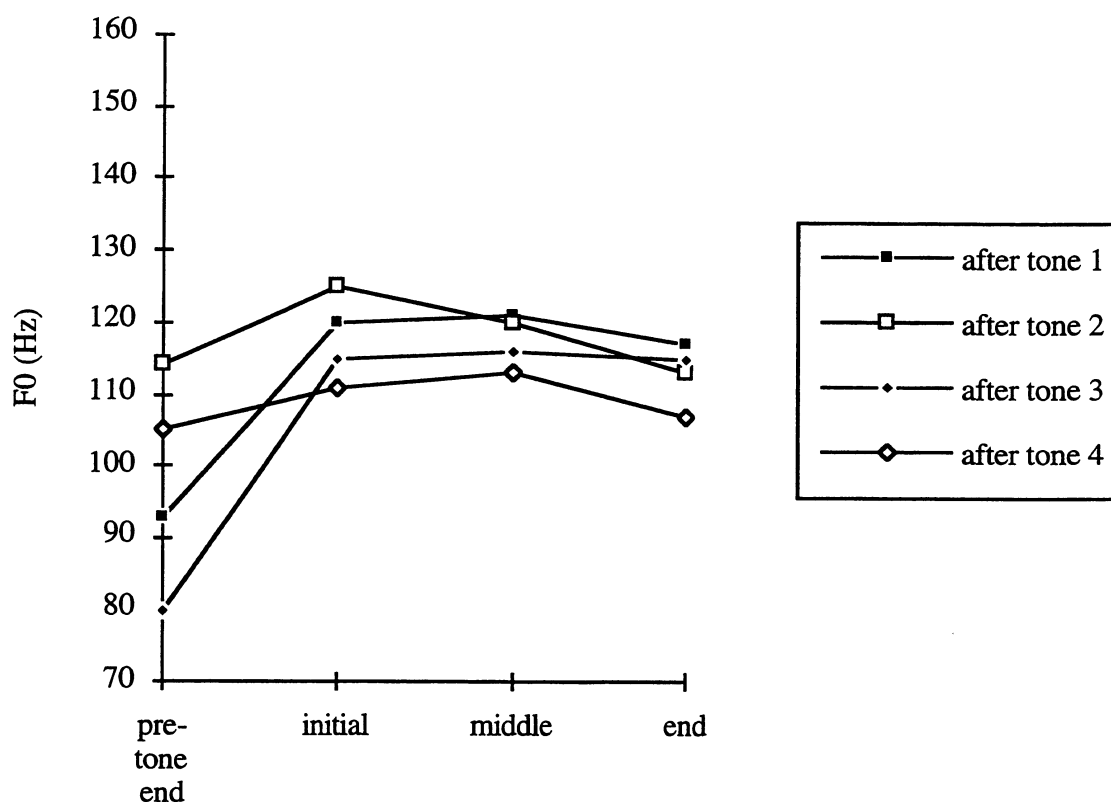


Figure 3: fundamental frequency of unstressed Tone 1 as a function of the preceding tone

The F0 data in Figure 3 show a fairly level contour for Tone 1, with fundamental frequency values clustering between 110-125 Hz. Regardless of the preceding tonal context, the F0 range covered by all four contours is quite restricted overall. Moreover, the target initial point appears to be independent of the pre-tone end. Particularly after Tone 3, which is low, the onset of the tone does not move gradually toward the F0 range, but rather achieves the target range directly at the onset; the target syllable jumps 35 Hz, then remains fairly level. Differences between almost all initial points are statistically significant. This is largely due to coarticulation effects from the preceding tones. For example, after Tones 1 and 2, which end high, the target is level and high, relative to the other instances of the target. After Tone 4, which usually ends low, the target tone is at its lowest F0 level. These results agree with Shen (1990) who reports that tones are higher after tones with high offsets, and lower after tones with lower offsets. Ending

points generally do not show significant differences. Thus, the unstressed Tone 1 appears to maintain its general F0 pattern which shows the canonical level Tone 1 shape. In addition, initial and ending F0 points are concentrated within a limited F0 range. Also, coarticulation effects are evident in the initial F0 points. Finally, ending points are generally focussed, rather than diffuse.

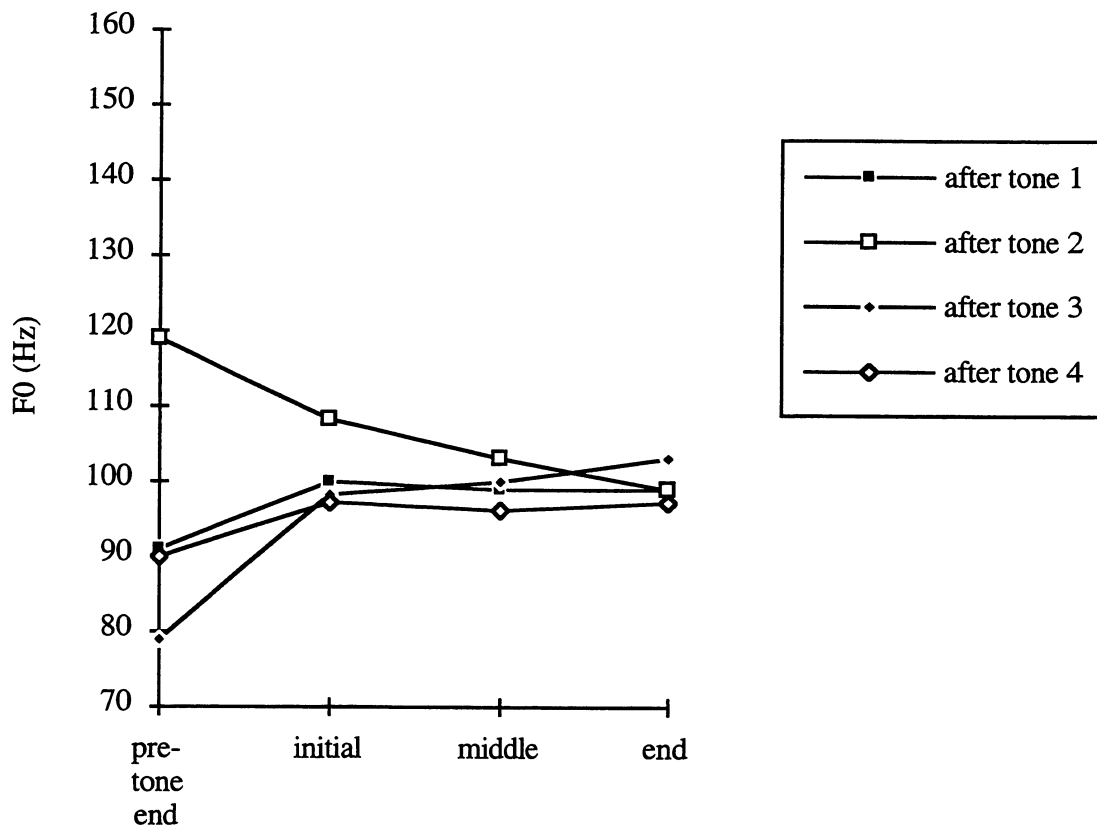


Figure 4: fundamental frequency of unstressed Tone 2 as a function of the preceding tone

Figure 4 shows that the F0 onset region for the Tone 2 stimuli is lower than the one for Tone 1, clustering between 95-110 Hz. The F0 Range data show an average overall rise of 9 Hz between Valley and Peak.⁴ The contours following Tones 1, 3 and 4 are initially very tightly clustered together, while the contour following a Tone 2 starts about 10 Hz

⁴The initial F0 for Tone 2 may not reflect the Low, since it may take several cycles from the onset of the tone to reach that F0 point. This may account for the "flatness" of the Tone 2 contour, though a phonetic sandhi rule may also be influencing the results (Shih and Sproat, 1993).

higher. Again, the contour following a Tone 4 is lower than the others. At the end of the tone, all four contours are quite focussed; differences between ending points are all non-significant.

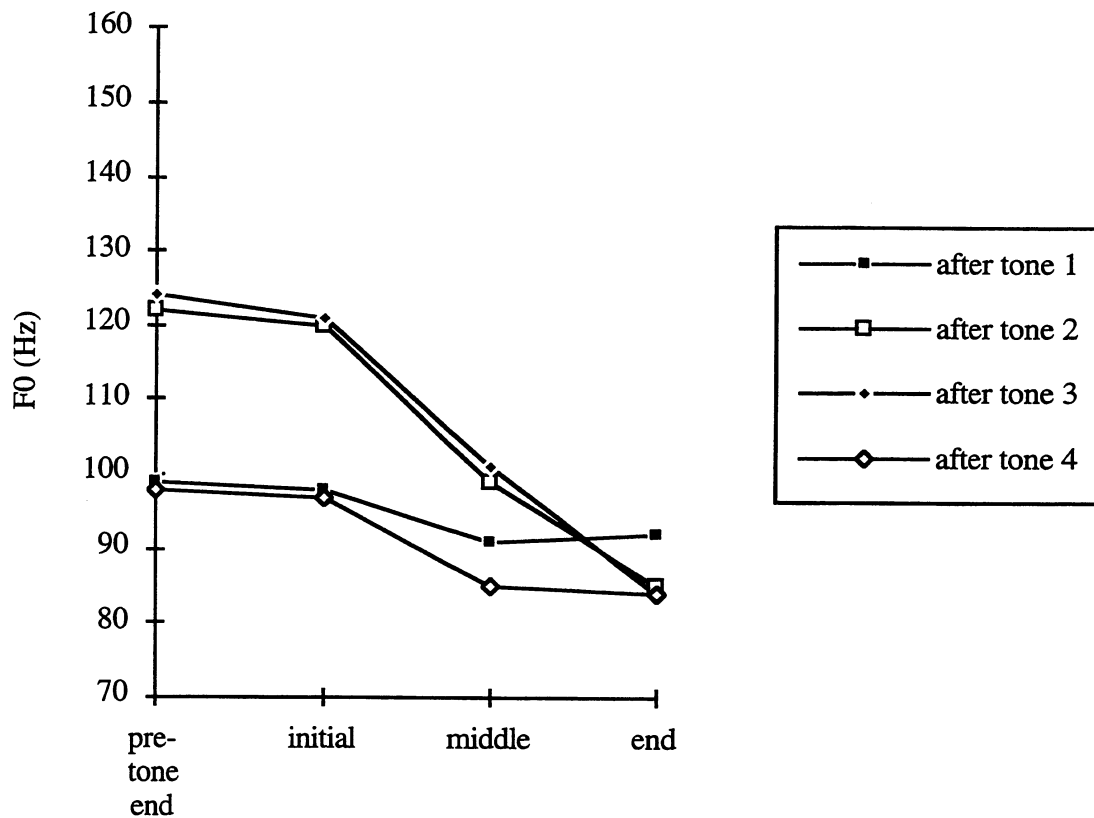


Figure 5: fundamental frequency of unstressed Tone 3 as a function of the preceding tone

Figure 5 shows that Tone 3 stimuli retain their falling property following each of the four tones, with F0 values ranging between 80-95 Hz at the lowest level. With regards to initial F0 points, the most obvious pattern among the Tone 3 contours is the clear separation between the F0 targets after Tones 2 and 3, and those after Tones 1 and 4. Rather than being an irregularity, this is instead a clear illustration of the third-tone sandhi rule, according to which in a string of two adjacent Tone 3 syllables, the first one changes to a Tone 2 (Shih, 1986), or, in rule form: $T3 \rightarrow T2 / _ T3$. This then predicts that there should also be two instances of the post-Tone 2 pattern represented in the graph. In fact, that is the case, as the contours following Tones 2 and 3 overlap. This example suggests

that tone sandhi is not sensitive to destressing, or occurs earlier in the phonology than stress rules. Similar to the other tones, the contour after Tone 2 is highest, and lowest after Tone 4. Thus, Tone 3 still retains several characteristic qualities related to its underlying specifications: the falling contour and distinctive lower range, occupying approximately 80-95 Hz, and the triggering of tone sandhi. As with the other lexical tones, the endpoints are focussed, showing no significant differences.

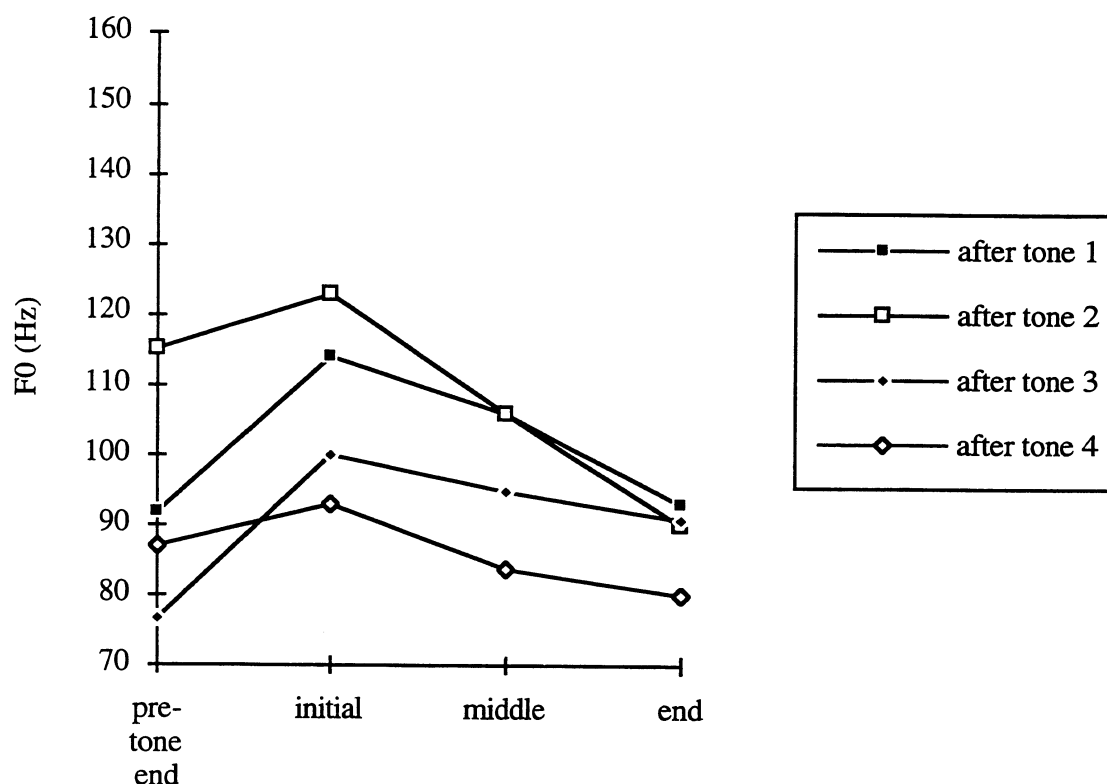


Figure 6: fundamental frequency of unstressed Tone 4 as a function of the preceding tone

The data for unstressed Tone 4 in Figure 6 show similarities with data of the previous tones. For example, initial F0 values are affected by coarticulation: the contour following a Tone 2 is highest, while the contour following a Tone 4 is lowest. With regards to contour, Tone 4 does appear to retain a falling F0 pattern in that each of the four F0 patterns in the chart show a general declination from higher to lower. In addition, there is still a general rise in F0 between the pre-tone end and the onset of Tone 4; crucially after tones which end high, Tones 1 and 2, the target tone rises to the typical higher starting level

for a Tone 4. Again, differences between endpoints are generally not significant. Thus, the Tone 4 data pattern with Tones 1-3 in three ways: 1) the phonetic F0 contour conforms with the citation falling shape, 2) F0 initial points are affected by coarticulation, and 3) ending points are concentrated, rather than diffuse.

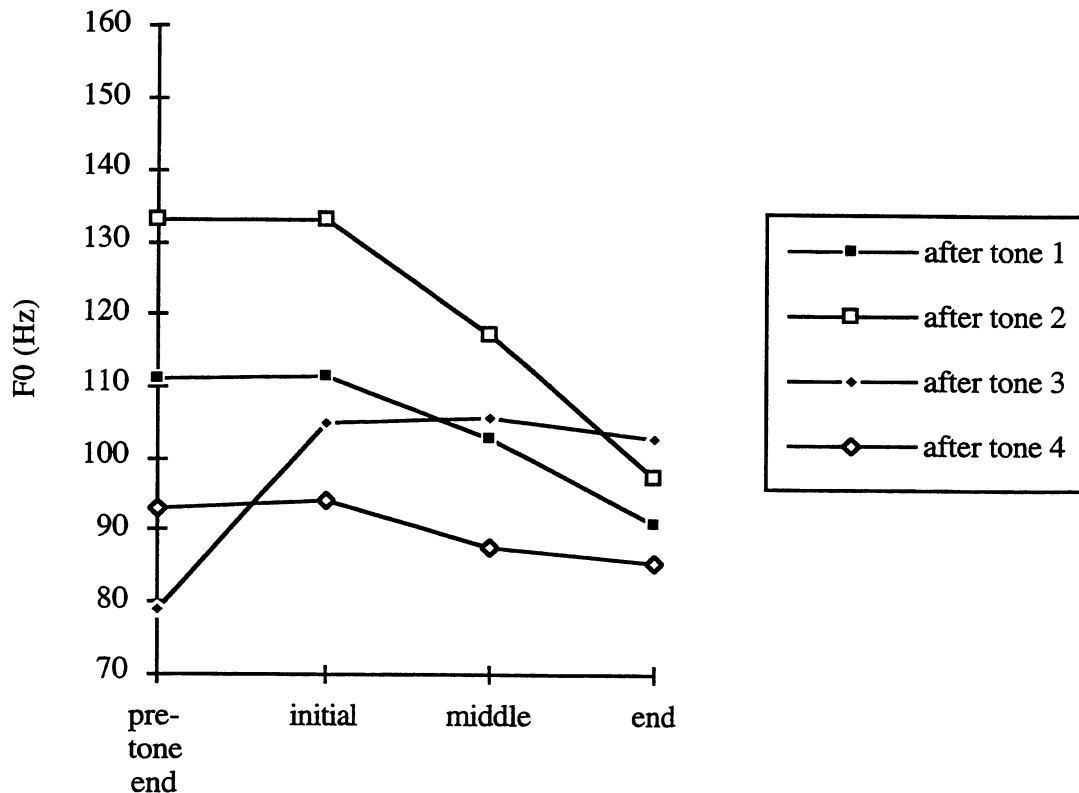


Figure 7: fundamental frequency of unstressed Tone 0 as a function of the preceding tone

The data in Figure 7 show F0 patterns of the neutral tone. The fundamental frequency at the end of the target syllable is highest after Tone 3, descending in F0 to Tone 2, then Tone 1, with the lowest F0 after Tone 4. These results are consistent with earlier reports (see references in section 1.2). Differences between initial and middle F0 points among the four contours of Tone 0 show no consistent patterns, however, as there appears to be no clear F0 focus until the end of the syllable. Moreover, although there is a general lowering in three cases, the fourth case, after Tone 3, is level after an initial rise. Since Tone 4 (high-falling) ends low, the low F0 value after Tone 4 may be assimilation. However, assimilation cannot account for the F0 level after Tone 1; although Tones 1 and 2 both end

high, the F0 pattern of the neutral tone does not take the same F0 "path" following each; Tone 2 ends high, but Tone 1 ends low.

Statistical analysis reveals a lack of cohesion in the beginning and ending fundamental frequencies, supporting the observation that the F0 pattern of Tone 0 is diffuse. All of the differences between beginning pitches are significant at the .05 level in paired, two-tailed t-tests, except for the points after Tones 1 and 3: for T0 after T1 compared with T0 after T2 [$t(15) = -4.55, p < .001$]; T0 after T1/T0 after T4 [$t(15) = 6.06, p < .001$]; T0 after T2/T0 after T3 [$t(15) = -5.21, p < .001$]; T0 after T2/T0 after T4 [$t(15) = 6.75, p < .001$]; and T0 after T3/T0 after T4 [$t(15) = 2.13, p < .007$]. This lack of focus in the neutral tone is further supported by the observation that beginning F0 points generally do not show distinctiveness from the pre-tone end; there is a gradual interpolation between the end of the previous tone and the ending point of the lexical tone, rather than a direct convergence toward a specific initial target. For example, in the contour following a Tone 1 the pre-tone end is 111 Hz, while the initial F0 value for that contour is also 111 Hz. The pre-tone end and initial F0 for the contour following a Tone 2 are also equivalent at 133 Hz. The contour after a Tone 3 may appear to be exceptional, since there is not the general F0 lowering exhibited by the other contours. This pattern is not unexpected, however, but rather supports the generally accepted hypothesis that the neutral tone receives a High from the preceding Tone 3, the High that cannot surface on a non-final Tone 3 syllable.

Ending F0 values for Tone 0 diverge to form two groups, one consisting of values after Tones 1 and 4, the other consisting of the F0 values after Tones 2 and 3. Paired, two-tailed t-test results are as follows: for T0 after T1/T0 after T2 [$t(15) = -2.97, p < .01$]; for T0 after T1/T0 after T3 [$t(15) = 2.78, p < .02$]; for T0 after T2/T0 after T4 [$t(15) = 3.20, p < .006$]; and for T0 after T3/T0 after T4 [$t(15) = 4.71, p < .001$].

To summarize, the neutral tone pattern shows diffuseness at the beginning of the tone, generally heading downward except after Tone 3, then forming two distinct categories at the end of the tone, one category consisting of the post-Tones 2 and 3 patterns, and one consisting of the post-Tones 1 and 4 patterns.

3.7 Discussion of F0 Contour Results

The F0 values and contour results above demonstrate that unstressed Tones 1-4 retain characteristic properties of their underlying contours in metrically weak positions. The F0 values for the lexical tones occur within a restricted range, despite their tonal environment, such that F0 targets are clearly defined. Sandhi effects are also evident where expected, as in the Tone 3 case. Furthermore, unstressed lexical tones exhibit the High, Mid (or upper

register Low), and Low (or lower register Low) distinctions along the F0 continuum, as would be predicted by their underlying tones. These phonological distinctions predict that Tone 1 initial targets, generated by an underlying High, should have higher F0 values than Tone 2 beginning points, and furthermore that the Tone 2 Low targets (represented here by initial points) should be higher than Tone 3 Low F0 values. This is indeed the case, as the average of Tone 1 High initial values is 118 Hz, significantly higher than the Tone 2 initial values of 101 Hz [$t(63) = 11.13$, $p < .001$]. Likewise, it would be expected that the Tone 2 Low targets (represented here by initial points) should be higher than Tone 3 Low F0 values. The results again confirm the prediction; the Tone 3 Low average (from Tone 3 ending F0 values) is 86 Hz, significantly lower than Tone 2 Low targets of 101 Hz, [$t(125) = 5.25$, $p < .001$].

While the lexical tones appear to occupy distinct F0 regions relative to each other, each of the tones shows differences among initial F0 points which appear to be the result of coarticulation. Observations of initial F0 points and their corresponding pre-tone endpoints indicate that tones with high offsets consistently raise the following initial F0, while tones with low offsets consistently lower it. This is most evident after a Tone 2, which ends high, and after a Tone 4, which ends low. Without exception, contours following a Tone 2 had the highest F0 contours initially, while at the low end, any contours following a Tone 4 had the lowest fundamental frequencies.

Coarticulation may also be responsible for at least some of the differentiation between Tone 3 contours. I noted above that the Tone 3 sandhi rule has applied to produce two post-Tone 2 contours. Since contours after Tone 2 overall show higher F0 values initially, it should not be surprising that the coarticulation effects generally evident in post-Tone 2 F0 points appear in both of these two contours. There is, however, a relatively large gap between post-Tone 2 and post-Tone 3 contours, and post-Tone 1 and post-Tone 4 contours, for which coarticulation may not be entirely responsible. Shih (1988) proposes an account of the post-Tone 2 higher F0 values, suggesting that it is a transposition of the Tone 2 F0 peak into the following syllable. She classifies the phenomenon as a phonological rule deleting the Tone 2 High before another High, where the Tone 2 aims for the High of the following syllable. However, this explanation would not account for the effect on the Tone 3 syllables in Figure 5, since Tone 3 is phonologically Low. Reasons for the larger gap between the Tone 3 contours may be discerned through further research.

Another somewhat surprising result occurs in the Tone 4 data. Although Tone 4 appears to be distinct from Tone 0, this contradicts the general belief that the particular target syllable, *ge*, is completely neutralized when unstressed (Dow, 1972, Tseng, 1981).

A more detailed comparison of the two tones uncovers further evidence of incomplete neutralization. For example, initial F0 values for unstressed Tone 4 appear to rise, rather than neutrally interpolate between the previous tone and its own ending F0 point. This is at least true for the Tone 4 contours following Tones 1 and 2. Statistical tests bear out this observation; while the initial F0 for a Tone 4, 114 Hz, is significantly higher than the 92 Hz offset of the preceding Tone 1 [$t(15) = -5.66, p < .001$], the initial F0 for Tone 0, 111 Hz, is not significantly different than the previous Tone 1 offset (111 Hz). The same is true for contours following a Tone 2, in that the initial F0 of Tone 4, 123 Hz, is significantly higher than the Tone 2 offset of 115 Hz [$t(15) = -3.32, p < .005$], while there is no significant difference between the Tone 0 initial F0, 133 Hz, and the Tone 2 offset of 133 Hz. Thus, while the Tone 4 F0 range corresponds with the range of the neutral tone, beginning and ending F0 points are more focussed than the Tone 0 points, and initial fundamental frequencies are significantly differentiated from the offset of the preceding tone in two cases.

A final point of discussion concerns the neutral tone. At this point, we can compare the F0 patterns reported here to several earlier studies. The chart below from Shen (1988:39) gives a summary of findings on the F0 level and/or contour of the neutral tone from several studies, most of which show pitch distinctions on a five-point scale where 1 is the lowest and 5 is the highest:

(8)

<u>author of study</u>	<u>level of Tone 0</u>	<u>contour of Tone 0</u>
Qi (1956)	2 after Tone 1 3 after Tone 2 4 after Tone 3 1 after Tone 4	
Dreher and Lee (1966)		41 after Tone 1 31 after Tone 2 23 after Tone 3 21 after Tone 4
Kratochvil (1968)	relatively low after high or falling tones mid after a rising tone high after a low tone	
Gao (1980) (levels are starting F0)	3 after Tone 1 3 after Tone 2 4 after Tone 3 2 after Tone 4	falling after Tone 1 falling after Tone 2 rising after Tone 3 falling after Tone 4

The list in (8) shows findings from previous acoustic studies on the neutral tone. Unlike several of these earlier experiments, the current study does not attempt to use the number scale to align the fundamental frequency levels⁵. However, both the results reported here and those in previous studies agree that the F0 pattern after Tone 4 leads to the lowest F0 level. Another point of agreement between all of the studies is that the pitch pattern after Tone 3 reaches the highest F0 level. One point of disagreement is where the F0 pattern after Tones 1 and 2 stand. The results from this study support Qi's (1956) placement of the post-Tone 2 F0 pattern being higher than post-Tone 1. By assigning a different numerical value to each contour, Qi does not demonstrate any grouping, however. Dreher and Lee (1966) show contours which place all ending F0 levels to be the same except after Tone 3. The numeric values of Gao's (1980) study are taken from the initial F0 level only, and hide any statistical distinctions, though the description of the contours is basically confirmed here. Kratochvil's (1968) generalizations are also supported by the data in this experiment.

In phonological work, Yip (1980) also describes the pitch of the neutral tone. Her description is in accord with the other studies in that she gives the pitch after Tone 3 as being the highest (number 4 on the 1-5 scale), and the pitch after Tone 4 as being the lowest (number 1), though it is unclear to what part of the tone contour her description refers. Yip's assigning of the same pitch (number 3) to both the F0 pattern after Tones 1 and 2 implies, contrary to the findings here, that these two F0 patterns should be phonetically alike.

To summarize the neutral tone discussion, the current experiment accords with previous work in the following conclusions: the F0 contour is generally falling, though somewhat more level after Tone 3; the ending F0 is highest after Tone 3, and lowest after Tone 4. Additional results of this experiment provide new evidence that ending F0 values for contours after Tones 1 and 4, and Tones 2 and 3 group together. Beginning F0 points generally show a gradual interpolation between the end of the previous tone and the ending point, rather than a direct convergence toward a specific initial target. Finally, the neutral tone generally has an initial F0 target which is non-distinct from the pre-tone endpoints. These characteristics provide a contrast to the unstressed lexical tone patterns, a contrast which is summarized in the next section.

⁵It is also unclear what point many of the previous studies took to be the one defining F0 level. Other studies, such as Gao (1980) and Dreher and Lee (1966) provide references to contours which correspond to the F0 patterns reported here.

3.8 Comparison of Unstressed Lexical Tones 1-4 with Tone 0

This section gives a brief recapitulation of the findings from the contour section above in order to compare instances of two types of unstressed tones: 1) unstressed lexical tones, and 2) the neutral tone, which derives from an atonic syllable.

The data above suggest that unstressed syllables retain underlying tone properties, both in F0 contour and in relative F0 level. The lexical tones appear to have F0 targets clearly defined. These unstressed tones reach distinguishable F0 targets as early as the beginning of the vowel, not showing gradient F0 levels throughout the vowel as in the case of the neutral tone. The distinctive patterns of the lexical tones can be related to their underlying tones by looking at general F0 patterns, the response to tone sandhi rules, and the restricted F0 ranges.

Tone 0, on the other hand, is widely divergent in fundamental frequency contour. Rather than clustering initially within a limited range, Tone 0 spans a range of about 40Hz. Statistical evidence supports the observation that initial F0 points do not converge toward a single locus, as opposed to the lexical tones. Furthermore, the ending F0 points are also more diffuse than in the lexical tones. These ending points divide into two groups, one at the higher F0 levels after Tones 2 and 3, and the other group at the lower levels after Tones 1 and 4. The initial fundamental frequency values for Tone 0 appear to be tied to the value of the preceding tone, such that the initial F0 value is virtually equivalent to the end of the preceding tone. The lexical tones show no such interpolation as a function of the preceding tone, rather, their initial F0 values appear to directly approach a target within a restricted F0 range. Finally, Tone 0 does not trigger tone sandhi, as the lexical tones do.

4.0 Conclusion

The purpose of this paper has been 1) to corroborate that two primary stress correlates in Mandarin Chinese are F0 range and duration, and 2) to examine the fundamental frequency patterns of unstressed lexical and atonic syllables, in order to determine to what extent tonal neutralization occurs in unstressed syllables.

The results of duration measurements confirmed that stress was indeed a factor in the duration of syllables and tones. In every case, stressed syllables had significantly longer durations than their unstressed counterparts. In addition, duration differences between tones were generally preserved when unstressed. Finally, all durations were significantly distinct from the duration of the atonic syllable.

Fundamental frequency range results also reinforce previous observations that stress affects the surface output of lexical tones. In particular, the data show that F0 range is

narrower in unstressed contexts, and expanded when stressed. Contrary to expected results, however, the data showed that the F0 range changes only in peak values, and that there is not a significant change in the low values. These results may provide evidence that phonological Lows do not contribute to stress and intonation contrasts to the same extent as Highs.

This study also presents new evidence that unstressed lexical tones are distinct from atonic syllables. In no case did the results indicate a loss of underlying tone, as might have been expected from the literature. In fact, although some assert that unstressed tones "lose" their underlying properties, these tones do not, however, take on properties of the neutral tone. Rather, the data presented here show that unstressed Tones 1-4 preserve their underlying contours and to a lesser extent, their relative fundamental frequency distinctions. Thus, there is a difference between lexical tones which are "neutralized" under lack of stress, and the "toneless" neutral tone, which never gets stressed. The target syllables in this study are more likely "weakly" stressed, though even at that level, categorical tonal properties are evident.

This investigation thus offers new evidence supporting a hypothesis that there are two categories of unstressed tones, each of which produces different tonal patterns. These two phenomena should not be lumped together, but rather examined separately, as they may give us important information about the phonological or phonetic rules of Mandarin Chinese tones.

The results of this experiment have several phonological implications. First of all, these data show that clitics in trisyllabic phrases do not necessarily undergo Tone Deletion or Tone Spreading, as often proposed (e.g., Yip, 1980). If Tone Deletion had occurred, we would have expected the Tone 0 fundamental frequency pattern to occur on the unstressed lexical tones. If Tone Spreading had occurred, the fundamental frequency patterns of the lexical tones would have taken on the more categorical properties of another tone. This was not the case. Another phonological question raised by this study concerns the representation of the neutral tone. As yet, no proposal has been made which accounts for why Tone 0 shares a low F0 value after Tones 1 and 4, but a higher F0 value after Tones 2 and 3. Any phonological proposal of Tone Spreading predicts that Tone 1 should pattern with Tone 2, and not with Tone 4, as in Yip (1980).

Finally, the facts presented here argue against complete neutralization of the classifier, *ge*, the Tone 4 syllable. Although this tone reduces in duration and F0 height under lack of stress, it patterns with the lexical tones in duration, contour, and in maintaining F0 initial and ending targets.

Studies in related issues may enable a deeper understanding of the phenomena reported in this study. For example, different atonic syllables in other phrasal configurations should be studied. Also, tones which appear as a result of phonological spreading rules may exhibit acoustic variations which are not present in lexically specified tones. In addition, the findings in this study could be enhanced by research on tonal coarticulation. Another area of inquiry includes the effects of intonation on weakly stressed tones: the findings in this study imply that stress itself may not be enough to delete a lexical tone; rather, intonation may have the greatest effect. The previous studies on intonation, or those using running speech (Tseng, 1981; Coster and Kratochvil, 1984; Shen, 1988; Shih, 1988) all mentioned the loss of inherent F0 contours and/or duration in unstressed syllables. The current experiment, however, controlled for intonation, finding much less variation than expected. Finally, perception studies based upon these results would improve our understanding of how acoustic variations due to stress are interpreted by listeners.

References

- Bo, Shi, and Zhang Jialu. 1986. Vowel intrinsic pitch in Standard Chinese. *Working Papers* 29. 169-190.
- Chao, Y-R. 1933. Tone and intonation in Chinese. *Bulletin of the Institute of History and Philology* 4. 121-134.
- Chao, Y-R. 1948. *Mandarin primer*. Cambridge: Harvard University Press.
- Chao, Y-R. 1956. Tone, intonation, singsong, chanting, recitative, tonal composition, and atonal composition in Chinese. In For R. Jakobson, 52-59. The Hague: Mouton.
- Chao, Yuen-ren. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Cheng, C-C. 1973. *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton.
- Cooper, J.M., and W.E. Sorensen. 1981. *Fundamental Frequency in Sentence Production*. Springer-Verlag: New York.
- Coster, D. C., and P. Kratochvil. 1984. Tone and Stress discrimination in normal Beijing dialect speech. *New Papers on Chinese Language Use*. 119-132.
- Davison, Deborah. 1991. Stress and Tonal Targets in Tianjin Mandarin. *UCLA WPP* 78. 1991. 58-77.
- Dow, Francis D. M. 1972. *An outline of Mandarin phonetics*. Canberra, Faculty of Asian Studies in association with the Australian National University Press.

- Dreher, J. and P-C. Lee. 1966. Instrumental investigation of single and paired Mandarin tonemes. Douglas Advanced Research Laboratory.
- Duanmu, San. 1990. A formal study of syllable, tone, stress and domain in Chinese languages. MIT Ph.D. dissertation.
- Fry, D.B. 1955. Duration and intensity as physical correlates of linguistic stress. *J. Acoust. Soc. Am.* 27. 765-8.
- Gao, Yu-zhen. 1980. Beijing hua de qingsheng wenti. *Yuyan Jiaoxue yu Yanjiu* 2. 82-98.
- Howie, John M. 1976. *Acoustical studies of Mandarin vowels and tones*. Cambridge University Press, Cambridge.
- Kratochvil, Paul. 1968. *The Chinese language today*. London: Hutchinson University Library.
- Kratochvil, Paul. 1971. An experiment in the perception of Peking dialect tones. *A Symposium on Chinese Grammar*, Scandinavian Institute of Asian Studies Monograph Series, No. 6. Lund. 7-20.
- Lehiste, I. 1970. *Suprasegmentals*. The MIT Press, Cambridge, MA.
- Li, Charles and Sandra Thompson. 1981. *Mandarin Chinese: A Functional reference Grammar*. University of CA Press, Berkeley.
- Liberman, Mark and Janet Pierrehumbert. 1984. Intonational Invariance under Changes in Pitch Range and Length. In M. Aronoff & R. T. Oehrle (eds.) *Language sound structure*. Cambridge, MA.: MIT Press. 157-233.
- Lindblom, B. 1963. Spectrographic study of vowel reduction. *J. Acoust. Soc. Am.* 35. 1773-1781.
- Nordenhake, Magnus, and Jan-Olof Svantesson. 1983. Duration of Standard Chinese Word Tones in Different Sentence Environments. *Working Papers* 25, Lund University. 105-111.
- Norman, J. 1988. *Chinese*. Cambridge University Press: Cambridge.
- Qi, Sheng-qiao. 1956. Hanyu de zidiao, tingdun yudiao de jiaohu guanxi. *Zhongguo Yuwen* 10. 10-12.
- Shen, Xiaonan. 1986. Phonology of the Prosody of Mandarin Chinese. *Cahiers de Linguistique - Asie Orientale*, vol. 15 (1). 171-178.
- Shen, Xiaonan. 1988. *The Prosody of Mandarin Chinese*. University of California Publications in Linguistics 118, Berkeley, University of California Press.
- Shen, Xiaonan. 1990. Tonal Coarticulation in Mandarin. *Journal of Phonetics* 18, 281-285.

- Shih, Chi-lin. 1986. The prosodic domain of tone sandhi in Chinese. Doctoral dissertation, UC-San Diego.
- Shih, Chi-lin. 1988. Tone and Intonation in Mandarin. in *Cornell WPP* 3. 83-109.
- Shih, Chi-lin, and Richard Sproat. 1993. Relative Prominence of Tonal Targets. A paper presented at the North American Conference on Chinese Linguistics 5, University of Delaware.
- Ting, A. C. 1971. *Mandarin tones in selected sentence environments: An acoustic study*. University of Wisconsin Ph.D. dissertation.
- Tseng, Chiu-yu. 1981. An acoustic phonetic study on tones in Mandarin Chinese. Ph.D. dissertation, Brown University.
- Yip, Moira. 1980. *The tonal phonology of Chinese*. MIT Press. Cambridge, MA.
- Zadoenko, T. 1958. An Experimental Study of the Atonic Syllables of Chinese. *Zhongguo Yuwen* 78. 581-587.