OSU #36: Papers from the Linguistics Laboratory 1985-1987, Sept. 1987.

A Study of Toishan FO

Gina Lee
Linguistics, Ohio State University

Abstract: Like other Chinese languages, Toishan uses tone to signal differences in meaning among words. With the exception of certain morphologically conditioned tone changes, which must be memorized by speakers, Toishan exhibits no tonal modifications; in particular, there appear to be no tonal modifications which are strictly phonologically governed. Given the absence of such changes, Toishan provides an ideal situation for examining the hypothesis that declination--a phonologically unmotivated lowering of FO, independent of tone--is a relevant component in a model of intonation. Three native speakers read paragraphs containing five pragmatically connected sentences of different length. Two target tones within each sentence were measured. No evidence for declination was found. The FO pattern is described in terms of an overall structure involving pitch range expansion for initial sentences and compression (or lowering) for final sentences. The results are discussed in light of evidence from other languages and implications for a model of local FO implementation.

1. Introduction

1.1. Downtrend and declination

In examining the properties of intonation, many have noted a general tendency for FO to gradually lower during the production of a syntactic unit. Such lowering of the intonation contour is believed by some to be universal. This phenomenon is known by a variety of terms; following Liberman and Pierrehumbert (1984), I will refer to the general lowering of FO as downtrend.

In some cases, downtrend may be the result of a phonologically conditioned rule. For instance, phonologically identical high tones within an utterance are sometimes not phonetically identical. In many African tone languages, each non-initial high tone in the sequence HLHLHLH is produced with lower fundamental frequency relative to the preceding high tone, due to a phonetic rule which lowers high tones after a preceding low tone. Such lowering is a special case of downtrend known as downdrift.

Recent work has proposed that in some situations FO lowering may be due to declination: a gradually declining phonetic frame of reference that is independent of tonal context (e.g. Pierrehumbert 1980). Various physiological explanations have been proposed. Among them is the argument that subglottal pressure decreases toward the end of a syntactic unit, and that--everything else being equal--this decrease corresponds to a decrease in FO. Although there is some doubt as to whether subglottal pressure actually decreases during the production of an utterance, and although other physiological factors may also contribute to FO lowering, it is

generally agreed that a decrease in subglottal pressure provides a clear physiological motivation for F0 lowering.

Unlike downdrift, declination is not dependent on specific phonological characteristics of the utterance but, rather, is dependent on time. Figures 1 (a) and (b) illustrate the relationship between utterance length and a declining F0 backdrop. Points x and y represent tones in an F0 contour. Distance d represents the difference between the heights of x and y. In 1 (b), the time interval separating x and y is longer relative to that in 1 (a). Because the pitch range falls to a greater extent in 1 (b), y is lower relative to x, i.e. d is larger in 1 (b). Thus, if declination is present in a set of utterances, and assuming no phonetic rules which might otherwise affect the F0, the difference between the F0 values of two points within a sentence is directly proportional to the length of the interval between the two points.

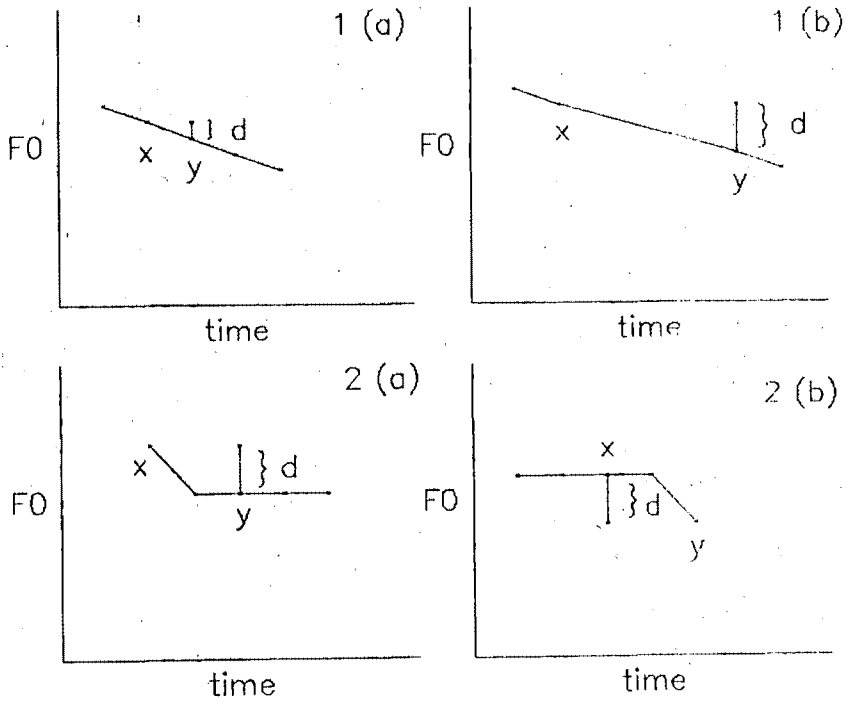

Figure 1: Declination as a function of time, and the consequential differences between the heights of points x and y, based on (a) short interval and (b) long interval lengths. Figure 2: Differences in height due to boundary effects of (a) initial raising and (b) final lowering, in each case without declination.

However, there is some question as to whether a declination component is necessary in a model of intonation. It may be the case that what appears to be a declination effect is actually the effects of initial raising and final lowering rules, particularly in short utterances. In 2 (a), the speaker raises the initial portion of the pitch range without applying declination. Since point x is higher, d is larger than it would be if x occurred later in the utterance. Final lowering, as described by Liberman and Pierrehumbert (1984), is a lowering and compression of the pitch range resulting from a speaker's anticipation of the end of the utterance. 2 (b) shows the effect of final lowering without declination. Point y is lower, and hence d is larger than it would be if y occurred earlier (relative to the end). If d can be accounted for exclusively in terms of such boundary effects--and if utterance length is not a significant factor--then it is unnecessary to posit a declination component.

In theory, the lack of tone sandhi in a language whose speakers exhibit downtrend should make it rather straightforward to design an experiment which tests for declination. And since there may be a physiological cause for declination, one would expect to find declination in such cases. The language under consideration is Toishan, a variety of Chinese spoken primarily in the Guangdong province of southern China, as well as Hong Kong.[1] The Columbus area has a small Toishan speaking community, consisting mostly of older immigrants. Toishan is a general term for the Chinese spoken in the Sze-yap ('four districts') region of Guangdong, consisting of Hoiping, Sunwui, Yanping and Toishan District. The Sze-yap varieties are an important part of the Yue "dialects". Of all the Yue "dialects", the best known is Cantonese.

In the varieties of Toishan described by McCoy (1969) and Cheng (1973), there are five underlying tones, three level and two falling, which I will indicate by a modified version of Cheng's notation: high level (55), high falling (52) (sometimes analyzed as mid falling (31)), mid level (33), low level (22), and low falling (21). In addition to the five basic tones, there are four "derived" tones which are strictly morphologically/syntactically motivated. These modifications are known traditionally as tone changes. They are similar to the more widely known Cantonese tone changes (for a detailed discussion of Cantonese and Toishan tone changes, see Wong 1982). For example, in some Toishan compounds, 33 tones become 21, regardless of the nature of the preceding tone, as in [kap] 33 'clip', [pɔ] 33 [tʃi] 55 [kap] 21 'newspaper clip'. Most tone changes occur in noun phrases or in verbs derived from nouns. Significantly, there are no modifications in tone that are strictly phonologically motivated, and there is no evidence for downdrift.

---

[1] This language is also known as Taishan(-ese). In the language itself it is called Hoisan-wa. Following Light (1981), who follows an old postal usage, and also to avoid confusion with phonetically similar names, I will use the name Toishan.

1.2. F0 implementation: global vs. local

In addition to the issue of declination, there is also the issue of the relationship between global trends and local events. In Thorsen's hierarchical representation of intonation, global patterns are direct, independent representations of a higher level component, such as the textual, sentence or phrase level. Using evidence from Standard Danish, Thorsen (1985) found that sentences in initial, medial and final positions within a text were differentiated by a sequential, evenly distributed declination. Lindau (1986) proposes a similar "layered" representation for Hausa, where the F0 contour is described in terms of an interaction between global and local factors.

In contrast, Liberman and Pierrehumbert (1984) argue that F0 is implemented only by means of locally triggered rules (a linear (or tonal) sequence view): " . . . we find no clear evidence for phrase level planning . . . The major factors shaping F0 contours appear to be local ones" (p. 166). In this model, the global trend is merely an artifact of the application of such local rules. A second point of interest (though not considered in this paper) involves the different claims these two views make concerning preplanning. In the "global" view, the implementation of F0 is by necessity preplanned. In the "local" view, preplanning is limited to phrase-level changes in pitch range specification (which has a direct effect only on the first phrasal pitch accent) and final lowering. The Toishan study may provide evidence for the preference of one view over the other.

2. Methods

The study involved three native speakers. Subject GFL was seventy-four years old at the time of the elicitation; he was born in Wailung Village, Sunwui District. He first came to the U. S. in 1927 before settling in Columbus after marriage. FKL, a sixty-seven year old female, was born in Taichong Village, Hoiping District, and came to the U. S. in approximately 1940. THY, born in the same village as FKL, was fifty-seven years old. He lived in Hong Kong before coming to the United States in 1982. GFL speaks fluent English; FKL understands some English, but does not speak it fluently. THY speaks little English. All three understand and can speak some Cantonese. None had any formal education beyond the grade school level.

Each read aloud a text consisting of four paragraphs. Each paragraph contained five pragmatically connected sentences. The story was about an elderly man and his two pets, a dog and a monkey. The sentences were neutral declarative sentences. For example, the first paragraph was: "Lɔ pak yu kaw toy. Kaw toy hay ɬay koy kaw toy. Kaw toy hay lɛŋ ɬuy. Lɔ pak tʃoŋyi həŋ kaw fan. Kaw toy han lɔ pak." (Which roughly translates as: 'Old Uncle has a puppy. The puppy is a small puppy. The puppy is two years old. Old Uncle likes to play with the dog. The puppy loves Old Uncle.') The corpus is given in its entirety in the appendix.

Each sentence had two target tones. The first target tone, which I will designate as t1, was always sentence initial; the second, t2, was

always in penultimate position. Since tone changes do not affect underlying high level tones in Toishan, all target tones were high level.

The intervening amount of material between the target tones was either two syllables--considered a "short" interval, or four syllables--a "long" interval. To test for declination, one would need to see if the difference between the F0 values in a sentence such as "Lɔ pak yu kaw toy"--where lɔ and kaw are the test syllables--is smaller than in "Ma law yet ɬɛŋ lam lɔ pak"--where a longer interval separates the test syllables ma and lɔ.

Paragraphs 1 and 3 were similar in structure, in that the interval between each target tone alternated between two and four syllables, with the pattern 2-4-2-4-2. Paragraphs 2 and 4 contained the reverse pattern, 4-2-4-2-4.

Since a minimal amount of background noise on the recordings does not affect a program's capability to track F0, I was able to record GFL and FKL at their home, and THY in a relatively quiet area at work. A potential advantage of this was that the subjects are as comfortable as possible, which may result in more natural readings. I recorded GFL and FKL using their SONY cassette deck; each wore a SONY ECM-16 clip-on microphone. THY was recorded with a portable cassette recorder containing a built-in microphone.

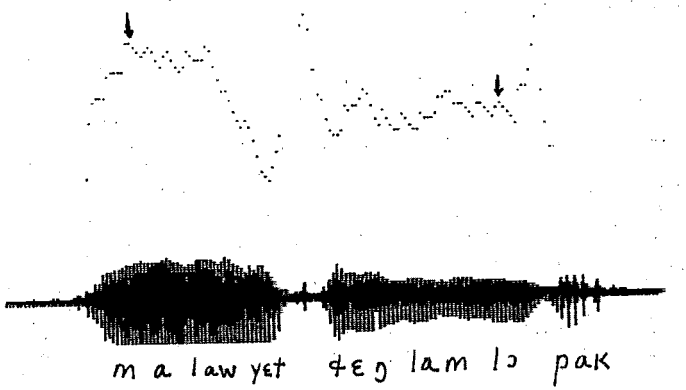

Figure 3: A sample F0 contour for speaker GFL. The arrows indicate measurement points.

Each speaker had a brief run-through session before the recording began. The subjects read each paragraph several times before going on to the next one. This appeared to be the best strategy (as opposed to reading the text in its entirety) because they were not fluent at reading aloud. They made numerous errors in reading (mainly by way of anticipatory mistakes), and reading by paragraph was the best way of obtaining good readings. The number of repetitions of each paragraph varied from paragraph to paragraph and from speaker to speaker; I simply asked them to repeat the paragraph until they had read it smoothly several times. FKL

had a minimum of four usable repetitions. GFL and THY's reading went more smoothly than FKL's; they each had a minimum of seven usable repetitions. In all, a total of nearly 400 tokens were analyzed.

The recordings were digitized at a sampling rate of 10000 Hz on a DEC PDP 11/23, using the ILS signal processing program. Using LPC analysis, I obtained the fundamental frequency contour for each sentence. Figure 3 shows how the measurements of the target tones were made. The target tone value was the first peak F0 after the onset of the vowel. The initials consisted of both sonorants and obstruents. Since obstruents affect the F0 of adjacent vowels, the measurements were taken far enough into the vowel to eliminate such segmental effects.

3. Results

3.1. Effects of interval and position

For all three speakers, the ratio of t1/t2 is, in general, slightly greater than 1.00. Thus, the value of t2 was usually lower than t1. All three speakers showed a small amount of downtrend which was significant at the .001 level. However, the downtrend cannot be attributed to declination. The results of a two-way ANOVA indicate that the difference between the short interval sentences and the long interval sentences was not significant at the .05 level for any of the three speakers. This is shown in Figure 4, where the t1/t2 ratio (y-axis) is plotted against sentence position and interval length (x-axis). The lower the value of t2 is relative to t1, the larger the t1/t2 ratio. The open circles represent short interval sentences, the filled circles, long interval.

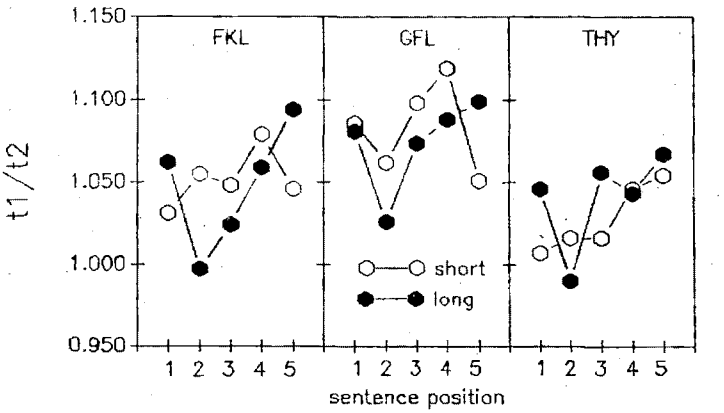

Figure 4: Average t1/t2 ratios (y-axis) according to both interval length and position within the paragraph (x-axis) for all three speakers. Open circles indicate short interval sentences; filled circles, long interval. If a declination effect is present, the filled circles should be consistently higher than the open circles. Hoever, none of the speakers show declination.

If declination is present, the filled circles should be consistently higher than the open circles (i.e. the ratios should have a relatively lower t2). This is clearly not the case for any of the three speakers. For GFL, the opposite occurs: the short interval sentences have a larger rate of downtrend in sentences 2 through 4, and roughly the same in sentence 1. FKL has a larger long interval downtrend only for positions 1 and 5, the peripheral sentences. THY likewise shows no declination effect.

The possibility of an interaction between position and interval was examined as well, where position refers to the order of each sentence within the paragraph. Perhaps early sentences with a short interval may show a smaller amount of downtrend than later sentences with a long interval. Across speakers, there was an unusual pattern involving a "dip" in the long interval sentences from sentence 1 to sentence 2, then a gradually increasing ratio from sentences 2 through 5, meaning that t2 became lower relative to t1. This is not true for the short interval sentences. This pattern, however, does not appear to be important. (A possible explanation is given below.) For all speakers, the interaction of position and interval was significant at the .05 level, but not at the .01 level.

Thus, there is no pattern of declination. But it would be interesting to see if the interval factor is significant in cases where the "long" interval is longer than four syllables--perhaps comparing two vs. eight syllable intervals. There is the possibility that declination may not have had an obvious effect in the current study because the difference between the syllable intervals was not large enough. The sentences were relatively short sentences because I did not want to make the reading difficult for the subjects. I had also hoped that they might with practice memorize the paragraphs and speak more naturally. (Because they made many errors, I did not attempt to have them memorize the sentences.)

3.2. A representation of F0 implementation

3.2.1. Comparison with the model of Pierrehumbert et al.

The role of sentence position is more obvious when considering the averages of the target tones in each position and interval, as shown in Figure 5. For all three speakers, F0 height for t1 in position 1 was higher relative to t1 in sentence position 2. The average magnitude of difference was roughly the same across speakers (FKL 6.903, GFL 8.6255, THY 8.929 Hz). For GFL and THY, t1 in position 1 is higher than t1 in the other four positions; in FKL's case it is higher than t1 only in positions 2, 3, and 5. In general the pitch range was expanded for the first sentence.

FKL's case is unusual because she expands the pitch range for sentence 4. The average value of t1 in sentence 4 is very close to that of t1 in sentence 1, showing a "resetting" of the pitch range. There was no semantic reason for her to do this. However, as noted earlier, FKL had difficulty reading the text, more so than the other speakers. The best explanation is that she split the text in half to make the reading easier.

In the production of sentence 5, GFL lowers and compresses the pitch range. t1 in position 5 is lower relative to positions 3 and 4, and the difference is significant at the .01 level. t2 in those positions remains the same. In the production of sentence 5, THY lowers the pitch range. t1 in position 5 remains the same, but t2 is lower relative to 3 and 4 (although the difference is not significant).

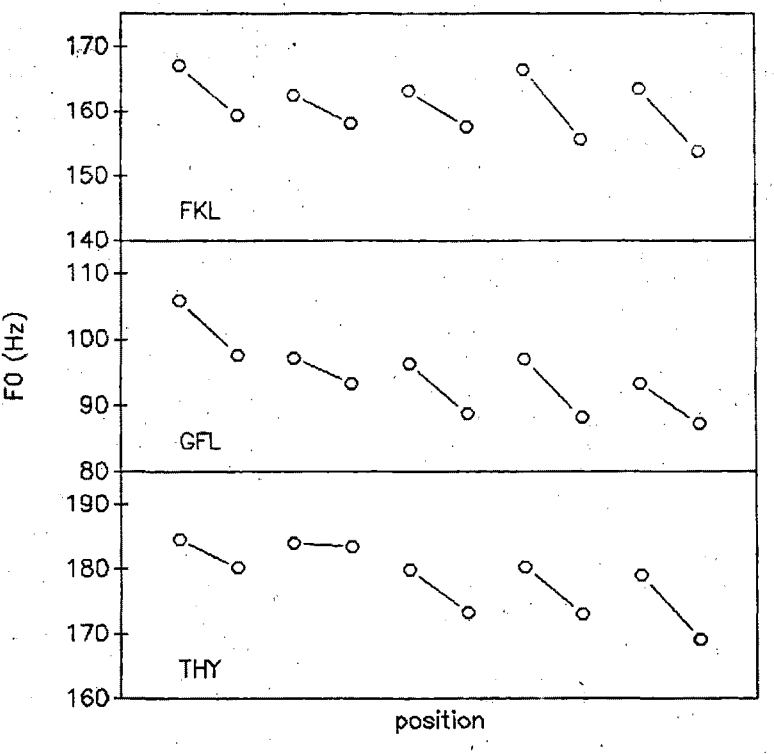

Figure 5: Averages of t1 and t2 by position for each speaker. The first connected pair represents t1 and t2 in sentence 1, and so on through t1 and t2 in sentence 5. The pitch range is expanded for the first sentence (which initiates the topic) and is compressed or lowered for the last sentence (which concludes the topic).

These results are consistent with Hirschberg and Pierrehumbert's descriptions of pitch range variations as pragmatic markers (1986). The position of the sentence within the discourse unit is signalled by its intonation. An initiation in topic is reflected by an increase in pitch range. The larger the change in topic, the larger the increase in the degree of expansion. End sentences are marked as well; if a sentence completes a topic, the pitch range is lowered and compressed to a relatively greater degree than for non-terminal sentences. The degree of

final lowering depends on the position of the sentence relative to the discourse structure.

3.2.2. Comparison with Thorsen's model

A second point which is brought out when looking at tonal averages involves Thorsen's idea of a progressive lowering in the textual contour. Thorsen (1985) found that, at the sentence level, later accents were lower than earlier accents. At the discourse level, accents in the second sentence were lower than the corresponding accents in the first sentence; in three sentence texts, accents in the third sentence were likewise lower than the corresponding accents in the second. Thorsen proposes a "nesting" of declination backdrops, whereby individual sentences have a declination slope which is superimposed upon a textual declination slope. However, Thorsen's 1985 study involved texts containing at most three sentences. Thus, the overall declination might also be due to effects of raising for the first sentence and lowering for the last sentence, where the pitch range was expanded to the greatest degree for the first sentence and compressed to the greatest degree for the last sentence.

This raises the question of whether texts containing more than three sentences would show a sequential lowering. At the time, Thorsen acknowledged that ". . . two or more medial sentences or clauses may not be further differentiated among themselves". In a follow-up study, Thorsen (1986) examined sequences of four sentences and found that the overall slope was not progressively lowered. The second and third sentences were generally not distinguished from one another. However, as discussed below, this was more true for the longer sentences than for the shorter sentences.

The evidence from Toishan also indicates that paragraphs containing more than three sentences do not show a progressive declination. In GFL and THY's data, sentences 3 and 4 are not differentiated. T-tests showed no significant difference between the onset and offset of position 3 versus those of position 4. The first three sentences follow a progressively downward slope, however. Sentence 2 has a slope different from that of 3 and 4 as well as a higher offset. The difference between t2 in position 2 and t2 in positions 3 and 4 was significant at the .01 level for both speakers.

However, the difference between position 2 and positions 3 and 4 could be explained in terms of experimental design. The vowels in the target syllables were not always of the same height within each sentence. The syllables containing target tones had mid and low vowels (i.e. [ɔ], [a], and [ɑw]). The measurement in [aw] was made at a point in the [a] portion, and thus [aw] was considered equivalent to [a]. In cases where a mid vowel (usually [ɔ]) occurred in t1 and a low vowel (always [a]) in t2, there is the possibility that F0 might decrease for t2, resulting in a higher ratio. Conversely, a low-mid pattern might cause F0 to increase for t2, resulting in an intrinsically lower ratio. It essentially boils down to a difference between [ɔ] and [a].

In all positions except for position 2, there was a mixture of three patterns (mid-low, low-mid and same vowel), with half of the tokens containing one pattern and one-fourth containing each of the other two.

Position 2 differs from the others in that half had the same vowel, while the other half had a low-mid pattern. This suggests that the value of t2 may be intrinsically higher here than in the other positions, and that is precisely the pattern found in GFL and THY's cases. This might also explain the drop in the t1/t2 ratio for position 2 in the long interval sentences. Ideally the vowels should be of the same height. However, the extent to which vowel height affects F0 in the case of [ɔ] and [a] is not clear--especially in this case, where both vowels are close in height.

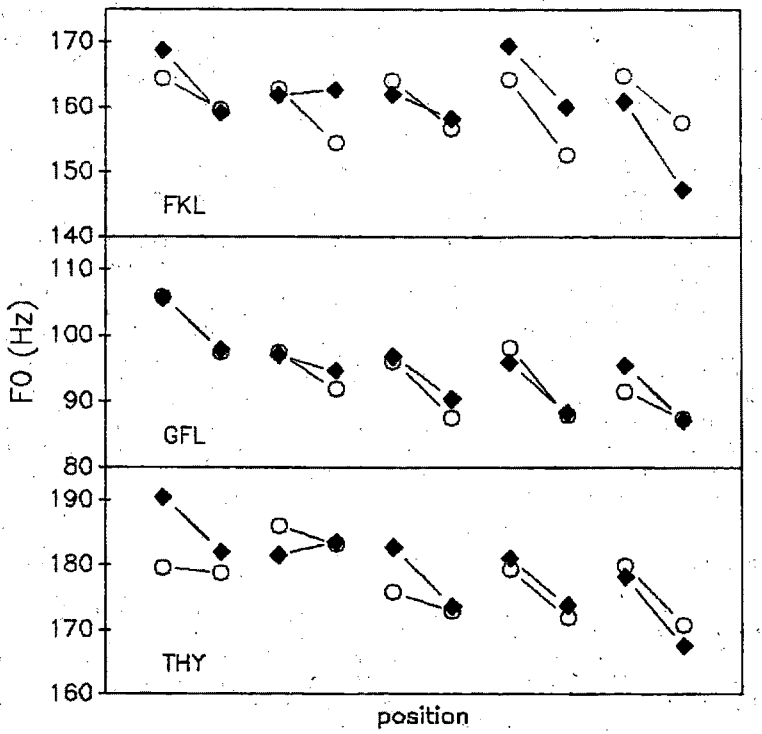

Figure 6: Averages of t1 and t2 by position for each speaker, separated by interval. The first connected pair represents t1 and t2 in sentence 1, and so on through t1 and t2 in sentence 5. Diamonds represent long interval sentences; circles, short interval. There is no difference between the onsets and offsets of long versus short sentences. Across speakers, there is no consistent pattern involving the relative degree of differentiation among medial sentences in both types of intervals.

Unlike the other speakers, FKL shows no difference between the second and third positions. This, together with the possibility that the relationship between the vowels in second position may have affected the quality of t2 for the other two speakers, suggests that under optimal conditions the three medial sentences might not have been distinguished from one another.

Thorsen (1986) also found that the shape of the textual contour differed depending on sentence length. The onsets and offsets of longer medial sentences (three stress groups) were higher and lower, respectively, than the onsets and offsets of shorter sentences (two stress groups). Moreover, the overall slope of the shorter medial sentences showed a more evenly distributed declination than the longer medial sentences. Her explanation is that each sentence must have a particular slope. In the production of longer sentences, there is less room within the speaker's pitch range to successively lower the accents and simultaneously preserve the sentence slopes.

However, unlike Thorsen's supplementary results, there was no consistent difference between long and short medial sentences in Toishan. Figure 6 shows the averages of the target tones separated by interval. The diamonds represent long interval sentences, the circles short interval. In general, there was no consistent difference between the onsets and offsets of short versus long medial sentences. As far as the relative degree of differentiation is concerned, there was no consistent pattern across speakers. In GFL's case, it is clear that in both lengths the medial positions are distinguished to the same extent from initial and final positions. For FKL, the short medial sentences cluster together, but are more differentiated from one another in the long interval. THY's long medial sentences cluster together; the short medial sentences are differentiated from one another, yet do not form an overall declining slope. Thus, there is no evidence for any "nesting" of declination backdrops in the Toishan data.

4. Conclusion

The results of this experiment do not support the existence of a declination component in Toishan. The amount of downtrend is not in any way correlated with the length of the intervening material between the target tones. Thus, we cannot assume that languages which do not have phonologically triggered downtrend exhibit declination, and that declination is an automatic, universal phenomenon.

The results are, on the other hand, consistent with a view of local F0 implementation. The positions of initial and final sentences are marked in the overall slope, and medial sentences as a whole are distinguished from the peripheral sentences (though probably not from each other). The hierarchical backdrop declination effects proposed in Thorsen's model did not appear in the five sentence paragraphs in the Toishan data. The results of the current study argue against a global implementation of F0.

Finally, the results are of particular interest when compared with the results of a parallel, simultaneous experiment involving Cantonese (see Johnson, this volume). Toishan and Cantonese are, to a large extent,

structurally similar. In addition to finding evidence for initial raising, Johnson found that interval length was a significant factor--suggesting the effect of a declination component. Not only is this finding important in terms of language relatedness[2], but also in terms of language contact. Both Toishan and Cantonese are spoken in the same geographical areas, and many speakers can understand both varieties. Many speakers of Toishan are fluent in Cantonese as well (although the reverse is not always true; this may be due to the sociolinguistic standing of the two varieties, since Toishan has the position of lower prestige). Thus, some aspects of downtrend are cross-linguistic, but others are language specific.

---

[2]Chiu-yu Tseng (1984) found no evidence for a declination component in Mandarin ("Sentence intonation is context/discourse oriented: evidence from Mandarin Chinese", abstract, Journal of the Acoustical Society of America, Suppl. 1, Vol. 75, Spring).

Acknowledgements

A preliminary version of this paper, "Declination in Toishan", was presented at the 111th meeting of the Acoustical Society of America, Cleveland, May 12-16 1986. My thanks to Mary Beckman for her advice and comments throughout the investigation. I am also grateful to Timothy Light for sharing his unpublished work on Toishan. Keith Johnson and Ken de Jong gave me much of their time in discussing these issues and (especially) in relieving my occasional frustrations with the 11/23.

References

Beckman, Mary E. and Janet B. Pierrehumbert. 1986. "Intonational Structure in Japanese and English". Phonology Yearbook 3:255-309.

Cheng, Teresa M. 1973. "The Phonology of Taishan". Journal of Chinese Linguistics, Volume 1, Number 2 (May):256-322.

Hirschberg, Julia and Janet Pierrehumbert. 1986. "The Intonational Structuring of Discourse". Proceedings of the 24th Annual Meeting of the Association for Computational Linguistics:136-144, June 10-13, Columbia University, New York.

Ladd, D. Robert. 1983. "Declination: A Review and Some Hypotheses". Phonology Yearbook 1:53-74.

_____. 1983. "Phonological Features of Intonational Peaks". Language 59:4 721-759.

Liberman, Mark and Janet Pierrehumbert. 1984. "Intonational Invariance under Changes in Pitch Range and Length". Language Sound Structure 157-233, Mark Aronoff and Richard T. Oehrle, ed. MIT Press.

Light, Timothy. 1981. "Toishan: An American Language (Problems for Learners of Standard Chinese)". Paper presented at the meeting of the National Association for Asian and Pacific American Education (April).

Lindau, Mona. 1986. "Testing a Model of Intonation in a Tone Language". Journal of the Acoustical Society of America 80:757-764.

McCoy, (William) John. 1969. "The Phonology of Toishan City". Orbis 18:108-122.

Pierrehumbert, Janet Breckinridge. 1980. The Phonology and Phonetics of English Intonation. PhD dissertation, MIT.

Thorsen, Nina Grønnum. 1985. "Intonation and Text in Standard Danish". Journal of the Acoustical Society of America 77:1205-1216.

_____. 1986. "Sentence Intonation in Textual Context--Supplementary Data". Journal of the Acoustical Society of America 80:1041-1047.

Wong, Maurice. 1982. Tone Change in Cantonese. PhD dissertation, University of Illinois. University Documents, Ann Arbor.

Appendix: Corpus

Target tones are indicated by an asterisk. Toishan also has a predictable alternation between long and short tones which is not under consideration here and, thus, is not indicated. In the noun phrase [lɔ pak], the tone on [pak] is underlyingly 33, which becomes 21 by a morphologically conditioned tone change. [pak] means 'paternal uncle', but can also be used, as in this situation, as a term of respect for an unrelated adult male. Because there is no exact equivalent in English, I have translated it as 'uncle'.

Paragraph 1:

*55 21 33 *55 55 lɔ pak yu kaw toy	'Old Uncle has a puppy.'
*55 55 31 31 3 *55 55 kaw toy hay ɬay koy kaw toy	'Puppy is a little puppy.'
*55 55 31 *55 33 kaw toy hay lɛŋ ɬuy.	'Puppy is two years old.'
*55 21 33 33 33 *55 55 lɔ pak tʃoŋyi haŋ kaw fan.	'Old Uncle likes (to) play with (the) dog.'
*55 55 31 *55 21 kaw toy han lɔ pak.	'Puppy loves Old Uncle.'

Paragraph 2:

*55 21 33 33 22 *55 55 lɔ pak yɛt yu fi ma law	'Old Uncle also has (a) fat monkey.'
*55 55 21 *55 33 ma law hay lɛŋ ɬuy	'Monkey is 2 years old.'
*55 55 33 33 33 *55 55 kaw toy m̩ tʃoŋyi ma law	'Puppy doesn't like Monkey.'
*55 55 33 *55 55 kaw toy tʃuy ma law	'Puppy chases Monkey.'
*55 55 55 55 22 *55 55 kaw toy lɛŋ ŋaw fi ma law	'Puppy wants to bite (the) fat monkey.'

Paragraph 3:

*55 55 55 *55 55 kaw toy haw ŋaw +ʃi	'Puppy (with) mouth bites paper.'
*55 55 55 33 55 *55 21 kaw toy i pɔ +ʃi lɔ pak	'Puppy gives (the) newspaper to Old Uncle'.
*55 21 33 *55 55 lɔ pak han kaw toy	'Old Uncle loves Puppy.'
*55 55 55 22 33 *55 21 kaw toy ɬɛŋ na ʃɛŋ lɔ pak	'Puppy wants to climb up on Old Uncle.'
*55 21 55 *55 55 lɔ pak lam kaw toy	'Old Uncle hugs Puppy.'

Paragraph 4:

*55 55 33 55 55 *55 21 ma law yɛt ɬɛŋ lam lɔ pak	'Monkey also wants to hug Old Uncle'.
*55 55 55 *55 55 ma law ŋaw kaw toy	'Monkey bites Puppy.'
*55 21 33 33 33 *55 55 lɔ pak fay[1] nɔ fi ma law	'Old Uncle quickly scolds Monkey.'
*55 21 33 *55 55 lɔ pak han ma law	'Old Uncle loves Monkey.'
*55 55 33 22 33 *55 21 ma law yɛt na ʃɛŋ lɔ pak	'Monkey also climbs up on Old Uncle.'

[1]Read as [kl̂t] by THY. THY also found the characters for 'monkey' unusual, but read them with the correct intonation.