VARIATIONS OF THE MANDARIN RISING TONE*

Chilin Shih
Department of East Asian Languages and Cultures
Rutgers University
College Avenue Campus
New Brunswick, NJ 08903

Richard Sproat
Linguistics Research Department
AT&T Bell Laboratories
600 Mountain Avenue
Murray Hill, NJ 07974

ABSTRACT

This paper uses the Mandarin rising tone (tone 2) as an example to illustrate the range of tonal variation in speech, from words read in isolation, words read in sentence frames, to words produced in conversation. It is shown that the 2nd Tone Sandhi Rule (2TS) described by Chao [1] is the result of a phonetic implementation rule, which applies to the low target of a rising tone in high tone context when the rising tone in question is in prosodically weak positions. The amount of pitch drop to the low target of a rising tone varies with the prosodic strength of the syllable. As a result, the pitch contour of an extremely weak rising tone in high tone context approaches the shape of a high level tone.

1. Background

Mandarin has four lexical tones: High level, Rising, Low, and Falling. The following table lists the tones in the traditional nomenclature, pitch contour, and tonal targets in H (high) and L (low).

Name	Pitch Contour	Targets
Tone 1	High level	H
Tone 2	Rising	LH
Tone 3	Low	L
Tone 4	Falling	HL

Table 1: Mandarin Tones

Tone 3 has other variants. Most notably, a falling-rising shape may surface in the sentence final position. Therefore the tone is also known as a dipping tone. In addition to the four tonal categories, some suffixes in Mandarin do not have an inherent tone. A syllable without tone is weak; its duration is typically only half of a corresponding tone-carrying syllable. The pitch value of a toneless syllable is determined primarily by the preceding tone. The lack of tone on a syllable is traditionally referred to as neutral tone, or tone 0.

We take Y.R. Chao's ([1]) description of the 2nd Tone Sandhi (2TS) as a starting point:

> "A tone sandhi of minor importance has to do with the change of the 2nd to a 1st Tone in three-syllable groups. If in a three-syllable word or phrase ABC, A is in the 1st or 2nd Tone, B in the 2nd Tone, and C in any except the neutral tone ... then B changes into the 1st Tone for speech at conversational speed, but does not change at a more deliberate speed." (p. 27-28).

The examples Chao gave include the following. (We use the Pinyin romanization system through out this paper. Angle brackets highlight a *changed* tone: [1] indicates a phonetic tone 1 with an underlying tone 2 source).

1. xi1 yang2 shen1 → xi1 yang[1] shen1
   west-ocean-ginseng
   "western ginseng"
2. cong1 you2 bing3 → cong1 you[1] bing3
   green-onion oil pancake
   "green onion pancake"
3. long2 fu2 si4 → long2 fu[1] si4
   Long Fu Temple
4. Shei2 neng2 fei1? → Shei2 neng[1] fei1
   who can fly
   "Who can fly?"

* This research is supported by National Science Foundation Grant BNS-9021274 and Rutgers University Research Council Grant 2-02364 to the first author. Many thanks to John Kingston and all the participants of the Prosody Workshop for valuable discussions.

5. hao3 ji3 zhong3 → hao[2] ji[2] zhong3
→ hao[2] ji[1] zhong3

quite several kind
"quite a few kinds"

Examples (1) and (2) show the rule applying after a tone 1 syllable, while (3-5) show the rule applying after tone 2. (4) and (5) show that the rule is not restricted to lexical items. It applies to sentences, questions as well as statements. Finally, (5) shows that the rule can affect a derived tone 2. The intermediate step in (5) shows the effect of another tone sandhi rule, Mandarin 3rd Tone sandhi (3TS), which changes the first of two tone 3 syllables to tone 2. In (5), the 3TS feeds the 2TS, allowing further change of the middle syllable to tone 1.

Furthermore, Cheng [2] shows that the 2TS may apply to a sequence of tone 2's derived via the 3TS, creating a stretch of tone [1]'s from underlying tone 3's. Cheng also notes that such a chain effect is only found in short phrases, which are not likely to extend beyond 6 or 7 syllables.

6. Lao3 Li3 mai3 hao3 jiu3
Old Li buy good wine
"Old Li buys good wine."

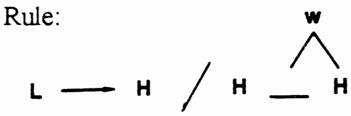
[2] [2] [2] [2] 3 --After 3TS
[2] [1] [1] [1] 3 --After 2TS

Our experimental results will be shown to support Chao's description in full, that the 2TS is frequently found in conversational speech, but much rarer in deliberate speech. The focus of this paper is the reasons behind the peculiar rule description. There are two issues involved: First, what is the motivation behind the rule description? Second, is the 2TS a phonological rule? Alternatively, could it be the result of phonetic implementation?

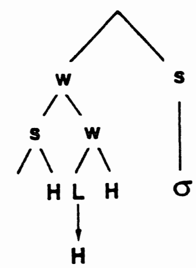
If the 2TS is a phonological assimilation rule, in which the L target of a rising tone assimilates to surrounding H targets, then two syllables in the rule description should be enough. Why should the presence of a third syllable be crucial, and yet the value of its tone have no effect? If the tone on the third syllable has no function, why can't it be a neutral tone?

Yip [3] and Zhang [4] both provide a stress account to explain the requirement of the third syllable. They both recognize that the real conditioning factor of the 2TS is stress, instead of syllable count. Their rules below specify that the L target of a Tone 2 will change to H only if the syllable is dominated by a weak (w) node. In other words, the 2TS only applies to prosodically weak syllables. Yip's rule does not explicitly refer to a third syllable as the conditioning factor; the rule suggests that the stress condition is all that is needed. Zhang's stress tree depicts the typical stress relation of a three syllable word in Mandarin: the middle syllable is weak, therefore allowing the 2TS to apply.

7. Yip's 2TS Rule:



8. Zhang's 2TS Rule:



Even though distinction of stress levels in a tone language such as Mandarin is subtle, it is generally agreed that a Mandarin word typically has final stress, and a three syllable structure has the stress pattern *secondary, tertiary, primary* [5, 6]. Hoa [5] has demonstrated the derivation of such stress relation in a left-branching word following the metrical structure and the rhythm rule of Liberman and Prince [7]. In a right-branching structure, similar stress relationship can be obtain by restructuring (Zhang [4], Hoa[5]), or by treating the three syllable structure in question as a single domain, ignoring internal branching (Shih [8]).

Interpreting stress as the conditioning factor of the 2TS explains successfully why the third syllable cannot have a neutral tone. Since neutral tone syllable are weak and cannot carry stress, its presence in the final position forces the middle syllable (being the last stress/tone carrying unit) to receive primary stress, therefore destroying the 2TS environment defined in Yip and Zhang.

A question left unanswered by Chao, Yip, and Zhang concerns the reason why a L target will change to H in a prosodically weak position. We believe that the answer lies in the way tonal targets are implemented in general. Turn to the second issue concerning Chao's description, whether the 2TS is a phonological rule, we will confirm with experimental data that 2TS is the consequence of phonetic implementation. We will show that the resulting tone 1 of the 2TS is a coincidence, so to speak, expected of an extremely weak L target in a high tone environment, where the strength of a L target is reflected in its distance away from the surrounding H tone environment. In H tone context, a weak L tone is implemented with higher

F0 value while a strong one has lower F0 value, as proposed in Pierrehumbert and Beckman [9] for Japanese.

Furthermore, we consider the 2TS to be part of a more general picture: a weak tone has less strength to break away from the interpolation line of surrounding strong targets of all types. The phenomenon is not limited to H L H sequences. Figure 1 shows the pitch track of the utterance *fan3-ying4 su4-du4*, "reaction speed", from our conversation data.

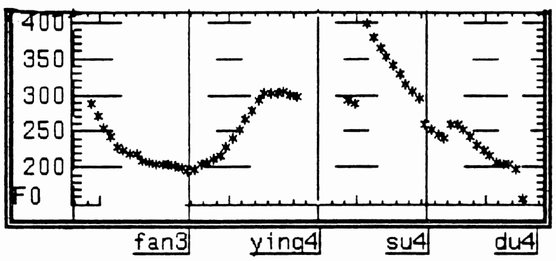


Figure 1: Pitch Track of *Fan3-Ying4 Su4-Du4*

The second syllable is in a weak position, and the falling tone (tone 4) surfaces with a rising shape. The surprising outcome is made possible when the syllable in question is so weak as to lose its own targets. The resulting pitch is just the transition from the previous syllable to the following syllable. In the pitch track shown, the preceding tone is low and the following tone is high, accounting for the transitional rising shape, which turns out to be the opposite of what is expected of the lexical tone. Since Mandarin speakers are typically sensitive to *wrong* tones, we play this speech segment to eight native speakers, asking them to identify the tone sequence, and ask them whether there is any thing wrong. All eight speakers identify the tone sequence as 3 4 4 4, the correct underlying tones. They didn't notice any aberrant behavior of the second syllable when the whole word was played to them, and were surprised later to hear a rising tone when the second syllable alone was played.

The following experiment is designed to investigate the two issues on the 2TS discussed above. First, we ask if the 2TS should be considered as a phonetic implementation rule rather than a phonological rule. Having answered this question in the affirmative, we will show that the phonetic implementation rule in question is sensitive to stress. We follow a straightforward working hypothesis of Liberman and Pierrehumbert [10] about the nature of phonological rule versus phonetic implementation: a phonological rule typically creates distinct classes on the surface, while phonetic implementation creates a continuum.

2. Experiment Design

We designed two sets of stimuli. The first, the control group, consisted of three and four syllable words with lexical high tones (tone 1) on every syllable. The second, the test group, were three or four syllable words with lexical high tones surrounding a single lexical rising tone (tone 2). These stimuli are described below:

Stimuli

I. Control group:

- 14 left branching three-syllable words with 1 1 1 tone pattern
- 9 right branching three-syllable words with 1 1 1 tone pattern
- 5 four-syllable words with 1 1 1 1 tone pattern

II. Test group:

- 22 left branching three-syllable words with 1 2 1 tone pattern
- 9 right branching three-syllable words with 1 2 1 tone pattern
- 9 four-syllable words with 1 2 1 1 tone pattern
- 6 four-syllable words with 1 1 2 1 tone pattern

In practice we found no effect of branching in the results for the three syllable words, so we will henceforth collapse the two conditions.

A female speaker of Beijing Mandarin was asked to produce utterances containing these words in the following contexts. Words with three and four high tones (1 1 1 and 1 1 1 1) are not used in reading contexts (IIb) and (IIc). Test words and sentences in each context are presented on cards to the speaker in random order. Contexts (IIb) and (IIc) are mixed. Each word/sentence is read once.

Contexts

I. Isolation

II. Read sentences with three different frames:

a. Ta1 shuo1 ____ chu1 mao2 bing4 le0.
"He says ___ went wrong."

b. Ta1 ke1 ABC de0 A zi4.
"He carves the character A of ABC"

c. Ta1 ke1 ABC de0 B zi4.
"He carves the character B of ABC"

III. Conversation

The isolated syllables *A* and *B* in contexts (IIb-c) are referred to as the *frame words*. The frame words are prosodically strong, carrying both the word stress and the sentential focus. In the case of 4 character words, the frame word B is the tone 2 syllable; it could be the second or third syllable of the word. The speaker pronounced the syllable *ke1* "carves" in (IIb-c) with a falling tone, which has noticeable lowering effect on the test words in (IIb-c) contexts, in comparison to the test words in the (IIa) context.

The conversation data are collected by having the speaker discuss the experiment with the experimenter after she had finished with the isolation and read sentence conditions for all words. Some of the typical questions addressed to the speaker include: "Do you like ___?" and "How often do you read ___?" All occurrences of words with the appropriate tonal combination produced by the speaker in the conversation are collected for analysis. Naturally, we were not able to elicit utterances for every stimulus word under these conditions. We have also collected a few words that are not used in the isolation and read sentence conditions.

One F0/time pair is measured for each syllable of the word for analysis. The pitch contours of most tone 1's and tone 2's in our data have an overall concave shape. Typically, we measure the F0 and time value of the F0 minimum for both types of tones. The lowest point of a tone 2 lies close to the center of the vowel region, and is considered to represent the L target of the rising tone. Occasionally, a tone 1 will have a convex shape; under those conditions, we measure the time and F0 value for the F0 maximum.

3. Results

Figure (2) shows three occurrences of the same word *zang1 mao2-yi1*, "dirty sweater", all collected from the conversation session. These pitch tracks show quite a variation in the implementation of the L target of the rising tone. Figure (2a) shows a pitch drop of around 50 Hz from the H target of *zang1* to the L target of *mao2*. The pitch drop in (2b) is considerably less, at around 25 Hz. Finally, there is hardly any pitch drop in (2c). The L target of a rising tone is realized at around the same pitch height of the surrounding H targets.

If the 2TS were a phonological rule, we would expect most of the samples to be similar to either Figure (2a), where the rule has not applied (note that the rule is optional), or to Figure (2c), where the rule has applied. Cases like Figure (2b) are not expected, so the frequency should be low, if they occur at all.

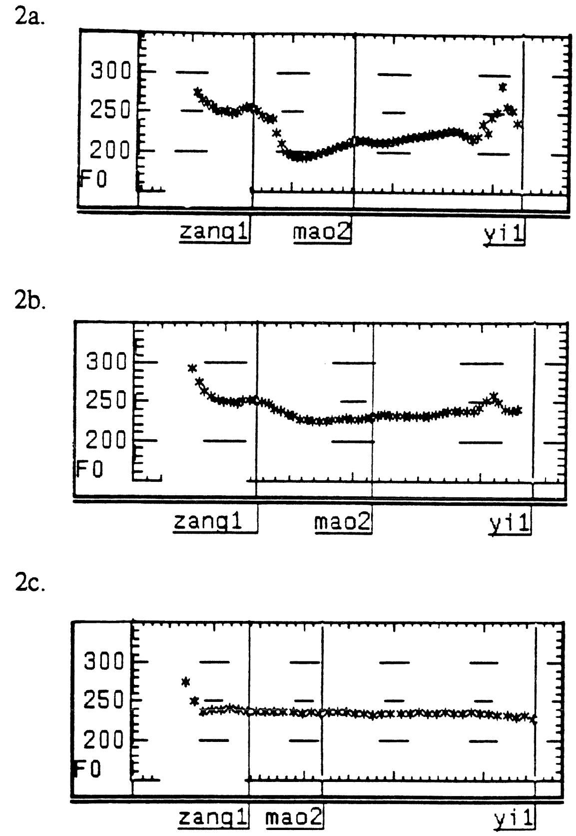


Figure 2: Variations in Pitch Drop

In the following, we will show that the amount of pitch drop is indeed gradient, and cases like (2b) are the norm, not the exception, supporting the view that the 2TS is the result of phonetic implementation. It is reasonable to ask whether the variations in pitch drop can be derived from duration: if the duration is short, there may not be sufficient time to reach down to a L target. We cannot find clear evidence supporting this hypothesis. Finally, we test whether stress is a factor, and find strong support in the sense that the amount of pitch drop correlates well with the phonologically defined stress levels.

3.1 Gradient Implementation

Figure 3 shows averaged tonal trajectories for all 1 1 1 and 1 2 1 words divided by condition. The 1 1 1 cases are represented by black plotting symbols, while the 1 2 1 cases are represented by open symbols; for example, the black triangle trajectory represents the average of all 1 1 1 words in the "Read Sentence" condition, and shows that the first syllable's high tone has a mean pitch value of 304Hz, the second syllable's high tone has a mean pitch value of 288Hz and the third syllable's high tone has a mean pitch value of 280Hz. The 1 1 1 words, with a string of H tones, show very similar behavior under all

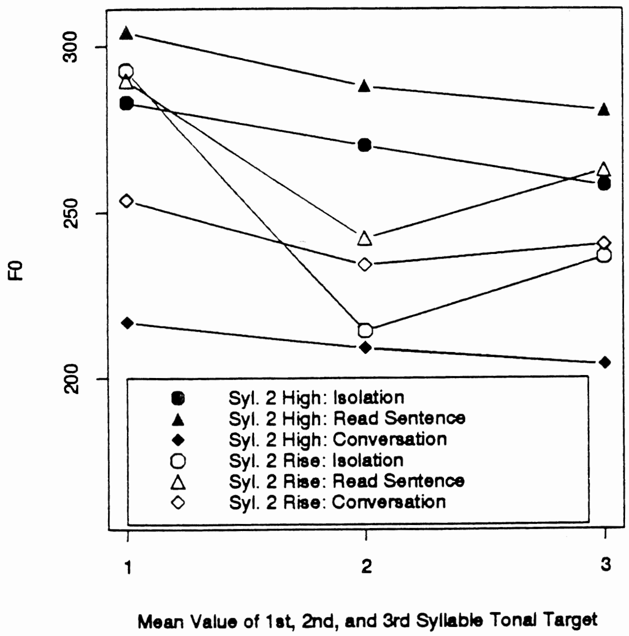


Figure 3: Tonal Trajectories of 3 Syllable Words

conditions: roughly speaking there is a drop of 15 to 25 Hz between the first and the third syllables, but otherwise the decline is almost linear. We assume that the significantly lower pitch value in the "Conversational" condition is due to many of these utterances having occurred fairly late in their contextual sentences, so that what we are seeing here is due to the accumulation of lowering effects such as declination and downstep.

The 1 2 1 words show a very different pattern. In the isolation context, represented by open circles, there is a drop of more than 75 Hz between the first syllable's tone 1 (H) and the second syllable's L tone. The third syllable's tone 1 only rises about 20 Hz, possibly due to the combined effects of downstep and final lowering. (See [10] for a discussion of these effects and their interaction.) In the read sentence context the drop between the first two syllables is less but still significant (47 Hz); since these words are not utterance-final, the final lowering effect is removed and the third syllable's tone 1 has a much higher value than in the isolation context. Most striking, however, is the conversational context where the drop between syllable one and two is minuscule (less than 20 Hz), and the overall pattern for the three syllables strongly resembles a sequence of three high tones. So, in the 1 2 1 contexts, we have what appears to be a gradient effect from the "Isolation" context, through the "Read Sentence" context, through the "Conversation" context.

That this behavior is indeed gradient is shown by the plot in Figure 4. Here we plot the F0 value for syllable 1 on the x-axis against the F0 value for syllable 2 on the y-axis, for both 1 1 1 and 1 2 1 words under all 3 conditions.

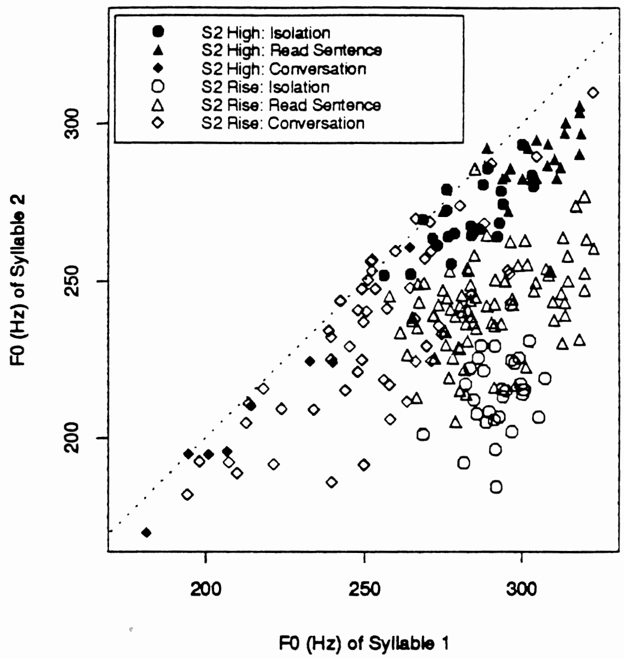


Figure 4: F0 of Syllable 1 vs. Syllable 2

The plotting symbols are the same as the previous plot: black symbols for the 1 1 1 cases; open symbols for the 1 2 1 cases; circles for the "Isolation" context; triangles for the "Read Sentence" context; and diamonds for the "Conversation" context.

The dotted line is the line x=y, and the fact that most points lie below that line, even for the 1 1 1 cases, is consistent with the overall declination effect (i.e., syllable 2 generally has a slightly lower pitch value than syllable 1, even if both are high tones). As can be seen in this plot all of the high tone cases cluster fairly close to the x=y line, as expected. The 1 2 1 cases, however, range from having syllable 2's F0 much lower than syllable 1's F0 -- e.g., the "Isolation" cases in open circles -- all the way to having the two syllables have roughly equal tone values -- e.g., many of the "Conversation" cases in open diamonds -- and thus being effectively indistinguishable from a sequence of three lexical high tones.

The gradient implementation of L target can be further supported by performing t-tests comparing the pitch difference between syllable 1 and syllable 2 for, on the one hand, all of the 1 1 1 cases, and on the other each of "Isolation", "Read Sentence" and "Conversation" for the 1 2 1 cases. These results are given in Table 2.

t-test: all 111 vs. 121		
	t(df)	p
121 Isolation	28.2 (82)	< .001
121 Read Sentence	13.4 (136)	< .001
121 Conversation	2.2 (105)	< .05

Table 2: T-Test Scores

For both the "Isolation" and "Read Sentence" cases we can reject the hypothesis that they are the same as the 1 1 1 cases, at the .001 level; note however that the actual t score for "Read Sentence" is less than that for the "Isolation", consistent with the observation that the "Read Sentence" cases have a smaller drop between syllable 1 and syllable 2 than the "Isolation" cases. For the "Conversation" cases, there is still a significant difference between them and the 1 1 1 set, but only at the .05 level. This is interesting for two reasons. First of all, it suggests that under conversational conditions the difference between the 1 2 1 and 1 1 1 cases is indeed reduced, consistent with Chao's rule. On the other hand, the fact that even under conversational conditions the 1 2 1 set cannot be treated as identical to the 1 1 1 set suggests that we must be dealing here with a gradient phonetic implementation rule, rather than a categorical phonological rule whose effect is to rewrite a tone 2 as a tone 1.

3.2 Duration

Given the rather strong evidence that the 2TS must therefore be a phonetic implementation rule, the next question is what the conditioning factors might be. One obvious possibility is speech rate or durational effects. When the duration is short, the speaker may not have enough time to reach the intended target, leading to the reduced pitch difference between a L target and the preceding H tone, and the apparent assimilation effect on the surface. To test this hypothesis, we checked the amount of pitch drop in 1 2 1 words against a number of duration measurements, such as the duration of each segment, the duration of each syllable, the duration of two syllables, and the duration of the whole word. We show one of these plots in Figure 5, the whole word duration against the amount of pitch drop. The correlation is among the best, and the pattern of the distribution is very representative of the cases where other duration measurements are used.

In Figure 5, the duration (in seconds) of 3-syllable 1 2 1 words is plotted on the x-axis, and the amount of pitch drop (in Hz) from the first syllable (H target) to the second syllable (L target) is plotted on the y-axis. Open circle, black circle, and open diamond represent the "Isolation" context, the "Read Sentence" context, and the "Conversation" context respectively. The dotted line represents the least-squares fit regression line for all the

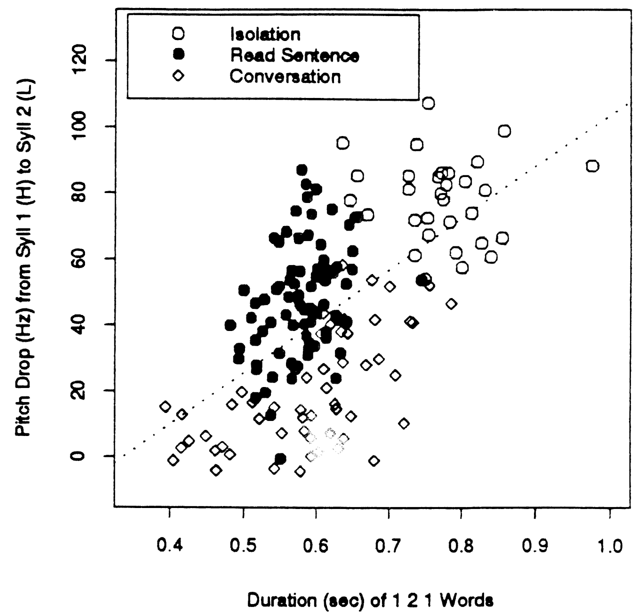


Figure 5: Duration vs. Pitch Drop in 1 2 1 Words

data.

As can be seen from the plot, there is some correlation between the duration variable and the pitch variable. For example, the "Isolation" context tend to be associated with longer duration (as might be expected) and they are also the cases with the largest pitch drop; contrariwise, some of the conversational cases have shorter durations and they also show the least pitch drop. However, the correlation is not good: the R-squared value for this correlation is only 0.36 (sample size is 170).

A striking feature of Figure 5 is the separation of the three contexts: On the one hand, the "Isolation" context do not overlap much in their time-axis distribution with the "Read Sentence" and "Conversation" contexts, but it overlaps considerably along the pitch-axis with the "Read Sentence" context. On the other hand, the "Read Sentence" context and the "Conversation" context show a very significant overlap in the time-axis, but are clearly separated in the pitch-axis. The complication of styles or test contexts makes it more difficult to generalize the relation between duration and pitch drop.

More importantly, there is no correlation between duration and pitch drop within the "Isolation" context, and the correlation is poor for the "Read Sentences" context (R-squared value is 0.13, sample size is 85), suggesting that the amount of pitch drop cannot be reliably predicted from duration for members within each of these two groups. The "Conversation" context has better correlation (R-squared value is 0.36, sample size is 54), but it is also clear that the distribution is triangle shaped: samples with short duration, especially those shorter than 0.5 second, have less than 20 Hz of pitch drop, but many samples with much longer duration still fall within the same range. It seems that only the extremely short duration is responsible for the undershoot of the target, represented by the 10 samples at the lower left corner of the plot. For the rest of the data, although duration or speech rate is surely a factor, there appears at the present time to be an additional irreducible effect of speech style in the phonetic implementation of the 2TS.

3.3 Stress

Another possible factor is stress. Indeed, we have hypothesized with Yip and Zhang that stress is the real reason behind Chao's "three syllable condition": only when a tone 2 occurs after a H target on a preceding syllable, and before some other non-neutral-toned syllable can it both be in the right tonal environment AND be in a metrically weak position. In order to evaluate the effect of stress, we turn now to the four-syllable cases. In Mandarin, in a four syllable sequence ABCD, the B syllable is generally the weakest prosodically (see summary in Zhang [4]). Thus, if stress is a factor in the implementation of 2TS, we would expect to find difference in behavior between the tone 2's in 1 2 1 1, where the tone 2 is on the weakest syllable, and 1 1 2 1, where it is on a stronger syllable. Assuming that stronger stress correlates with a more pronounced rendition of a tonal target, we would expect the tone 2 in 1 2 1 1 to show more 2TS effects than 1 1 2 1. This expectation is confirmed. In Figure 6 we plot data from "Conversation" and "Read Sentences" contexts the F0 value for the syllable preceding the tone 2 against the F0 value for the tone 2 itself. The 1 2 1 1 cases are represented by open triangle, and the 1 1 2 1 cases by black circle. The dash line is the x=y line. As can be seen, the value for tone 2 is significantly further away from the x=y line for the 1 1 2 1 cases than for the 1 2 1 1 cases, consistent with the stronger stress for the tone 2 in 1 1 2 1.

Both cases presented in Figure 6 are relatively weak, comparing to some strong positions that are expected to receive primary word stress, such as a monosyllabic word, the tone-carrying final syllable of a word, or the syllable before a neutral tone. If stress is a factor in the implementation of the 2TS, we would expect even more pitch drop in the strong cases then in 1 1 2 1. This prediction is again bourn out. The difference between the 1 1 2 1 case and other cases with even stronger stress is a good evidence against the phonological analysis of Zhang [4], where it is claimed that 2TS does not apply in the 1 1 2 1 cases because there is some level of stress dominating the tone 2 syllable.

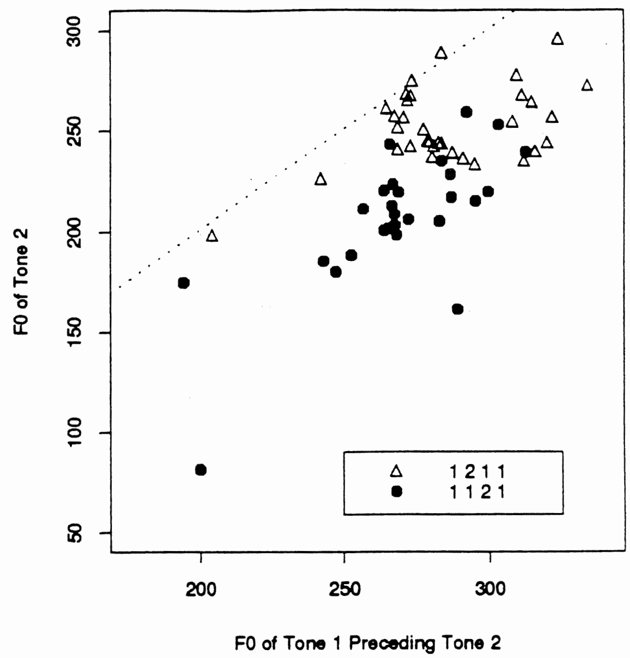


Figure 6: Pitch Drop in 1211 and 1121

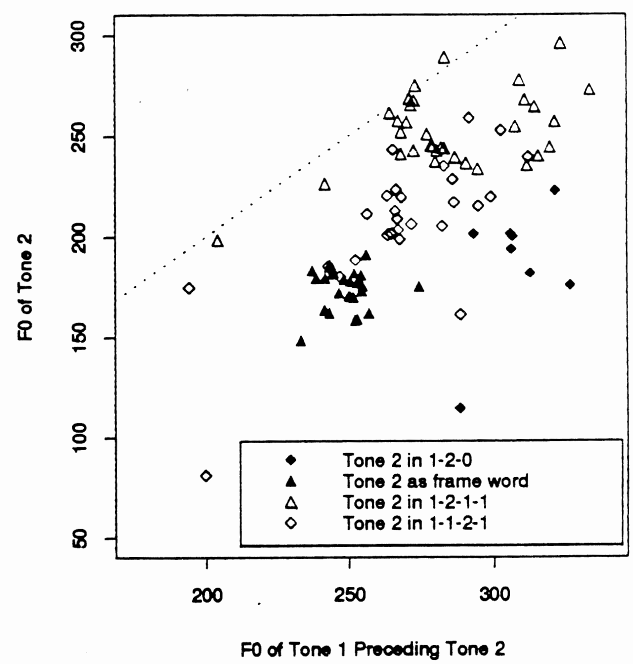


Figure 7: Pitch Drop in Weak Positions (1211, 1121) and Strong Positions (120, 2)

Figure 7 adds two other cases where tone 2 is expected to be strong (in black), contrasting to the tokens in Figure 6 (in white). One of the strong cases we used is the tone 2 in tonal combination 1 2 0, where the third syllable has a neutral tone (tone 0), and cannot receive stress. The middle tone 2 will be assigned primary stress instead because it is now the last tone/stress carrying member of the word. In this case we plot the value of the preceding high tone (x-axis) against the value of the tone 2 (y-axis) in black diamonds. Another presumably strong case is tone 2 of a monosyllabic word. We plot the value of tone 1's (on x-axis) from the frame word in Read sentence condition (IIb) against the tone 2 (on y-axis) from the frame word in Read sentence condition (IIc). (Given a stimulus word ABC, with tonal combination 1 2 1, we plot A (tone 1) from the sentence "**A** of ABC" against B (tone 2) from the sentence "**B** of ABC".) Black triangle is used as the plotting symbol. It is clear that all of the expected strong tone 2's have comparatively lower F0 value for their L targets.

Figure 7 gives strong support to the notion that the value of L target in H tone context is sensitive to stress: L target in stronger positions is implemented with lower F0 value, with bigger excursion away from the surrounding environment; L target in weaker positions is implemented with higher F0 value, which is more similar to the H tone environment.

4. Conclusion

To summarize, we have shown that Chao's 2TS rule is best viewed as a phonetic implementation rule that raises the pitch value of the L excursion in a tone 2 when that tone 2 occurs after a H tone. The tone 2 must be followed by a non-neutral toned syllable, which provides the minimum environment in which the tone 2 can be metrically weak. We have shown that stress is indeed an important factor by showing a robust difference among the implementation of the tone 2 in 1 2 1 1, 1 1 2 1, 2, and 1 2 0 environments. Moreover, in both 1 2 1 1 and 1 1 2 1 cases the tone 2 is in a metrically relatively weak syllable, but in 1 1 2 1 it is nonetheless in a somewhat stronger position than in 1 2 1 1. Tonal implementation is sensitive to the difference. Note that this difference does not follow from Chao's statement of the rule, and that the stress condition thus has the nice property of both explaining an odd feature of Chao's rule (the three syllable condition) and actually offering a better account of the data than his rule. We have also tested whether duration or speech rate may be a factor in the implementation of the L target, and show that there are some problem in interpreting the correlation of duration and the amount of pitch drop. Furthermore, consistent with Chao's original description, there seems also to be an irreducible effect of style.

5. References

[1] Chao, Y. R. *A grammar of spoken Chinese*. Berkeley: University of Califomia Press, 1968.

[2] Cheng, C. C. *A synchronic phonology of Mandarin Chinese*. The Hague: Mouton, 1973.

[3] Yip, M. *The tonal phonology of Chinese*. Ph.D. dissertation, MIT, 1980.

[4] Zhang, Z. S. *Tone and tone sandhi in Chinese*. Ph.D. dissertation, Ohio State University, 1988.

[5] Hoa, M. *L'accentuation en Pekinois*. Paris: Centre de Recherches Linguistiques sur l'Asie Orientale, Editions Langages Croises, 1983.

[6] Lin, M., Yan, J. and G. Sun. The stress pattern and its acoustic correlates in Beijing Mandarin. pp. 504-514, in *Proceedings of the 10th International Congress of Phonetic Sciences*, 1983.

[7] Liberman M. and A. Prince. On stress and linguistic rhythm, pp. 249-336, in *Linguistic Inquiry* 8.2, 1977.

[8] Shih, C. L. Mandarin third tone sandhi and prosodic structure. In. J. Wang and N. Smith eds, *Studies in Chinese phonology*, Foris Publication, in press.

[9] Pierrehumbert J. and M. Beckman. *Japanese tone structure*, Cambridge: MIT Press, 1988.

[10] Liberman M. and J. Pierrehumbert. Intonational invariance under changes in pitch range and length, pp. 157-233, in M. Aronoff and R. Oehrle eds. *Language Sound Structure*, Cambridge: MIT Press, 1984.