Journal of Phonetics (1997) **25,** 61–83

Contextual tonal variations in Mandarin

Yi Xu

Department of Communication Sciences and Disorders, Speech and Language Pathology, Northwestern University, 2299 North Campus Drive, Evanston, IL 60208, U.S.A.

Received 29th March 1996, and in revised form 5th September 1996

The present study examines acoustic variations of tones in Mandarin under the influence of different tonal contexts. In particular, variations in the four Mandarin tones due to anticipatory and carry-over effects are analyzed by examining the time course of f_0 contours of bi-tonal sequences. Using balanced nonsense sequences produced in different carriers with balanced tonal structures, this study establishes a baseline for local contextual tonal variation in Mandarin. It is found that anticipatory and carry-over tonal influences differ both in magnitude and in nature. Carry-over effects are mostly assimilatory: the *starting* f_0 of a tone is assimilated to the *offset* value of a previous tone. Anticipatory effects, on the other hand, are mostly dissimilatory: a low *onset* value of a tone raises the *maximum* f_0 value of a preceding tone. While the magnitude of the carry-over effect is large, anticipatory effects are relatively small.

1. Introduction

1.1. *Background*

Lexical tones in Mandarin and other East Asian languages are known for their gliding f_0 contours (Bai, 1934; Pike, 1948; Chao, 1956, 1968; Abramson, 1962, 1976; Lin, 1965, 1988; Howie, 1976; Tseng, 1981; Shih, 1986). When produced in isolation, these contours seem well defined and quite stable (Chao, 1968; Abramson, 1962, 1976; Lin, 1965, 1988). When produced in context, however, the tonal contours undergo certain variations depending on preceding and following tones (Chao, 1968; Abramson, 1979; Wu, 1984; Kratochvil, 1984; Shih, 1986; Wu, 1988; Gandour, Potisuk, Dechongkit, & Ponglorpisit, 1992*a,* 1992*b*; Lin & Yan, 1991; Shen, 1990; Shih & Sproat, 1992; Xu, 1993, 1994*a,* 1994*b*; Gandour, Potisuk, & Dechongkit, 1994; Potisuk, Gandour, & Harper, 1996).

An important issue in the study of contextual tonal variation concerns the magnitude and directionality of the contextual effect. For Vietnamese and Thai, it has been found that, in general, the effect of a preceding tone is greater in magnitude than that of a following tone. For example, Han & Kim (1974) examined the influence of the Vietnamese tones on one another in disyllabic utterances. They found that f_0 variations of the tones were greater in the second syllable than in the

0095-4470/97/010061 + 23 $25.00/0/jp960034

first syllable, indicating greater carry-over than anticipatory effects in Vietnamese. Similarly, in several recent studies of tonal coarticulation in Thai, Gandour and his colleagues found Thai tones to be influenced more by carry-over than by anticipatory coarticulation (Gandour *et al.,* 1992*a,* 1992*b,* 1994).

The findings on Mandarin tonal variation, however, have been mixed. Shen (1990), for example, analyzed all possible Mandarin tri-tonal combinations and found both carry-over and anticipatory effects, leading her to conclude that the bi-directional effects were symmetric. A different line of evidence, however, suggests asymmetrical bi-directional effects for Mandarin tones. Chao (1948, 1968) proposed a now well-cited sandhi rule, by which Tone 2 (a mid-rising tone) becomes Tone 1 (a high-level tone) when preceded by Tone 1 or Tone 2 and followed by any other stressed tone. This rule seems to indicate that Tone 2 is influenced more by a preceding tone than by a following tone (i.e., the preceding tone must be either Tone 1 or Tone 2, both of which have a high offset, whereas the following tone can be any of the four lexical tones, having either high, mid, or low onset). Although its phonological status has been questioned (Shih & Sproat, 1992; Xu, 1994*a*), the acoustic aspects of Chao's rule were largely supported by Shih & Sproat (1992) and by Xu (1994*a*). In addition, Xu (1994*a*) found that Tone 4 was also influenced more by a preceding tone than by a following tone. In a study of tonal coarticulation in tetrasyllabic tone combinations in Mandarin, Lin & Yan (1991) observed that tonal coarticulation in Mandarin was unidirectional, namely, each tone in Mandarin was affected either by a carry-over effect, or by an anticipatory effect. The findings by Shih & Sproat (1992), Lin & Yan (1991), and Xu (1994*a*) seem to indicate an asymmetry in the tonal coarticulation patterns in Mandarin, suggesting a similarity among Mandarin, Thai, and Vietnamese in this respect.

To further complicate the picture, there is experimental evidence that carry-over and anticipatory effects are different in nature. Data from Mandarin and Thai suggest that the anticipatory tonal influence is dissimilatory rather than assimilatory (Gandour *et al.,* 1992*b,* 1994; Xu, 1993, 1994*b*; Potisuk *et al.,* 1996). That is, a tone with a low onset seems to raise the f_0 value of the tone immediately preceding it rather than lowering its f_0 value. It is thus necessary to look at the issues of both directionality and the nature of anticipatory and carry-over tonal effects. The present study is designed to examine in detail both carry-over and anticipatory tonal influences in Mandarin.

1.2. *Possible patterns of contextual tonal variation*

Before presenting the experimental work, it is helpful to consider first the theoretically possible transitional patterns between adjacent tones. When two lexical tones are produced in sequence, especially when there is uninterrupted voicing through the two syllables that carry them, some kind of tonal transition has to occur (even between two identical tones, e.g., when two rising tones or falling tones are next to one another). To illustrate some of the possible transition patterns, Fig. 1 shows what may happen when a high-level tone is followed by a rising tone. Fig. 1(a) shows a case of Mutual Independence, where the f_0 value shifts down abruptly without any transition. Articulatorily, this is an impossible scenario because the

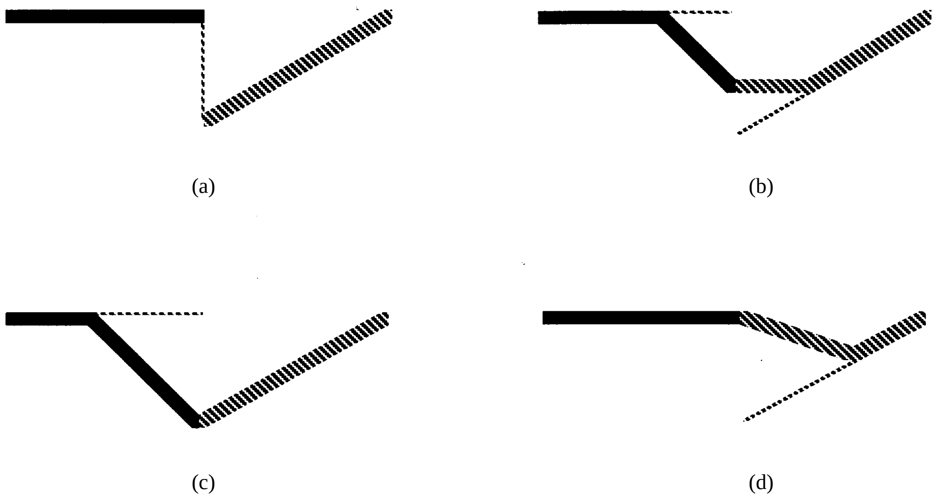

Figure 1. Possible tonal transitions between a high-level tone and a rising tone: (a) Mutual Independence; (b) Two-way Compromise; (c) Exclusive Anticipation; and (d) Exclusive Carry-over.

glottis is incapable of shifting from one state to another instantaneously. The other three configurations are all articulatorily possible. Fig. 1(b) shows a case of Two-way Compromise, where both tones accommodate to the other by adjusting their contours so that they meet halfway at the junction. This is a situation in which there are equal amounts of anticipatory and carry-over assimilation. (The term "assimilation" used here and in the rest of the paper refers only to a phonetic assimilatory effect. Phonological assimilation is not considered.) Fig. 1(c) shows a case where the high tone accommodates to the following rising tone, while the rising tone remains the same. As a result, much of the high-level contour of the high tone changes to falling. In this pattern, there is Exclusive Anticipatory Assimilation without any carry-over assimilation. In Fig. 1(d), the rising tone accommodates to the preceding high tone. Consequently, a large portion of the rising contour becomes falling. This is a situation in which there is Exclusive Carry-over Assimilation without any anticipatory assimilation. There are also other conceivable patterns, but the patterns shown here represent the major possible transition patterns when assimilation alone is taken into consideration.

In the combination patterns displayed here, at least a portion of each canonical tone contour remains intact, indicating that the tonal influence is local. It is theoretically possible, however, to have even more extensive assimilation. When that happens, the entire shape of a tone contour is affected.

What happens in actual tonal realization should tell us the relative importance of onset and offset values of the tones. If tone onset is more important (phonologically, articulatorily, or perceptually), then the final portion of a preceding tone would accommodate to the onset of the following tone, hence we should find the pattern in Fig. 1(c). However, if tone offset is particularly important, the pattern shown in Fig. 1(d) should be seen more often. If both onsets and offsets are equally important, then the pattern in Fig. 1(b) should be seen more often.

As discussed earlier, the findings made by Shih & Sproat (1992) and by Xu (1993) seem to suggest that pattern (d) in Fig. 1 is more likely to be found. If, however, there are equal amounts of anticipatory and carry-over assimilation, as suggested by Shen (1990), then the transition patterns should look more like (b) in Fig. 1. If the anticipatory influence is dissimilatory, as some previous studies have suggested (Gandour 1992*b*, 1994; Xu, 1993, 1994*b*; Potisuk *et al.*, 1996), more complicated bi-tonal patterns than those shown in Fig. 1 should be found. The following experiments were therefore designed to provide answers to these questions. In addition, the possibility of cross-syllabic tonal influences was examined.

2. Method

2.1. *Material*

Two reading lists were used for eliciting the Mandarin tones. The monosyllabic reading list, as shown below, consisted of the syllable /ma/ with the four Mandarin lexical tones. Phonologically, the four tones have the values of High-Level (Tone 1, as in mā), Mid-Rising (Tone 2, as in má), Falling-Rising (Tone 3, as in mǎ), and High-Falling (Tone 4, as in mà).

Monosyllabic reading list

Pinyin	mā	má	mǎ	mà
Character	妈	麻	马	骂
Gloss	"mother"	"hemp"	"horse"	"scold"

The disyllabic reading list consisted of the sequences /mama/ with all 16 possible bi-tonal combinations of the four Mandarin tones (see below). Among them, only the sequence 妈妈, "mother", is a word in Mandarin. The others are either nonsense combinations or sequences that are unnatural combinations in Mandarin, though they could be considered grammatical.

Disyllabic reading list

māmā 妈妈	māmá 妈麻	māmǎ 妈马	māmà 妈骂
mámā 麻妈	mámá 麻麻	mámǎ 麻马	mámà 麻骂
mǎmā 马妈	mǎmá 马麻	mǎmǎ 马马	mǎmà 马骂
màmā 骂妈	màmá 骂麻	màmǎ 骂马	màmà 骂骂

The initial nasal consonant /m/ was used to insure an unbroken f_0 contour throughout each syllable. At the same time, nasal-vowel boundaries are relatively easy to determine in the waveform, due to the abrupt shift in the waveform shape and amplitude level commonly seen at the boundary.

The monosyllables were produced in isolation, and the /mama/ sequences were produced in different carrier sentences as well as in isolation. The four carrier sentences were:

Carrier 1	Wǒjiaō____liańluò. 我教_____联络	"I teach____to communicate."
Carrier 2	Wǒjiaō____liaǹxí. 我教_____练习	"I teach____to practice."
Carrier 3	Wǒjiaò____liańluò. 我叫_____联络	"I tell____to communicate."
Carrier 4	Wǒjiaò____liaǹxí. 我叫_____练习	"I tell____to practice."

These carriers were designed in such a way that their phonetic as well as syntactic structures are very similar, except for their tones. There are two different pre-target syllables, /jiaō/ and /jiaò/. The former has a high tonal offset whereas the latter has a low offset. Likewise, one post-target syllable has a low tonal onset, /liań/, and the other a high onset, /liaǹ/. Except for the word 妈妈 "mother", none of the disyllabic combinations makes much sense in the carrier sentences.

The use of well-balanced nonsense utterances as shown here, instead of more natural but less-balanced utterances, is desirable at the current stage of studying contextual tonal variation, because fundamental frequency is influenced by many factors, only some of which are known to us. As will be shown later, the effects found in the present study can actually be seen scattered through the data of some previous studies. However, partly due to insufficient control, those studies could not reach systematic conclusions regarding these effects. In the present study, it is assumed that the same set of basic mechanisms govern contextual tonal variations both in natural utterances and in nonsense utterances, the only difference being the degree of variation, i.e., greater variation in natural utterances than in nonsense utterances. Hence, while the contextual effects found in this study cannot be considered to be maximal in magnitude, they should be relatively free of the influences of extraneous factors, thereby allowing for relatively straightforward interpretation.

2.2. *Recording*

Eight male native speakers of Mandarin, all of them graduate students studying in the United States, produced the target sentences. All of the speakers were born and raised in Beijing, and so were native speakers of Beijing Mandarin.

The monosyllabic reading list and the disyllabic reading list were printed in Chinese. The carrier sentences were printed (also in Chinese) below the /mama/ sequences. A pre-recorded pacing tape was used to control the speaking rate. On the tape were groups of six beeps with 3 s intervals. The beginning of each group was signaled by a double beep, and the end of each group by an extra long beep.

The speakers first produced the monosyllable /ma/ with the four lexical tones in isolation. They then produced all 16 /mama/ sequences in isolation. Finally, they produced all the /mama/ sequences in each of the four carrier sentences. They repeated each item six times, each repetition following a beep from the pacing tape. Successive repetition was used instead of interspersed repetition throughout a random list because a pilot recording session had found that some native subjects had difficulty reading the randomized list fluently without pause and frequent errors, since the target sequences differed from one another only in tone.

Half of the speakers produced the disyllabic sequences, both in isolation and in carrier sentences, in the same order as shown in the disyllabic reading list; the other half produced them in reverse order. The speakers were asked to make all their productions as natural as possible so that no pause was inserted between the two syllables. As a result, they produced both syllables in each /mama/ sequence with roughly the same stress. However, for the word 妈妈, "mother", they had to be told to say it with the same even stress pattern as the other sequences, instead of the usual trochee stress pattern associated with that word. They had no difficulty in following this instruction.

2.3. f_0 *extraction*

The utterances by four of the speakers were digitized at a sampling rate of 20 kHz at Haskins Laboratories, using the Haskins PCM system (Whalen, Wiley, Rubin, & Cooper, 1990). f_0 extraction was carried out manually. The waveform of each utterance was displayed on the computer screen in such a way that each glottal cycle was shown clearly. The onset of each glottal cycle in the waveform of a /mama/ sequence was marked with a label (no labeling was done on the carrier sentences). The time location of the labels was recorded. The utterances by the other four speakers were digitized at a sampling rate of 16 kHz at the MIT Research Laboratory of Electronics. The vocal pulse labeling was first made using the "epoch" program in the ESPS signal processing software package by Entropic Inc. The label files were then hand-edited to correct spurious labeling and to add boundary marks that divide the /m/ and /a/ segments.

All the label files were then processed by a separate computer program written by the author. The program transformed the intervals between successive labels into f_0 values and the f_0 curves obtained were smoothed using a simple window function incorporated into the program, eliminating discontinuities in the curves. In order to line up the raw f_0 curves for proper averaging, each f_0 curve was then equalized in time, using the same program, within each of the four segments in the /mama/ sequence, namely, the first nasal, the first vowel, the second nasal, and the second vowel.

2.4. f_0 *and duration measurements*

Several f_0 measurements were taken from the obtained f_0 curves: maximum f_0 in each segment, minimum f_0 in each segment, and f_0 values at five points in each segment (beginning, one quarter, midpoint, three quarters and end of the segment). Duration of each segment was measured as the distance between the first and last labels in the segment.

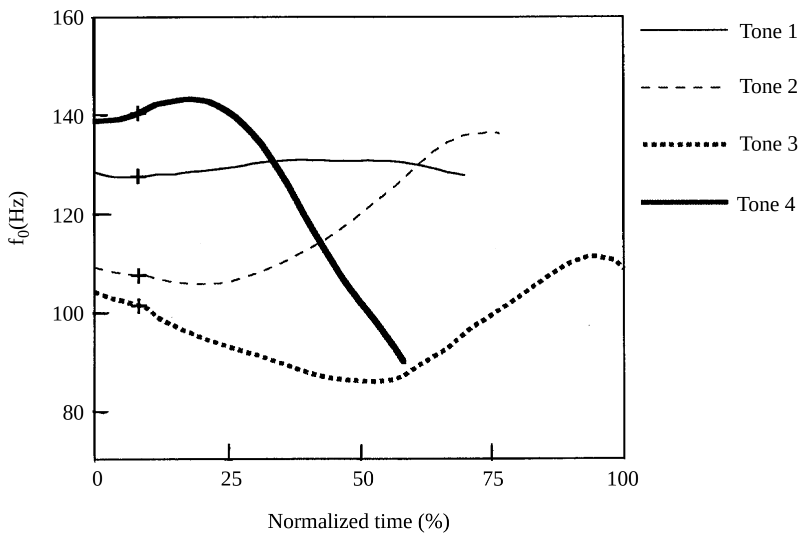

Figure 2. Mean f_0 contours (averaged over speakers and tokens; $n = 48$) of four Mandarin tones in the monosyllable /ma/ produced in isolation. The time is normalized, with all tones plotted with their average duration proportional to the average duration of Tone 3.

3. Results

3.1. *Tones produced in isolation*

To establish the canonical forms of the four Mandarin tones produced by the speakers in this study, tones produced in isolation were analyzed first. Fig. 2 shows the average f_0 contour of the syllable /ma/ in the four tones, obtained by averaging over all tokens produced by all eight speakers (48 utterances for each tone) and plotted as functions of average relative time of each segment, with the average duration of Tone 3 as the reference duration (100%). The cross symbol "+" indicates boundaries between /m/ and /a/. Tone 1 starts with a high f_0 value (near 130 Hz) and stays around that level throughout the syllable. Tone 2 starts with a low f_0 (near 110 Hz), then falls slightly before rising (starting at 20% into the vowel) throughout the remainder of the syllable. Tone 3 starts with an f_0 value slightly lower than the onset of Tone 2, falls to the lowest f_0 of all the four tones (about 90 Hz) right at the vowel midpoint, then rises sharply to the end of the syllable. Tone 4 starts with the highest f_0 value of the four tones (140 Hz), continues to rise before reaching the maximum about one fifth of the way into the vowel, then falls sharply to the end of the syllable.

In terms of duration, Tone 4 is the shortest tone (214 ms on average), Tone 3 is the longest (349 ms), while Tones 1 and 2 have intermediate durations, with Tone 2 (273 ms) longer than Tone 1 (247 ms). These duration patterns match those found in previous studies (e.g., Lin, 1988).

The f_0 patterns of tones produced in isolation reflect relatively directly the canonical forms of the tones. In multisyllabic utterances, the canonical forms will be distorted by various factors, including the adjacent onset and offset values of the neighboring tones. As suggested earlier and demonstrated by Xu (1994*a*), it is

Table I. Onset and offset values of Mandarin lexical tones

		OFFSET	
		High	Low
ONSET	High	Tone 1	Tone 4
	Low	Tone 2	Tone 3

probably the onset or offset values of a tone that are crucially affecting the f_0 contour of the tone next to it. Table I lists the canonical onset and offset values of the four Mandarin tones. Note that the phonological value ''Mid'' is not used here for the onset of Tone 2. It is assumed here that the onset of Tone 2 exerts similar, although smaller, contextual effects as those of Tone 3. Also in need of explanation is the offset value of Tone 3, which is listed here as Low rather than Mid or High as its canonical form seems to suggest. This is because many studies have shown that the final rise seen in isolation is usually absent in non-prepausal positions (Chao, 1968; Wu, 1984; Shih, 1986; Xu, 1993).

3.2. *Carry-over effects*

3.2.1. *Examining f_0 excursions in the /mama/ sequences*

Although the tones of /mama/ sequences produced both in isolation and in carrier sentences were analyzed, only the latter are presented here, largely because essentially the same patterns were found for both conditions, differing only in magnitude. Fig. 3 shows the f_0-contour variation due to the influence of the preceding tones in the /mama/ sequences. Each panel in the figure plots the same tone in the second syllable when preceded by four different tones. Each curve was obtained by averaging over all six repetitions produced by all eight speakers (48 utterances). The time scale is equalized for all the curves. A nasal segment is plotted with 10 points, while a vowel segment is plotted with 20 points. The f_0 contours plotted with duration of each segment proportional to the averaged actual duration of segment were also obtained. However, those plots did not provide more information about the effects to be discussed than the time-equalized plots shown in this and other similar figures, and so are not presented in the present paper.

Three observations concerning carry-over effects can be made from Fig. 3. (1) In the first syllable, different tones have distinct f_0 values throughout the syllable. (2) At the boundary between the two syllables, the starting f_0 of a given tone in the second syllable varies enormously (as large as 55 Hz on average) depending on the tone of the first syllable. (3) The differences due to the preceding syllable decrease gradually over time: during the initial nasal consonant, there are rapid f_0 movements, which are larger when the adjacent values of two neighboring tones are far apart than when they are similar; the f_0 differences are still quite substantial at the onset of the vowel in the second syllable (as large as 35 Hz), remain fairly sizeable at vowel midpoint (as large as 17 Hz), and are still discernible for Tone 1 and Tone 2 at vowel offset (7 Hz for Tone 1, and 4 Hz for Tone 2). Due to this gradual reduction

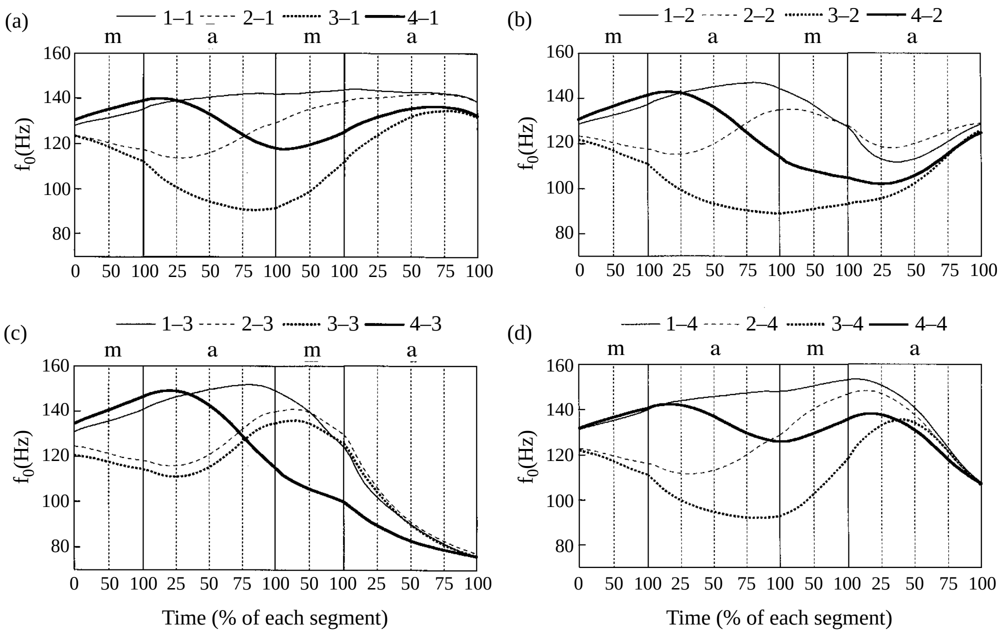

Figure 3. Carry-over effects; effects of preceding tone on f_0 contour of following tone in /mama/ sequences in Mandarin. In each panel, the tone in the second syllable is held constant (Tones 1–4 in (a) to (d) respectively), and the tone of the first syllable is varied. Thin (solid or dashed) lines indicate a high offset (Tone 1 or Tone 2) of the first syllable tone; thick lines indicate a low offset (Tone 3 or Tone 4) of the first syllable tone. Each curve is the average of 48 contours (six tokens per speaker).

of carry-over effects, a considerable portion of the f_0 curve for Tone 1 (phonologically, a high level tone) has a rising contour when the preceding tone is Tone 3 or Tone 4, both of which have a low offset. In the case of Tone 2, a mid-rising tone, f_0 does not start to rise until it is 35% into the vowel when following Tone 1, a high-level tone (compare with 20% in the isolated tone). In the case of Tone 4, a high-falling tone, f_0 does not start to fall until it is 40% into the vowel when following Tone 3, a low tone (compare with 20% in the isolated tone). The only case where a non-assimilatory carry-over effect is seen is when Tone 3 is preceded by another Tone 3. Due to a phonological tone-sandhi rule (Chao, 1948, 1968), the first Tone 3 assumes a rising f_0 contour that resembles that of Tone 2. (The slight separation of the contours for the derived Tone 2 and the underlying Tone 2 indicates that tone sandhi did not fully neutralize the differences phonetically. See Xu (1993) for detailed discussion.)

3.2.2. *Statistical analysis*

To examine the immediately adjacent as well as the cross-syllabic carry-over effects, two sets of analyses of variance were conducted. The first is a set of three-factor repeated-measure ANOVAs. The independent variables were the offset value of the syllable in the carrier sentence immediately preceding the /mama/ sequence (pre-target offset: high/low), offset value of the first syllable in the /mama/

sequence (pre-tonal offset: high/low), and tone of the second syllable (Tone 1–4) in the /mama/ sequence. The dependent variables were the f_0 values at vowel onset, one quarter, one half, and three quarters into the vowel, and at the end of the vowel of the second syllable in the /mama/ sequence. Pre-target and pre-tonal offset is defined as high when the tone of the syllable in question is Tone 1 or Tone 2 and as low when the tone is Tone 3 or Tone 4. Due to the phonological tone sandhi that occurs when Tone 3 is preceded by another Tone 3 [Fig. 3(c)], the sequence 3–3 is not included in the analysis. Statistical analysis of carry-over effects in individual speakers indicated the same patterns of f_0 variation and a similar magnitude of statistical significance as in the pooled analyses. Therefore, only results of pooled analyses are reported.

Table II lists the main effects and the two-way interactions obtained in the ANOVAs. None of the three-way interactions is significant, and they are not listed here.

First, starting with carry-over effects on the immediately following tone, as can be seen in Table II, the effect of pre-tonal offset is significant at the first four positions in the vowel, but not at the end of the vowel. However, there is a significant interaction between pre-tonal offset and Tone at vowel offset. Fig. 4 shows carry-over effects on different tones at different locations in the immediately following syllable. As can be inferred from Fig. 4, this interaction is due to a larger effect of pre-tonal offset on Tones 1 and 2 than on Tones 3 and 4. Newman–Keuls *post hoc* tests showed that, at vowel offset, the effect of pre-tonal offset was significant for Tone 1 and Tone 2 (at 0.05 significance level), but not significant for Tone 3 and Tone 4. Apparently, tones that have low offsets are less likely than tones with high offsets to show carry-over effects throughout the length of the vowel. Figure 4 also clearly shows the gradual reduction of carry-over effects over time, which reach an asymptote near the end of the vowel.

Turning now to cross-syllabic effects, also shown in Table II, the effect of pre-target offset is not significant at any of the positions, and there is no interaction between pre-target offset and pre-tonal offset. However, as shown in Fig. 5, the mean f_0 of Tones 1 and 2 at all positions of the second vowel is higher when pre-target offset is high than when it is low. The same is true of Tone 4 at the first four locations in the vowel. For Tone 3, the difference is very small in the first two locations, and is in the wrong direction in the last two locations in the vowel. It is apparent that the cross-syllabic carry-over effects, if any, are not large enough in the present experimental setting to reach a statistically significant level.

The second set of analyses concerned the carry-over effect of the carrier on the tone of the first syllable in the /mama/ sequence. These were repeated-measure ANOVAs with f_0 measurements at five positions in the first vowel of the /mama/ sequence as dependent variables. The independent variables were the offset value of the syllable in the carrier sentence immediately preceding the /mama/ sequence (pre-target offset: high/low), and tone of the first syllable (Tone 1–4) in the /mama/ sequence. The results are shown in Table III. Compared with Table II, the immediate carry-over effect shown in Table III seems somewhat smaller; the effect of pre-target offset is significant only at the first three positions in the vowel. Similar to Table II, however, are the significant interactions between pre-target offset and Tone when the main effect of pre-target offset is not significant. Examination of the effect on individual tones showed it to be much

TABLE II. ANOVA results for the effects of pre-target offset, pre-tonal offset and tone at five positions in the second vowel of the sequence /mama/

		Position									
		0		0.25		0.5		0.75		1	
Effect	df	F	P	F	P	F	P	F	P	F	P
Pre-target offset	1,7	1.09	0.331	1.09	0.332	1.34	0.285	0.84	0.389	0.22	0.657
Pre-tonal offset	1,7	37.87	0.001*	29.57	0.001*	19.37	0.003*	9.49	0.018*	3.06	0.124
Tone	3,21	53.84	0.001*	72.32	0.001*	72.43	0.001*	60.73	0.001*	51.91	0.001*
Pre-target offset × pre-tonal offset	1,7	1.90	0.211	2.07	0.193	2.44	0.162	1.04	0.342	1.10	0.328
Pre-target offset × tone	3,21	2.05	0.138	1.34	0.289	1.47	0.253	2.32	0.105	2.03	0.141
Pre-tonal offset × tone	3,21	4.80	0.011*	0.78	0.516	2.81	0.065	4.24	0.017*	3.69	0.028*

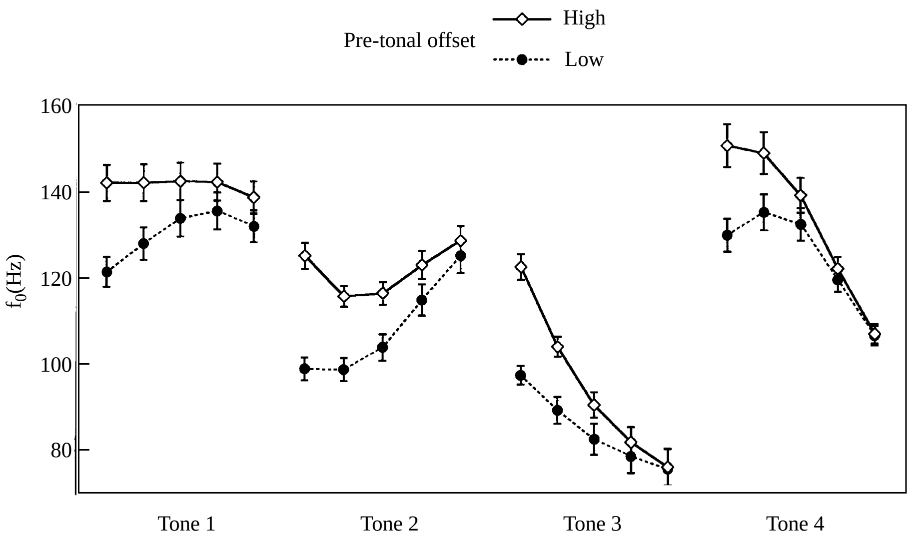

Figure 4. Effects of pre-tonal offset (high/low; see text) on mean f_0 values of Tones 1–4 at five temporal locations in the second vowel (0, 0.25, 0.5, 0.75, and 1). The horizontal bars indicate standard error.

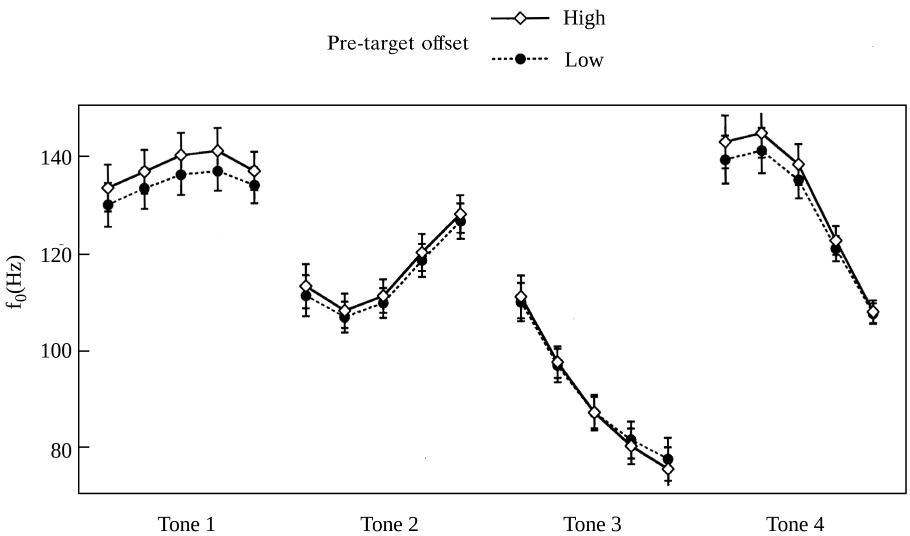

Figure 5. Cross-syllabic carry-over effects; effects of pre-target offset (high/low) on mean f_0 values of Tones 1–4 at five temporal locations in the second vowel (0, 0.25, 0.5, 0.75, and 1). The horizontal bars indicate standard error. None of the effects is statistically significant.

smaller on Tones 3 and 4, both of which have low offsets, than on Tones 1 and 2, whose offsets are both high.

A two-factor repeated-measure ANOVA was performed on the duration of the second vowel, with tone of the second syllable (1–4) and tone of the first syllable (1–4) as independent variables. The durational variation due to the tone of the

TABLE III. ANOVA results for the effects of pre-target offset and tone at five positions in the first vowel of the sequence /mama/

		Position									
		0		0.25		0.5		0.75		1	
Effect	df	F	P	F	P	F	P	F	P	F	P
Pre-target offset	1,7	15.48	0.010*	9.68	0.017*	6.39	0.039*	3.98	0.086	2.70	0.144
Tone	3,21	51.08	0.001*	52.92	0.001*	62.19	0.001*	63.67	0.001*	56.07	0.001*
Pre-target offset × first tone	3,21	4.60	0.013*	2.87	0.061	2.72	0.070	3.60	0.031*	4.11	0.019*

TABLE IV. Mean duration (in ms) of second vowel in the sequence /mama/ according to tone of first syllable (top row) and second syllable (bottom row)

	Tone			
	1	2	3	4
Duration of	131	133	136	134
second vowel (ms)	137	140	121	135

second syllable was significant, $F(3, 21) = 19.11$, $p < 0.001$, but the variation due to the tone of the first syllable was not. The mean durations of the vowel in the second syllable are listed in Table IV according to the tone of the first syllable (top row) and the second syllable (bottom row).

3.2.3. *Summary of carry-over effects*

Examination of the f_0 contours of Mandarin /mama/ sequences reveals substantial effects of the tone of the first syllable on the f_0 contour of the second syllable. Not surprisingly, the effects are due primarily to the offset value of the earlier tone. The influence is assimilatory, that is, a tone with a low offset lowers the f_0 of the following tone, and a tone with a high offset raises the f_0 of the following tone. The magnitude of the assimilatory effects decreases over time: during the initial nasal consonant, there are rapid f_0 movements, which are larger when the adjacent values of two neighboring tones are far apart than when they are more similar to each other; the f_0 differences due to carry-over effects remain quite sizeable during the vowel, though with reduced magnitude. The high f_0 region seems to be more susceptible to carry-over effects, and the lowest f_0 region seems to have strong resistance to the effects. The cross-syllabic carry-over effects were weak, although a trend continuing the immediate carry-over effects could still be seen in the f_0 means.

The finding that the contour of a tone is significantly influenced by the offset of the preceding tone suggests that the transitional pattern at the junction of two adjacent tones is that of either Mutual Comprise as shown in Fig. 1(b), or Exclusive Carry-over as shown in Fig. 1(d). Selecting between the two will depend on examination of the anticipatory effects.

3.3. *Anticipatory effects*

3.3.1. *Examining f_0 excursions in the /mama/ sequences*

Fig. 6 shows variations in f_0 contour of the tones when followed by different tones in the /mama/ sequences. In contrast to the strong tendency for the f_0 of the second syllable to assimilate to the offset value of the first syllable, the contours of the first syllable are much less affected by the identity of the following tone (the only exception being the 3–3 sequence due to phonological tone sandhi).

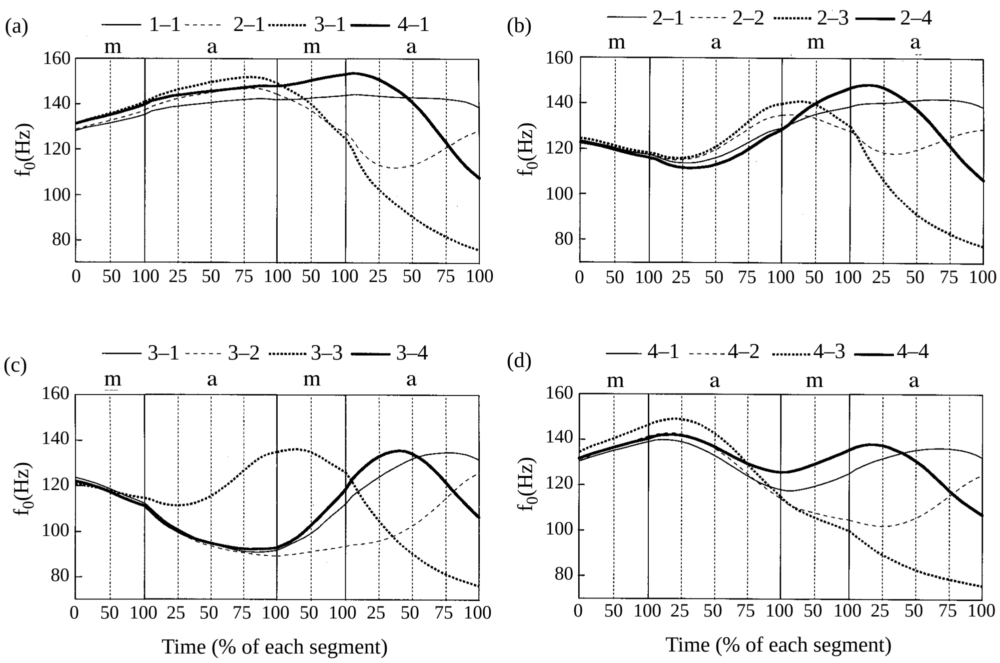

Figure 6. Anticipatory effects; effects of following tone on f_0 contour of preceding tone in /mama/ sequences in Mandarin. In each panel, the tone in the first syllable is held constant and the tone of the second syllable is varied. Solid (thick or thin) lines indicate a high onset (Tone 1 or 4) of the second syllable tone; dashed lines indicate a low onset (Tone 2 or 3) of the second syllable tone. Each curve is the average of 48 contours (six tokens per speaker).

More surprisingly, in most of the cases where the f_0 contour of the first syllable varies slightly, the effect seems to be dissimilatory. When the first syllable is followed by a tone with a low onset (i.e., Tones 2 and 3), its f_0 contour is somewhat higher than when it is followed by a tone with a high onset (i.e., Tones 1 and 4). The effect can be most clearly seen when the first syllable has Tone 2 [Fig. 6(b)] or Tone 4 [Fig. 6(d)]. Interestingly, while the effect is the greatest in the final portion of the f_0 contour of the first syllable for Tone 2, for Tone 4, it is greatest in the early portion. In both cases, it is the maximum f_0 value in the tonal contour that is being raised the most. The strongest anticipatory effect is exerted by Tone 3 in the second syllable; the maximum f_0 of a tone preceding Tone 3 is always higher than when followed by other tones. A similar raising effect is seen in Tone 2 when it follows another Tone 2. For Tone 1 and Tone 4, the raising effect due to a following Tone 2 is not obvious, and the maximum f_0 is no greater than when followed by Tone 4.

3.3.2. *Statistical analysis*

Because it seemed that it was the maximum f_0 of the first syllable that was affected the most by the onset value of the following tone, as seen in Fig. 6, the maximum f_0 of the preceding tone was used as the dependent variable in a three-factor

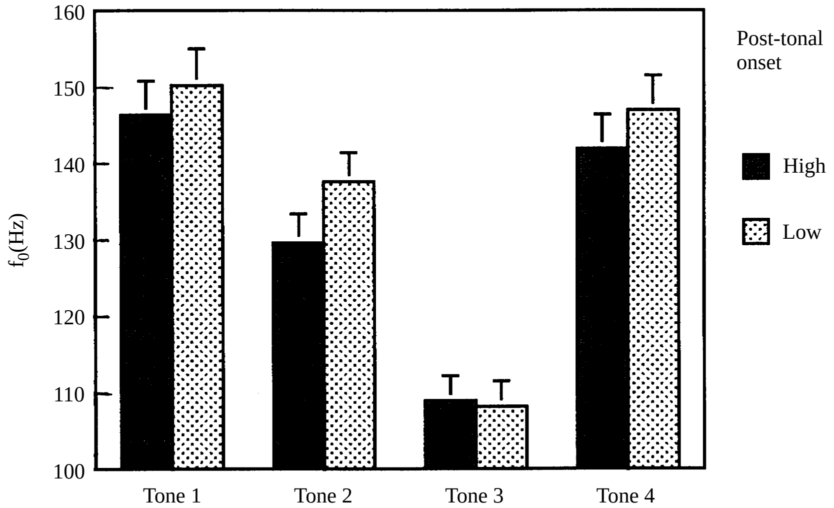

Figure 7. Effects of post-tonal onset (high/low; see text) on mean f_0 values measured at f_0 maximum of first vowel. The horizontal bars indicate standard error.

repeated-measure ANOVA for testing anticipatory effects. The independent variables were onset value of the following tone (post-tonal onset: high/low), onset value of the syllable in the carrier phrase immediately following /mama/ (post-target onset: high/low), and the tone of the first syllable (Tone 1–4). Due to the phonological tone sandhi that occurs when Tone 3 is followed by another Tone 3 [Fig. 6(c)], the sequence 3–3 is not included in the analysis. As in the case of carry-over effects, because anticipatory effects in individual speakers had the same patterns of f_0 variation and a similar magnitude of statistical significance as in the pooled analyses, only results of pooled analyses are reported.

The effect of post-tonal onset is highly significant, $F(1, 7) = 22.67$, $p < 0.005$, indicating that there is indeed anticipatory dissimilation as observed in Fig. 6. There is also a highly significant interaction between tone and post-tonal onset, $F(3, 21) = 23.40$, $p < 0.001$. As can be inferred from Fig. 7, the highly significant interaction is primarily due to lack of anticipatory effects on Tone 3, which is in sharp contrast with the other three tones. Two factors might have contributed to this. First, because the sequence 3–3 is not included in the ANOVA due to the phonological tone sandhi, only the sequence 3–2 is included in this condition. As can be seen in Fig. 6, Tone 2 does not raise the maximum f_0 of the preceding tone as much as Tone 3. Second, and probably more importantly, Tone 3 is the only tone in Mandarin that does not have a true high in its contour. The fact that there is virtually no difference in Tone 3 whether the following tone has a high or low onset suggests that the anticipatory effect is simply to raise the high f_0 region of the preceding tone.

Anticipatory effects due to the onset of the carrier phrase that follows the /mama/ sequence may also affect the f_0 of the second syllable in the /mama/ sequence. To check this effect, a two-factor ANOVA was conducted, with post-target onset and tone as independent variables, and maximum f_0 of the second /a/ as the dependent variable. The effect of post-target onset is not significant; neither is the interaction between tone and post-target onset. However, for all

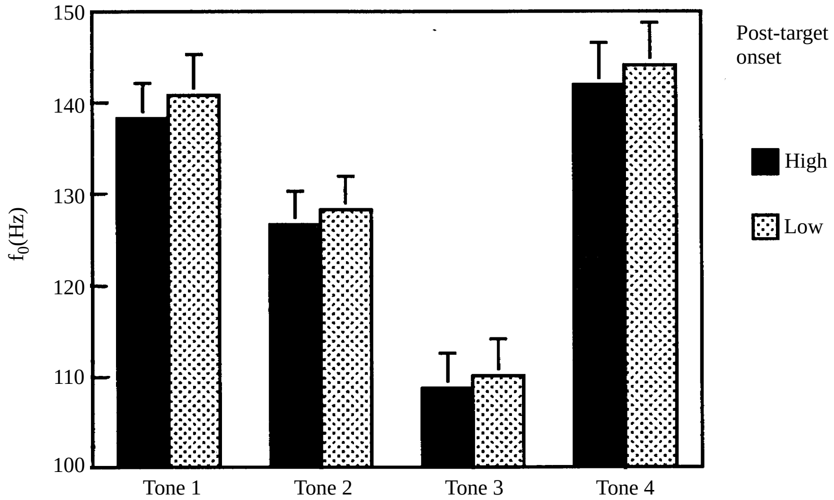

Figure 8. Effects of post-target onset (high/low) on mean f_0 values measured at f_0 maximum of second vowel. The horizontal bars indicate standard error.

tones, the direction of the variation is the same as what could have been seen if true dissimilation occured, as shown in Fig. 8. It is possible that because /mama/ was the focus word in the sentence, and thus received more stress, it had greater resistance to the anticipatory influence that the following tone might otherwise have exerted on it.

A two-factor repeated-measure ANOVA was performed on the duration of the first vowel in the /mama/ sequence, with tone of the first syllable (1–4) and tone of the second syllable (1–4) as independent variables. The durational variation due to the tone of the first syllable was significant, $F(3, 21) = 11.97$, $p < 0.001$, and the variation due to the tone of the second syllable was also significant, $F(3, 21) = 11.42$, $p < 0.001$. The mean durations of the vowel in the first syllable are listed in Table V according to the tone of the first syllable (top row) and the second syllable (bottom row). From Table V, it appears that the significant effect of the second tone in the /mama/ sequence is mainly due to the lengthy duration when the following syllable has Tone 3. Table IV shows that when the second syllable has Tone 3, the duration of the second vowel is the shortest (121 ms). It seems that there is a compensatory force at work that lengthens the duration of the first vowel when the second vowel is too short, although the compensation is not enough to make up the durational differences in the entire disyllabic sequence. The variation of the total

Table V. Mean duration (in ms) of first vowel in the sequence /mama/ according to tone of first syllable (top row) and second syllable (bottom row)

	Tone			
	1	2	3	4
Duration of first vowel	118	122	111	115
	114	116	121	115

duration of the sequence due to the tone of the second syllable is still significant, $F(3, 21) = 4.63$, $p < 0.05$.

3.3.3. *Summary of anticipatory effects*

The immediate anticipatory effects within the /mama/ sequences are largely dissimilatory. More precisely, a tone with a low onset tends to raise the maximum as well as the overall f_0 of the preceding tone, except in cases where the preceding tone is Tone 3. It is not clear whether there are any cross-syllabic anticipatory effects: the cross-syllabic main effect is not significant; the interaction between the cross-syllabic effect and the tone of the first syllable indicates that the possible effect could be a complicated one. Further study is needed to clarify this.

4. Discussion

4.1. *Exploring the nature of anticipatory dissimilation*

Although the anticipatory dissimilation found in the present study was never systematically reported for Mandarin before Xu (1993, 1994*b*), traces of it already existed in the data collected in several other studies. Partly due to lack of rigorous control of various factors affecting f_0, or lack of systematic examination of all tonal contexts, the traces did not lead to systematic conclusions regarding the anticipatory effect. Reexamination of these previous studies in the light of the findings reported in the present study made the earlier observations readily interpretable. For example, in a study of Beijing Mandarin stage speech (recordings of real stage speech), Kratochvil (1984) observed that in sequences of Tone 2 + Tone 4 and Tone 4 + Tone 4, the f_0 range of the first syllable was narrower than the average f_0 range of the particular tone group it belonged to, whereas in all the other cases, the f_0 range of the first syllable was wider than the average f_0 range of their respective tone groups. In other words, Tone 4, which has a high onset value, reduced the range of f_0 rising in the preceding Tone 2 and the range of f_0 falling in the preceding Tone 4. In a study on tones of Mandarin spoken in Taiwan, Shih (1986) noticed that Tone 3, a low tone, seemed to raise a high tonal target in the preceding tone. She observed that the high values of both Tone 4 and Tone 2 were higher when followed by Tone 3 than by other tones. A reexamination of her data, however, also shows a similar trend in other cases. Fig. 9 plots the starting and ending f_0 values of the syllable /fu/ in four tones when followed by different tones, measured by Shih (1986). Plotted in this way, it is now clear that not only do Tone 2 and Tone 4 have higher values before Tone 3, but Tone 1 does as well. In addition, Tone 2 appears to have a similar raising effect on preceding Tones 1 and 2. Even Shen (1990), who concluded that tonal coarticulation in Mandarin was symmetric, noticed that Tone 1 and Tone 2 "have the highest overall tonal values" when preceding Tones 2 and 3 (p. 285).

In addition to Mandarin, Thai, a language which also has contour lexical tones, was found to exhibit similar anticipatory dissimilation (Gandour *et al.*, 1992*b*, 1994). Gandour *et al.* (1994) found that in Thai "a low onset at the beginning of the influencing tone raises the height of the f_0 contour in the preceding syllable in target tones with higher f_0 offsets, whereas a high onset of the influencing tone lowers the height of the f_0 contour in the preceding syllable in target tones with higher f_0 offsets." (p. 490).

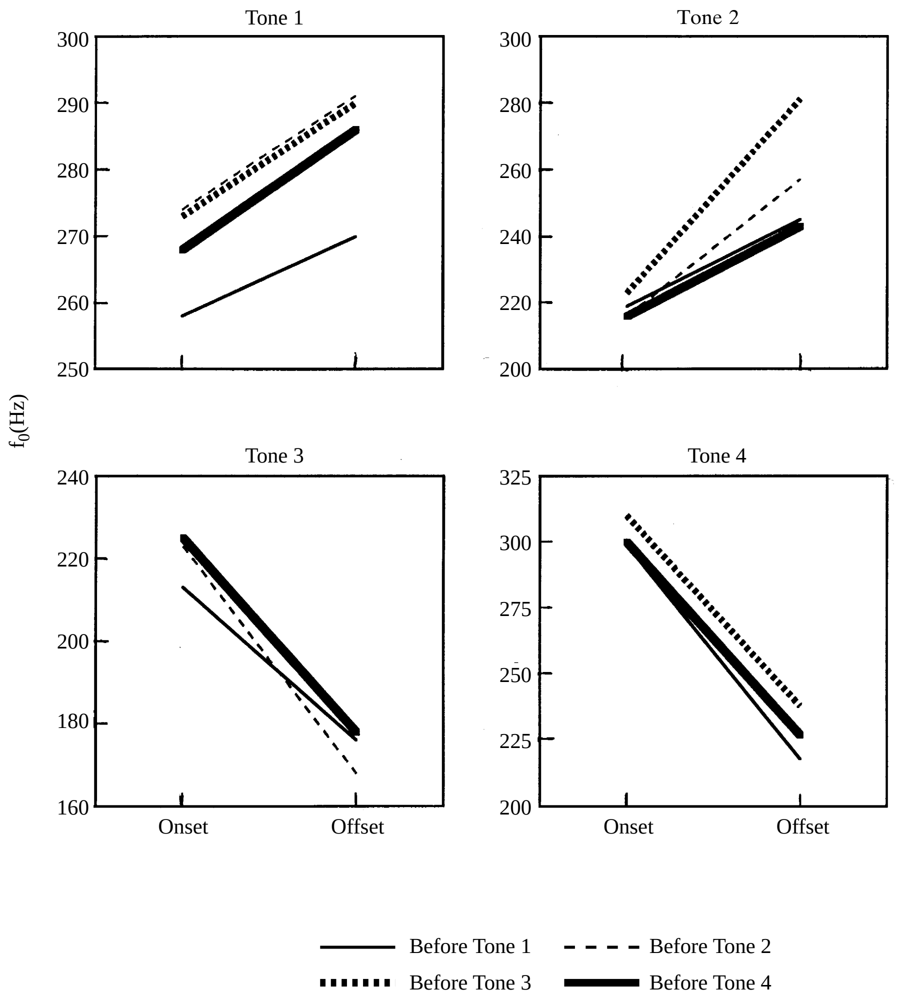

Figure 9. Starting and ending f_0 values of the Mandarin syllable /fu/ in four tones when followed by different tones. Measurements taken from Shih (1986), Table 5.

More interestingly, the raising of f_0 before a low tonal target found in the present study for Mandarin and in previous studies for Thai bears some similarity to two phonological rules that are fairly well known, namely, downstep and H-Raising, proposed for English and some African tone languages. In downstep, when a tone sequence of HLH (High Low High) is in the same prosodic unit, the second H is lower in frequency than the first H (Pierrehumbert, 1980). In H-Raising, an H in an HL sequence early in the sentence has higher values than in other contexts (Connell & Ladd, 1990; Laniran, 1992). Laniran (1992), however, suggested that in Yorùbá, both phenomena may be viewed as the consequence of "regressive upstep," a process by which an L raises an earlier H. In a recent study, Laniran & Clements (1995) also found further evidence for long distance regressive upstep in Yorùbá. (See Xu & Kim (1996), however, for somewhat different evidence.) Laniran's suggestion for Yorùbá appears to be similar to what was found in the present study for Mandarin and in earlier studies for Thai. At present, however, we only hint at the similarities between the anticipatory dissimilation found in Mandarin and Thai and the regressive upstep found in Yorùbá. Moreover, at the moment, we can only speculate about the actual mechanism that produces these phenomena.

It was shown earlier that minimum f_0 of Tones 3 and 4 was affected neither by anticipatory effects nor by carry-over effects, suggesting that a low tonal target is more resistant than a high tonal target to contextual influences. An adequate account of anticipatory tonal dissimilation ought to be able to explain why carry-over assimilation as well as anticipatory dissimilation affect the high frequency range more than the low frequency range. Two possible accounts are offered here.

First, it is possible that it is more difficult for a speaker to reach a low tonal target than a high tonal target due to phonatory constraints. As a result, the upper range in the preceding tone is extended. Supporting evidence for this account may be found in the observation that a low offset in a tone (as in Tone 3 or 4) does not exhibit either carry-over or anticipatory effects, as observed earlier, indicating that the low frequency range is much less flexible than the high frequency range. Physiological support for this assumption can be found in the data reported by Erickson (1976) for the production of tones in Thai. She observed that while the production of low f_0 involves the activities of the strap muscles (mainly thyrohyoid, sternohyoid, and sternothyroid), these muscles only contribute actively to lowering f_0 when it is to drop below a threshold level, usually near the midrange. This indicates that to reach the lower f_0 range, extra effort by the speaker may be needed. Thus, the lower an f_0 target, the more difficult it is to reach. Similar evidence is also found in a recent study on English (Erickson, Honda, Hirai, & Beckman, 1995).

Alternatively, the anticipatory dissimilation may serve to counteract declination, (i.e., "the tendency of pitch to drift downwards over the course of an intonation groups"; Pierrehumbert, 1979, p. 363). Declination has been found in many languages including Mandarin (Maeda, 1976; Gårding, 1987; 't Hart & Cohen, 1973). Because the normal declination is already going from high to low, when the tone pattern is HL across a tonal boundary, there is potential for confusing the tone pattern with the declination pattern. To reduce this potential confusion, the difference between the H and L target may be exaggerated by the speaker. However, due to the physical limit of the lower threshold, this exaggeration is accomplished by fully implementing the H target rather than by lowering the L target. Indirect support for this hypothesis comes from the finding by Pierrehumbert (1979) that when two stressed syllables are heard as equal in pitch, the second is actually lower, indicating that speakers normalize for declination in judging the relative height of peaks in the intonation contour. Thus, in the case of tone perception, this normalization process has the potential of reducing or even obliterating the actual relative pitches if the pitch sequence across the tone boundary is HL.

4.2. *Tones produced in succession: how they interact with one another*

Examination of contextual tonal variation in Mandarin reveals very different patterns for carry-over and anticipatory effects. Considering for the moment only carry-over effects, the pattern in Fig. 1d, the "Exclusive Carry-over" pattern, presents the best match with the bi-tonal patterns in Mandarin. When two tones are produced in succession, the final portion of the first tone closely follows its intended trajectory to the end of the syllable; the second tone, however, has to start from where the previous tone ends, and only approaches its targeted curve towards the later portion of the syllable. One possible benefit of this way of realizing both tones is that the potential conflict at the neuro-muscular level is avoided when the offset of the preceding tone and the onset of the following tone have different values. By guaranteeing only the offset value of each tone, unnecessary muscle strain and waste of energy would be prevented. It is also possible that this pattern of tonal interaction reflects the importance of f_0 offset for pitch perception. Heinz, Lindblom, & Lindqvist (1968) found in a tune-matching experiment that, when asked to match the pitch of a steady-state sinusoid to that of a sweeping sinusoid, listeners matched the steady-state tone to the terminal frequency of the sweeping tone rather than to its average frequency, whether or not the sweeping sinusoid was ascending or descending in frequency. If the perception of contours of lexical tones is also heavily dependent on f_0 offsets (as in the pitch perception of sweeping sinusoids), it would then be quite natural for the speakers to guarantee only the shape of the tonal offset whenever there is a conflict at the boundary between two adjacent tones. At least one tone perception study has indicated that this could be the case (Whalen & Xu, 1992).

In contrast, the anticipatory effects found in the present study do not at all match the "Exclusive Anticipatory" effects in Fig. 1(c). When a tone other than Tone 3 is in the first syllable, its maximum f_0 varies *inversely* with the onset of the tone of the second syllable; higher when followed by a low onset, and lower when followed by a high onset. Hence, the final portion of the f_0 curve of Tones 1 and 2 is raised; for Tone 4, the earlier portion of its f_0 curve is raised. Clearly, the mechanism for producing the anticipatory effects is completely different from the mechanism for producing the carry-over effects. Although the true nature of this anticipatory dissimilation cannot be determined by the present study, some inferences may be drawn from other studies as suggested earlier.

What also remains unclear is the domain of the anticipatory and carry-over effects. The present study did not find statistical evidence for carry-over effects beyond the immediately adjacent syllable. However, as mentioned earlier, this study was not designed to find maximum contextual effects. Nonsense disyllabic sequences produced in carrier frames in which the disyllables constitute the focus of the sentence are not likely to be susceptible to maximum contextual influences. In all likelihood, the results of the present study reflect rather near minimum contextual effects. To explore the fuller domain of the contextual effect found in the present study, studies designed particularly for that purpose are needed.

5. Conclusion

Although the full extent of possible contextual tonal variation awaits further study, the near minimum contextual tonal effects found in the present study are quite interesting. As long as there is no pause between two adjacent syllables, even if the sequence forms a nonword (and so could not have been said many times by the speaker before the experiment), the f_0 contour of either tone in the sequence is substantially influenced by the surrounding tones. The most apparent influence is from the preceding tone, whose offset value virtually determines the starting f_0 of the following tone, and whose influence can be seen at least two thirds of the way into the vowel of the following syllable and sometimes even through the end of the following syllable. Due to the gradual reduction of the carry-over influence, the contour shape of a following tone can sometimes be distorted beyond easy recognition by the eye without taking the preceding tone into consideration. The influence from the following tone is more subtle, but no less interesting, than the carry-over effects. A low tonal onset raises, rather than lowers, the f_0 of the preceding tone. The raising effect is most clearly seen in the location of maximum f_0 in the tonal contour of the preceding tone, rather than always at the offset of the contour. While the actual cause of the anticipatory effects can only be speculated on at this moment, it is already fairly obvious from the data of the present study that the anticipatory and carry-over effects are due to different mechanisms. Naturally, further studies will be needed to investigate the full scope of both kinds of effects as well as their underlying mechanisms and their interaction with other prosodic and segmental factors in speech production and perception.

The author would like to thank Ignatius Mattingly, Bruno Repp and Bruce Smith for valuable help at various stages of the study, and Donna Erickson and Philip Rose for helpful comments on an earlier version of the paper. Part of the data in the study was reported in Xu (1993, 1994*b*).

References

Abramson, A. S. (1962) The vowels and tones of standard Thai: acoustical measurements and experiments. Bloomington: Indiana University Research Center in Anthropology, Folklore, and Linguistics, Pub. 20

Abramson, A. S. (1976) Thai tones as a reference system. In Thai linguistics in honor of Fang-Kuei Li (T. W. Gething, J. G. Harris & P. Kullavanijaya, editors), pp. 1–12. Bangkok: Chulalongkorn University Press

Abramson, A. S. (1979) The coarticulation of tones: an acoustic study of Thai. In *Studies in Tai and Mon-Khmer phonetics and phonology in honour of Eugénie J. A. Henderson* (T. L. Thongkum, P. Kullavanijaya, V. Panupong & K. Tingsabadh, editors), pp. 1–9. Bangkok: Chulalongkorn University Press

Bai, D. (1934) Guanzhong shengdiao shiyan lu [Experiments with tones of Guanzhong dialects]. In *Shiyusuo Jikan* [*A Collection by Shiyusuo*], pp. 355–361

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 Roman Jakobson* (M. Halle, editor), pp. 52–59. The Hague: Mouton

Chao, Y. R. (1968) A grammar of spoken Chinese. Berkeley, CA: University of California Press

Connell, B. & Ladd, D. R. (1990) Aspects of pitch realization in Yorùbá, *Phonology,* **7**, 1–29

Erickson, D. M. (1976) A physiological analysis of the tones of Thai. Doctoral dissertation, The University of Connecticut

Erickson, D., Honda, K., Hirai, H. & Beckman, M. E. (1995) The production of low tones in English intonation, *Journal of Phonetics,* **23**, 179–188

Gandour, J., Potisuk, S., Dechongkit, S. & Ponglorpisit, S. (1992*a*) Tonal coarticulation in Thai disyllabic utterances: a preliminary study, *Linguistics of the Tibeto-Burman Area,* **15**, 93–110

Gandour, J., Potisuk, S., Dechongkit, S. & Ponglorpisit, S. (1992*b*) Anticipatory tonal coarticulation in Thai noun compounds, *Linguistics of the Tibeto-Burman Area,* **15**, 111–124

Gandour, J., Potisuk, S. & Dechongkit, S. (1994) Tonal coarticulation in Thai, *Journal of Phonetics,* **22**, 477–492

Gårding, E. (1987) Speech act and tonal pattern in Standard Chinese, *Phonetica,* **44**, 13–29

Han, M. S. & Kim, K. (1974) Phonetic variation of Vietnamese tones in disyllabic utterances, *Journal of Phonetics,* **2**, 223–232

Heinz, J. M., Lindblom, B. E. F. & Lindqvist, J. C. K.-G. (1968) Patterns of residual masking for sounds with speech-like characteristics, *IEEE,* **AA-16**, 107–111

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

Kratochvil, P. (1984) Phonetic tone sandhi in Beijing dialect stage speech, *Cahiers de Linguistique, Asie Orientale,* **13,** 135–174

Laniran, Y. (1992) Intonation in tone languages: the phonetic implementation of tones in Yorùbá. Doctoral dissertation, Cornell University

Laniran, Y. O. & Clements, G. N. (1995) A long-distance dependency in Yorùbá tone realization. In *Proceedings of the XIIIth international congress of phonetic sciences,* vol. **2**, pp. 734–737. Stockholm

Lin, M. (1965) Yingao xianshiqi yu Putonghua shengdiao yingao texing [The pitch indicator and the pitch characteristics of tones in Standard Chinese], *Acta Acoustica Sinica,* **2**, 8–15

Lin, M. (1988) Putonghua shengdiao de shengxue texing he zhijue zhengzhao [Perceptual cues for tones in Standard Chinese], *Zhongguo Yuwen* [*Chinese Linguistics*], **204**, 182–193

Lin, M. & Yan, J. (1991) Tonal coarticulation patterns in quadrasyllabic words and phrases of Mandarin. In *Proceedings of the XIIth international congress of phonetic sciences,* vol. **3**, pp. 242–245

Maeda, S. (1976) *A characterization of American English intonation.* Doctoral dissertation, MIT, Cambridge, MA

Pierrehumbert, J. (1979) The perception of fundamental frequency declination, *Journal of the Acoustical Society of America,* **66**, 363–369

Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation.* Doctoral dissertation, MIT, Cambridge, MA

Pike, K. L. (1948) *Tone languages.* Ann Arbor: University of Michigan Press

Potisuk, S., Gandour, J. & Harper, M. P. (1996). Contextual variations in trisyllabic sequences of Thai tones, *Phonetica,* **53,** 200–220

Shen, X. S. (1990) Tonal coarticulation in Mandarin, *Journal of Phonetics,* **18**, 281–295

Shih, C. (1986) *The phonetics of the Chinese tonal system.* Technical memorandum. AT & T Bell Laboratories

Shih, C. & Sproat, R. (1992) Variations of the Mandarin rising tone. In *Proceedings of the IRCS workshop on prosody in natural speech No.* 92–37, pp. 193–200. Philadelphia: The Institute for Research in Cognitive Science, University of Pennsylvania

't Hart, J. & Cohen, A. (1973) Intonation by rule: a perceptual quest, *Journal of Phonetics,* **1**, 309–327

Tseng, C. (1981) *An acoustic phonetic study on tones in Mandarin Chinese.* Doctoral dissertation, Brown University

Whalen, D. H., Wiley, E. R., Rubin, P. E. & Cooper, F. S. (1990) The Haskins Laboratories' pulse code modulation (PCM) system, *Behavior Research Methods, Instruments & Computers,* **22**, 550–559

Whalen, D. H. & Xu, Y. (1992) Information for Mandarin tones in the amplitude contour and in brief segments, *Phonetica,* **49,** 25–47

Wu, Z. (1984) Putonghua sanzizu biandiao guilü [Rules of tone sandhi in trisyllabic words in Standard Chinese], *Zhongguo Yuyan Xuebao* [*Bulletin of Chinese Linguistics*], **2**, 70–92

Wu, Z. (1988) Tone-sandhi patterns of quadra-syllabic combinations in Standard Chinese, *Report of Phonetic Research, Institute of Linguistics* (*CASS*), *Beijing, China,* PL-ARPR/1988, 1–13

Xu, Y. (1993) *Contextual tonal variation in Mandarin Chinese.* Doctoral dissertation, The University of Connecticut

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

Xu, Y. (1994*b*) Asymmetry in contextual tonal variation in Mandarin. In *Advances in the study of Chinese language processing,* (H.-W. Chang, J.-T. Huang, C.-W. Hue & O. J. L. Tzeng, editors), Vol. **1**, pp. 383–396. Taipei: Department of Psychology, National Taiwan University

Xu, Y. & Kim, J. (1996) Downstep, regressive upstep, H-raising or what?—sorting out the phonetic mechanisms of a set of "phonological" phenomena, *Journal of the Acoustical Society of America,* **100,** 5aSC7