

The Effects of Subject Themmatization on the Tonal Pattern of Japanese Sentences

H. J. Warkentyne

University of Victoria

1.0 INTRODUCTION

It is widely held by students of Japanese that some, but not all, Japanese sentences contain a theme or topic occurring at, or near, the beginning of the sentences. However, there are two prevailing views on what constitutes a theme or how it is recognized. One interpretation is that given by Susumu Kuno (1975), and the other is elaborated by Samuel Martin (1975).

For Kuno, generic and anaphoric sentence-initial noun phrases marked by the enclitic particle wa are thematic or contrastive. Noun phrases preceding the contrastive wa have greater accentual prominence than those preceding the thematic wa. Considerable attention has been focussed on wa when it occurs with the subject noun phrase in which instance it contrasts with the subject-marking particle ga. This contrast can be illustrated by paraphrasing the translation 'That book is red' of the following Japanese sentences. (ano 'that', hon 'book', akai 'is red')

Ano hon ga akai.

'It is that book that is red.'

Ano hon wa akai.

'Speaking of that book, it is red.'

Sentence constituents, other than subject noun phrases, can also be thematized by wa attachment and preposing to the sentence-initial position.

In Martin's view, thematization does not depend on wa attachment to preposed sentence elements. A sentence-initial phrase is thematic only if followed by a major juncture, regardless of which enclitic particle occurs with that

phrase. The relative height of the pitch contours of the two phrases on either side of the major juncture is its primary phonetic feature (Martin 1970:432). In the presence of a major juncture the pitch peak of the second phrase is as high as, or higher than, that of the phrase preceding the juncture. The post-junctural peak is higher for tonic phrases (i.e., phrases containing accented lexical items) than for atonic phrases (i.e., comprised of unaccented lexical items). In addition to relative pitch height, major juncture is accompanied by a slowing of tempo and blocking of phonetic assimilation. Martin also describes a pitch differential between the low-pitched final syllable of an initial tonic phrase and a lexically low-pitched initial syllable of the post-junctural phrase, the latter being somewhat higher. In the absence of juncture, there is no such pitch differential.

Recent studies reported by Neil Galliaiford (1983 a,b) on pitch declination or downdrift in Japanese show that amongst a large set of unrelated and uncontextualized Japanese sentences read by informants, the mean values of the fundamental frequency (Fo) peaks of three- and four-word sentences decline on each successive peak, the first one being the highest. In longer sentences, however, the Fo declination line is more frequently broken so that, for example, the third peak is higher than the second. This suggests that each stretch of utterance corresponding to an unbroken pitch declination line constitutes a single accent phrase and that the resetting of the declination line signals an accent-phrase boundary.

In order to determine if the accent-phrase boundary observed in the previous experiment corresponds to Martin's major juncture, we designed an experiment by which we could elicit sentence utterances with and without thematic elements, and then measure and compare their fundamental frequency contours. The utterance of thematized sentences was elicited by placing them into a context where the thematic elements were clearly anaphoric, yielding Martin's 'subdued theme' and also satisfying Kuno's condition for thematization. The experimental procedure followed is similar to that used by Sorensen and Cooper (1977,1980) in their studies of fall-rise patterns of English.

2.0 METHOD

2.1 The Subjects

The subjects for this experiment were ten female native speakers of Japanese who were raised and educated in the Tokyo area. These women, who were from 20 to 50 years of age, had been resident in Canada for periods ranging from eight months to 25 years. Those residing in Canada for long periods had had close contact with Japan through frequent visits. Although two were part-time teachers of Japanese, none had any special training in linguistics. The subjects were told that we were interested in certain aspects of the language, but were not told what those aspects were.

2.2 The Sentence Data

The data for this experiment consisted of four matched pairs of sentences which were recorded as part of a larger corpus with the target sentences interspersed with filler sentences. The four pairs of sentences appear in Table 1. As can be seen from Table 1, the sentences of each pair are identical in phonetic content except that the A sentences contain the subject-marking ga, whereas the B sentences have the theme marker wa in the same position.

TABLE 1
The Sentence Data

- | | | | | |
|----|----|----------|--------|-------------|
| 1. | A) | zoo ga | hana o | itamemasita |
| | B) | zoo wa | hana o | itamemasita |
| | | elephant | nose | injured |

The elephant injured his trunk.

- | | | | | |
|----|----|----------|------------|------------|
| 2. | A) | Taroo ga | sono hon o | yomimasita |
| | B) | Taroo wa | sono hon o | yomimasita |
| | | Taroo | that book | read |

Taroo has read that book.

3. A) kanojo ga odoriko desu
 B) kanojo wa odoriko desu
 she dancer is

She is the dancer.

4. A) Taroo ga Yokohama e ikimasita
 B) Taroo wa Yokohama e ikimasita
 Taroo Yokohama went

Taroo went to Yokohama.

2.3 Procedure

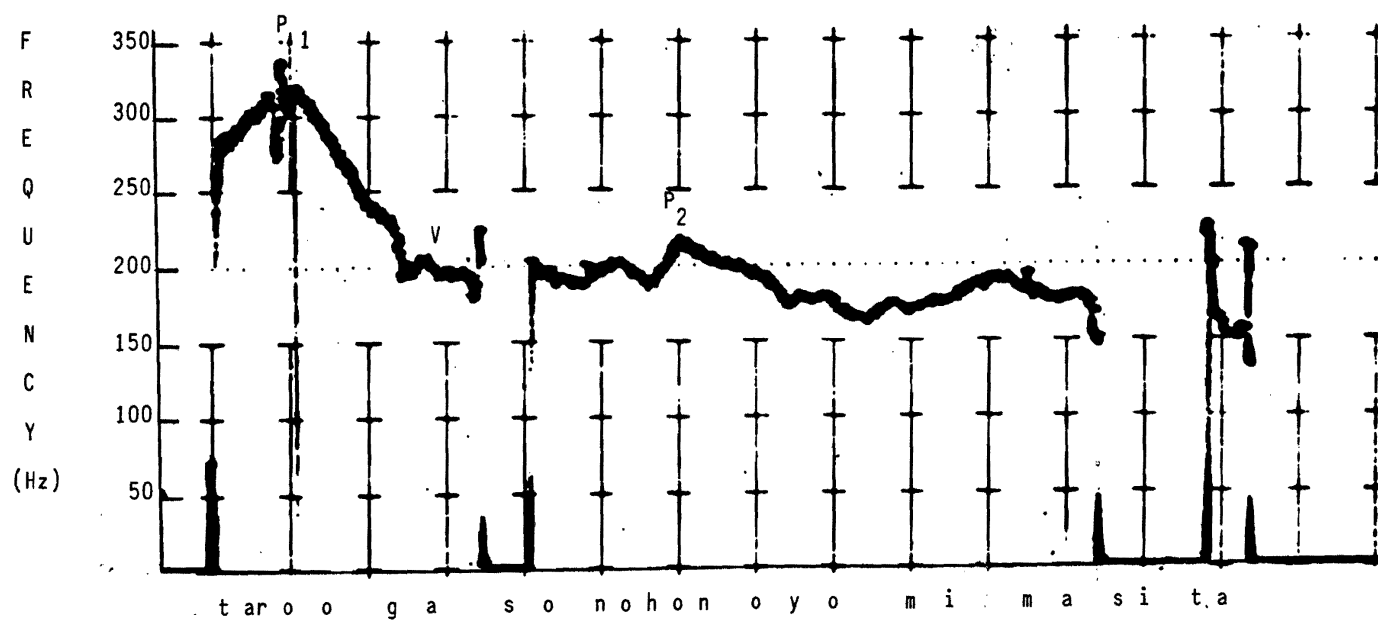
The eight target sentences, interspersed with thirteen other sentences, were written in the conventional way, using kanji and hiragana, on index cards. The cards were arranged in two different quasi-random orders with filler sentences appearing at the beginning and end of each order. Five subjects read the sentences in one order, five in the other. The subjects sat in a sound insulated room with the index cards in front of them. They heard instructions, in Japanese, over a set of stereo headphones. The subjects were instructed to read the sentences in response to prompting sentences which provided contexts for each target sentence. The contexts were designed to elicit the target sentence in a natural way.¹ Each context sentence was repeated twice. The first time the subjects heard the context sentence they were to read the associated target sentence silently to themselves. On hearing the cue sentence a second time they read aloud, or recited, the target sentence. The subjects were instructed to use a normal conversation style rather than a reading or a highly dramatic style. If a subject misread a sentence, she simply said 'repeat' and uttered it again. The subjects' utterances were recorded using a Sony ECM 200T condenser microphone, a Revox Model A77 dual track tape recorder and 1.5 mil Scotch 176 Audio Recording Tape run at 7.5 ips. The whole recording session took about 15 minutes and none of the subjects reported having any difficulty with the task. Ultimately, four subjects were re-recorded uttering the same sentences presented this time in the other order.

2.4 Measurements

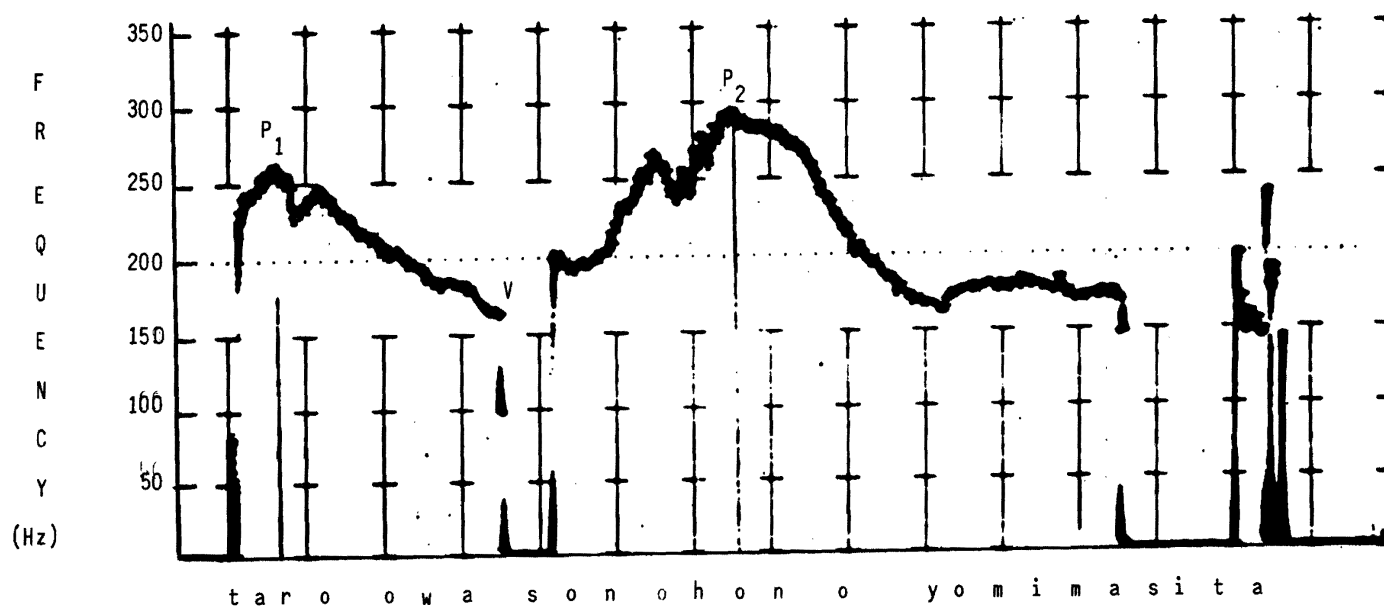
Measurements of the fundamental frequency patterns of the target sentences were obtained using an FJ Electronics Fundamental Frequency Meter (Type FFM 650), the output of which was directed to a Honeywell Model 1806A Visicorder Oscillograph. The Fo meter uses a sequence of high and low pass filters to extract the fundamental frequency of speech input and the Visicorder provides a hardcopy trace of Fo versus time. The system was calibrated using input signals of known frequency, the traces of which were used to construct a scale against which to measure the Fo traces of the target sentences. A sample of Fo trace appears as Figure 1.

Three measurements of Fo were obtained for each sentence. Since we were interested in the fall-rise pattern in Fo associated with the ga and wa enclitics, the peak in Fo (P1) associated with the initial word in each sentence was located and its frequency noted. Similarly the Fo valley (V) associated with the enclitic and the Fo peak (P2) associated with the following word were also noted. The fall-rise pattern is then defined by the fall in Fo from P1 to V and the rise in Fo from V to P2. The times of occurrence of each of these points (T1, TV, and T3) were also noted, time zero being onset of voicing in the first voiced segment in the utterance. These times were later used to calculate elapsed times between the data points in order to characterize the temporal structures of the fall-rise patterns. The total time of utterance from onset of voicing to the cessation of voicing in the last voiced segment of the utterance was also recorded.

In order to avoid selecting Fo peaks which were the result of short time perturbations in Fo associated with preceding voiceless segment (J.-M. Hombert 1978), no peak could be located within 20ms of such a segment. As well, the Fo of a possible peak had to be consistent with the values for Fo on either side of the peak. Where momentary peaks occur (duration < 10ms) an estimation of Fo was made based on values on either side.



Sentence 2A



Sentence 2B

Figure 1. Fo Traces of Sentences 2A and 2B, spoken by Subject 5

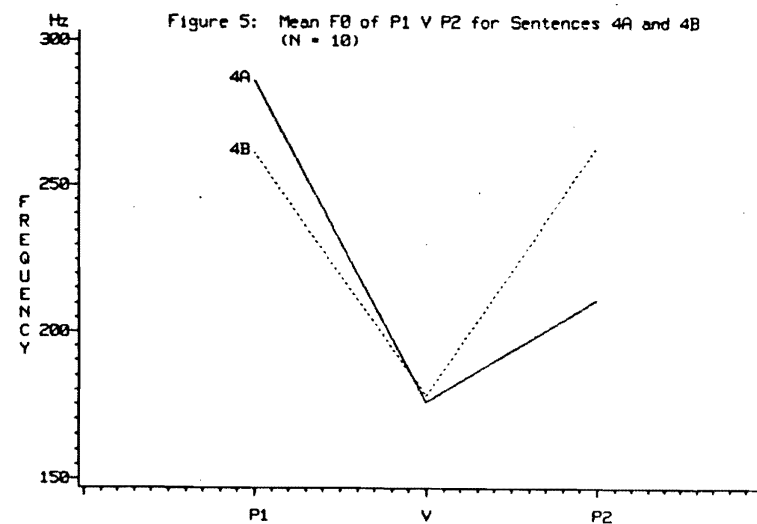
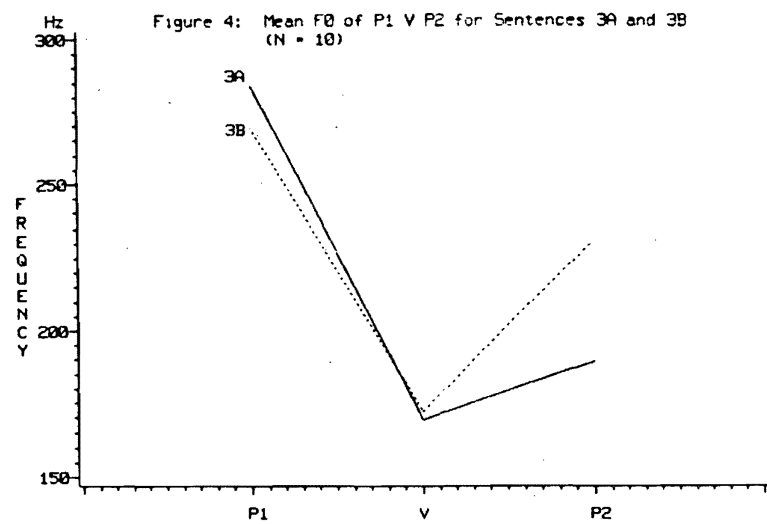
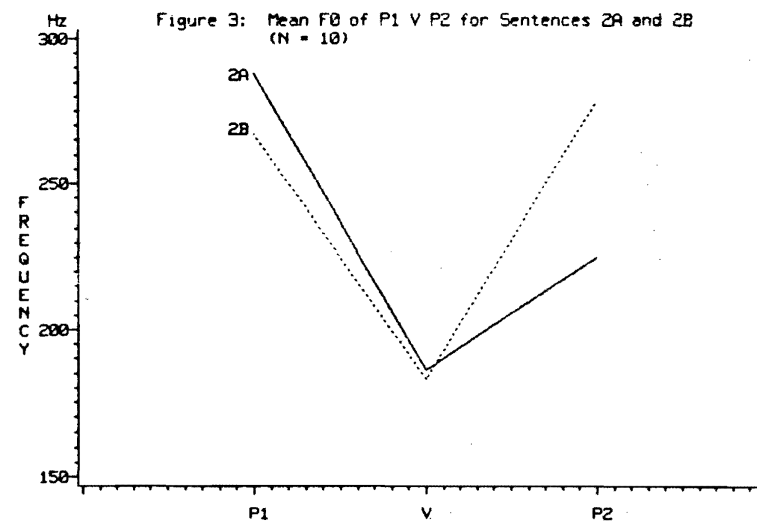
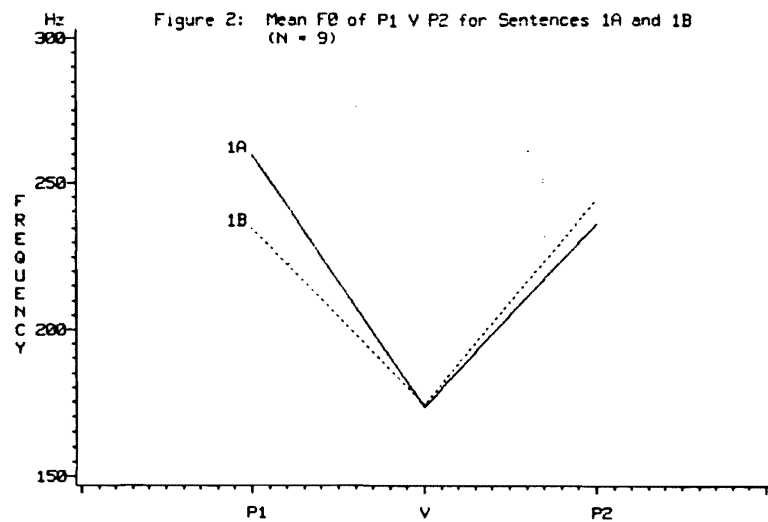
2.5 Results and Discussion

The means and standard deviations of the F_0 of the peaks and valleys together with the means and standard deviations of the elapsed times from P1 to V, V to P2, and P1 to P2 for each sentence appear in Table 2. The F_0 data in Table 2 reveal that there is a consistent difference in the F_0 contours of the matched sentences. In general terms we can say that in the B sentence of each pair P1 tends to be lower and P2 tends to be higher than in the A sentence; there is little difference in the mean F_0 of the valleys within the sentence pairs. (See Figures 2 through 5.)

TABLE 2

Means and standard deviations of F_0 for P1, V, and P2, and of elapsed time from P1 to V, V to P2, and P1 to P2 for all sentences

		Fo (in Hertz)			Duration (in milliseconds)			
		P1	V	P2	P1-V	V-P2	P1-P2	sentence total
1A		260	173	236	185	263	448	1253
	SD	32.6	17.0	20.3	26	66	60	65
1B		235	174	244	188	286	474	1253
		32.1	15.6	25.7	35	50	55	98
2A		288	186	225	258	266	524	1409
		34.3	17.8	29.7	51	109	120	73
2B		267	183	278	244	349	592	1393
		34.2	18.9	30.9	41	42	59	74
3A		284	169	189	334	136	470	960
		24.6	14.2	17.2	73	33	71	100
3B		269	172	231	290	238	528	916
		43.3	15.5	26.8	53	54	78	75
4A		286	176	211	279	220	499	1348
		32.0	15.0	23.5	46	45	76	82
4B		261	178	263	264	262	526	1331
		32.5	12.7	30.9	22	62	67	83



In order better to characterize the differences in the fall-rise patterns associated with the ga and wa enclitics, the changes in F_0 between P1 and V, between V and P2, and between P1 and P2 were calculated in semitones.³ We compared these changes in pitch by means of paired t-tests. We found that the difference in the pitch fall (from P1 to V) in the A and B sentences of each pair was significant in pairs 1 and 4.⁴ The mean fall in pitch was greater in the A sentences than in the B sentences for these pairs. Pairs 2 and 3 showed a statistically non-significant trend in the same direction. The difference in the pitch rise (from V to P2) was significant in pairs 2, 3, and 4, but not in pair 1.⁵ In the three pairs which showed a significant difference, the pitch rise was much greater in the B sentences than in the A sentences. In pair 1, which showed no significant difference, the pitch rise in sentences 1B was greater than the rise in 1A for 5 of the 9 speakers for which we had data.

Since the F_0 of the valleys was reasonably stable in each pair, the shift in the fall-rise pattern seems mainly to involve an adjustment in the peaks. A comparison of the pitch changes from P1 to P2 in the paired sentences reveals that there was a significant difference in the relative heights of the two peaks in all of the pairs. This difference was significant at the 0.001 level in pairs 2, 3, and 4, and at the 0.05 level in pair 1. Out of a total of 39 A sentences, 37 sentences showed a fall in pitch for P1 to P2; 2 sentences utterances (both 1A) had a rise in pitch from P1 to P2. In slightly more than half of the B sentences (20/39) the pitch rose (or stayed the same) from P1 to P2. Of the remaining 19 B sentences, 18 had less of a pitch fall from P1 to P2 than did the A sentences. This means that in 38 of 39 B sentences pitch either rose from P1 to P2 or fell less than it did in their A counterparts.

It seems safe to say that there is a significant difference in the behaviour of pitch in sentences with thematic and non-thematic subjects. This difference in fall-rise pattern involves a lowering of the first F_0 peak in the sentence and a raising of the second peak if the sentence contains a thematized subject.

As noted earlier, according to Martin major juncture is also evidenced by slowing of tempo, in addition to the feature concerning pitch contours. The mean values for the various time variables in Table 2 suggest that there may be consistent differences in the duration of the falls and rises and in the fall-rise pattern as a whole. The mean duration from P1 to V is greater in the A sentences than in the B sentences for three of the four pairs; in pair 1 the durations are nearly equal. Conversely, the mean duration of the pitch rise is greater in the B sentences for all sentence pairs. The mean duration of P1 to P2 is greater in the B sentences than in the A sentences. This greater duration of the fall-rise pattern in the B sentences does not seem to be accompanied by an increase in the duration of the whole sentence -- the mean duration of the A sentences is longer than that of the B sentences in three of the four pairs; in pair 1 they are the same. However, a closer look at the individual sentences reveals that the means are perhaps deceptive in their consistency. The speakers are about evenly divided as to whether the pitch fall takes longer in the A or B sentences. The B sentences more often have a greater duration for the pitch rise than do the A sentences, but in one-quarter of the utterances the pitch rise takes longer in the A sentences than in the B sentences (9/39). The speakers are once again about evenly divided as to whether the fall-rise pattern as a whole is longer in the A or the B sentences. In pairwise t-tests of all time variables for matched sentences only in two cases, Nos. 2 and 3, do the differences between A and B sentences reach statistical significance. The difference in the duration of the pitch rise in sentence pair 2 is significant at the 0.05 level, and the difference in the duration of the rise in pair 3 is significant to the 0.001 level. In the other two pairs, the differences are non-significant on the basis of the t-test metric. Similarly, the overall duration of the sentences was not significantly different for the matched pairs.

3.0 CONCLUSIONS

Our experiment has demonstrated that subject thematization results in the lowering of the Fo peak of that noun phrase relative to the non-thematized form. In three of the four cases examined, there is clear evidence of resetting the

F₀ declination line following the theme, from which fact the presence of a major intonation or accent phrase boundary can be inferred. Sentences 1B, 2B, and 4B also conform to Martin's conditions for a major juncture in respect of relative heights of the pitch contours flanking the boundary. Evidence for the temporal characteristics associated with major juncture is less clear, although a slight tendency towards slowing the tempo was observed. A greater pitch differential between final and initial syllables of the first and second pitch contours in the case of the thematized forms was also observed, but no measurements were made.

Sentence pair 3 represents a form different from the rest. These are 'identificational' sentences (Martin 1975:239-242) and it is possible that 3B has a focussed theme (ibid:240) rather than a subdued theme, in which one would not expect a lowered F₀ peak on the theme and, being a two-word sentence, no resetting of the declination line is present. However, sentence 3B shows the strongest temporal cue for major juncture, of all B sentences the mean difference in P1 to P2 duration between 3A and 3B being significant at the .0001 level. It may be that there are alternate phonetic cues for major junctures or major accent phrase boundaries where theme is concerned.

FOOTNOTES

¹ For example, the context sentences for 1A and 1B were: 1A) kyoo Ueno-doobutsu-en de, taihen na koto ga okorimasita. 'A terrible thing happened at Ueno Zoo today.' and 1B) Zoo wa genki ga arimasen ne? Doositan desyoo ka? 'The elephant is not well, is he? What happened to him?'

² N = 10 for all pairs except pair 1 where N = 9. Sentences 1A and 1B as spoken by Subject 1 could not be used because of poor recording quality.

³ The following formulas (based on Weitzman 1970) were used to calculate three derived measures of pitch change -- Delta1 represents the pitch change from P1 to V, Delta2 the change from V to P2, and Delta3 the relative heights of P1 and P2:

$$\begin{aligned}\text{Delta1} &= 39.86 \cdot \log(V/P1) \\ \text{Delta2} &= 39.86 \cdot \log(P2/V) \\ \text{Delta3} &= 39.86 \cdot \log(P2/P1).\end{aligned}$$

⁴ Pair 1: $p < 0.05$, $t = -3.26$, $df = 8$
Pair 4: $p < 0.01$, $t = -3.47$, $df = 9$

⁵ Pair 2: $p < 0.001$, $t = -5.22$, $df = 9$
Pair 3: $p < 0.001$, $t = -5.99$, $df = 9$
Pair 4: $p < 0.001$, $t = -5.29$, $df = 9$

⁶ Pair 1: $p < 0.05$, $t = -2.54$, $df = 8$
Pair 2: $p < 0.001$, $t = -5.07$, $df = 9$
Pair 3: $p < 0.001$, $t = -5.23$, $df = 9$
Pair 4: $p < 0.001$, $t = -8.04$, $df = 9$

REFERENCES

- Cooper, William E. and John M. Sorensen. 1977. Fundamental Frequency Contours at Syntactic Boundaries. JASA 62:3:683-92.
- _____. 1981. Fundamental Frequency in Sentence Production. New York: Springer Verlag.
- Gallaiford, Neil. 1983a. The Effect of Downdrift on Major Phrase Intonation in Japanese. WPLC 3:1,37-44.
- _____. 1983b. Pitch Declination in Japanese: An Instrumental Study, University of Victoria M.A. Thesis.
- Hombert, J.-M. 1978. Consonant Types, Vowel Quality and Tone. In V. Fromkin (ed.) Tone -- A Linguistic Survey. New York, Academic Press.
- Kuno, Susumu. 1975. The Structure of Japanese. M.I.T. Press.
- Martin, Samuel E. 1970. Junctural Cues to Ellipsis in Japanese. Studies in General and Oriental Linguistics, 429-46, Tokyo (TEC).
- _____. 1975. A Reference Grammar of Japanese, New Haven and London, Yale University Press. (Second printing 1983).
- Weitzman, Raymond S. 1970. Word Accent in Japanese, Studies in the Phonology of Asian Languages No.9, Mieko S. Han (Series ed.), Detroit, Management Information Services.