Ann. Bull. RILP, No. 14, 225-236 (1980)

COMPREHENSION OF RELATIVE CLAUSE CONSTRUCTION AND PITCH CONTOURS IN JAPANESE*

Tazuki Uyeno**, Hideo Hayashibe***, Kunihiko Imai****, Hiroshi Imagawa and Shigeru Kiritani

1. Introduction

Sentences in speech are characterized by segmental and suprasegmental components. When we comprehend these sentences, we must go through the process of reconstructing the given phonological sequence into hierarchical constructions matching the corresponding meaning. In this process, it is assumed that intonation plays an important role in representing the internal structure involved in a sentence, besides its function of conveying the speaker's intention by designating the sentence as declarative, interrogative, imperative, etc. (Hadding & Studdert-Kennedy, 1973; Hakoda & Sato, 1978).

Our previous study (Uyeno et al., 1979), which treated declarative complex sentences in Japanese, revealed that sentence intonation plays a crucial role in a structurally ambiguous sentence. On the basis of our findings, a listening test was conducted to investigate how the difference in pitch contours of relative clause constructions affects comprehension by speakers of Japanese.

In Japanese relative clause constructions, pitch pattern differs depending on the location of the relative clause. When the relative clause is located at the sentence-initial position, which is called "left-branching (hereafter, LB), the pitch contour is characterized by a high rise at the onset of the utterance, gradually declining toward the end of the utterance. When the relative clause is embedded in the center of the matrix sentence, which is called "center-embedding" (hereafter, CE), the pitch contour is characterized by a high rise in pitch at the onset of the utterance and by a repeated high rise in the relative clause.

Three ambiguous relative clause constructions have been used in the experiment:

(1)	Ototoi	koronda	otona ga	waratta.
	day before yesterday	fell	adult	laughed

The adult who fell the day before yesterday laughed.

or: The adult who fell laughed the day before yesterday.

* This study was supported in part by a Grant-in-Aid for Scientific Research from the Japanese Ministry of Education Grant No. 410207

** The National Language Research Institute
*** Tokyo Gakugei University
**** Tokyo Metropolitan University

(2) Anoto'ki ugo'ita ano'ko ga osa'eta.

that time moved that child caught

That child who moved at that time caught (it).

or: That child who moved caught (it) at that time.

(3) A'rutoki n'igeta e'nzi ga k'aita.

at one time ran away kindergarten child drew

The kindergarten child who at one time ran away drew (it).

or: The kindergarten child who ran away drew (it) at one time.

Sentence (1) uses lexicons without marked accent, (2) those with high pitch accent on the word-medial mora, and (3) those with high pitch on the word-initial mora. An accented mora is indicated by the mark ['] immediately following it.

Each sentence is constructed with a chain of four phrases: adverbial phrase - verb phrase - noun phrase - verb phrase. The sentence-initial adverb may modify either the immediately following verb or the verb located in the sentence-final position, as in (4), where the adverb ototoi in (1) can modify the immediately following verb in LB or the sentence-final verb in CE.

(4) a. LB: [Ototoi koronda] otona ga waratta.

The adult who fell the day before yesterday laughed.

b. CE: Ototoi [koronda] otona ga waratta.

The adult who fell laughed the day before yesterday.

2. Experiment

2.1 Stimulus

2.1.1 Procedure of Analysis-Synthesis

Six test sentences, uttered by an adult male speaker of the Tokyo dialect one at a time without pause, were recorded in a sound-proof room. The speech signals were then lowpass-filtered to 5kHz, sampled at 10kHz and digitized in 10 bits. PARCOR analysis was made every 6.4 msec. for Hamming-windowed speech of 19.2 msec. The order of the analysis was 12. Voiced/unvoiced judgment was performed based on the values of the maximum peak in the autocorrelation function of the residual wave for the period of 38.4 msec. In the case where is was judged 'voiced', the pitch period was determined by detecting the location of the maximum peak in a range of 2.5–16 msec. in the autocorrelation function of the residual wave. The extracted pitch contours of the sentences (hereafter, original pitch contours) together with the speech envelopes are shown in Figs. 1–6.

Fig. 1 Pitch Contours of Sentence 1: LB

Fig. 2 Pitch Contours of Sentence 2: LB

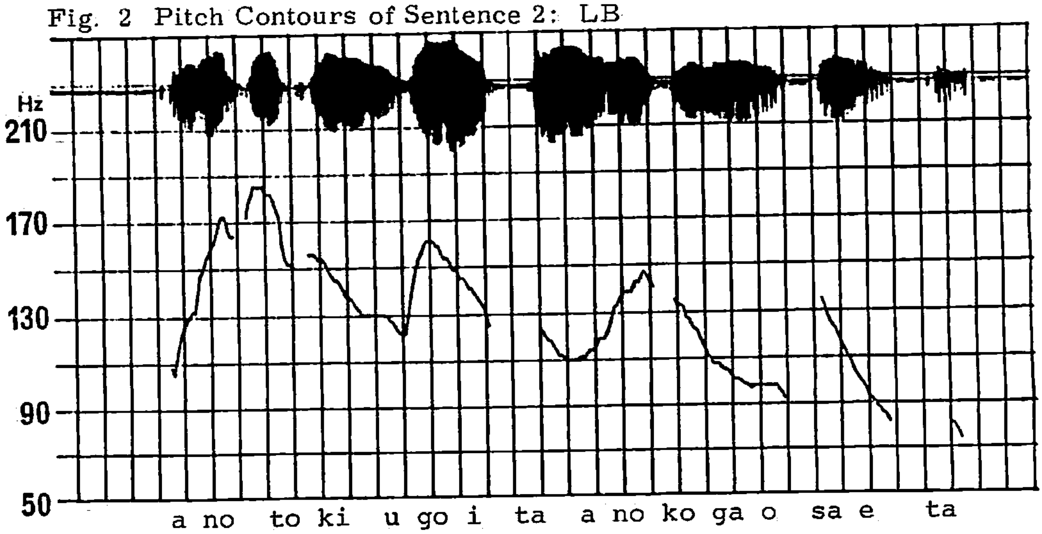


Fig. 3 Pitch Contours of Sentence 3: LB

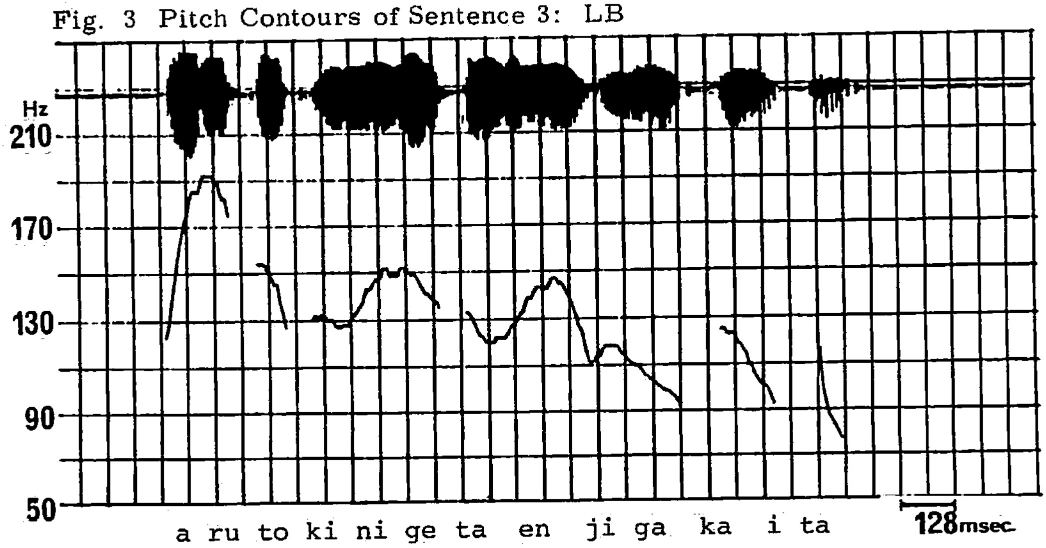


Fig. 4 Pitch Contours of Sentence 1: CE

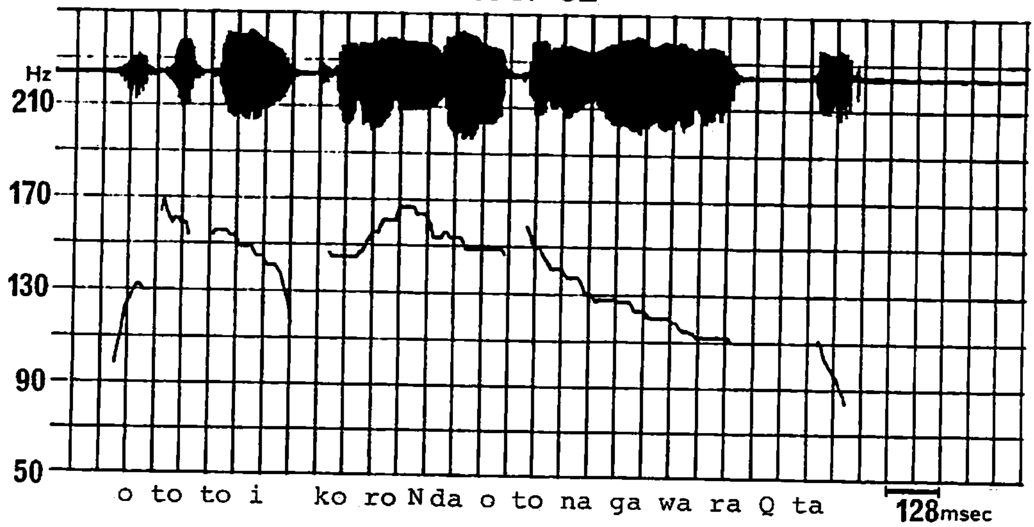


Fig. 5 Pitch Contours of Sentence 2: CE

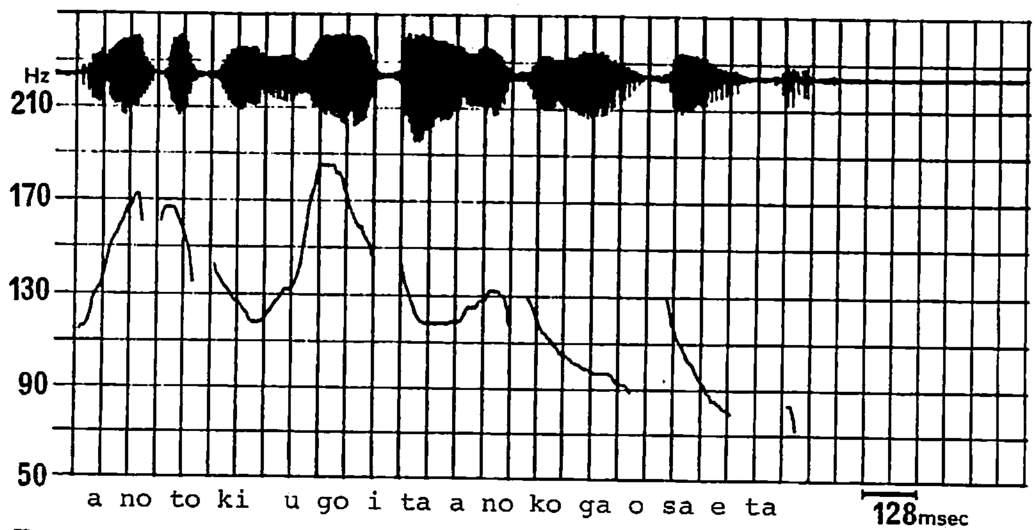
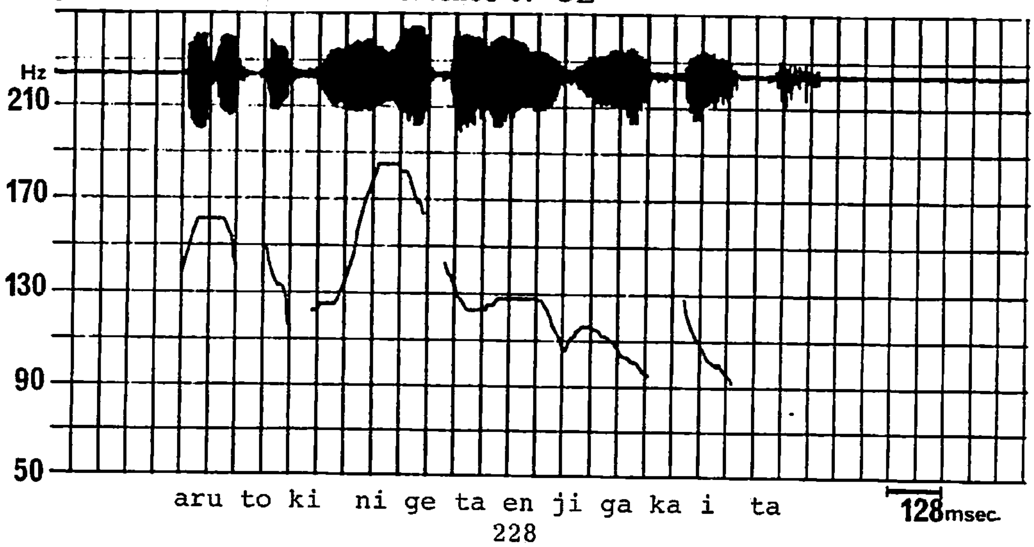


Fig. 6 Pitch Contours of Sentence 3: CE



The original pitch contours of the sentences in CE are modified as described below, and the stimulus sounds with modified pitch contours were synthesized by the PARCOR synthesis method. These three sentences in CE are hereafter called test sentences 1, 2 and 3.

2.1.2 Synthesis of the Pitch Contours:

In the original pitch contours of the three test sentences, the pitch rises first at the beginning of the adverbial phrase, then descends once, and rises again at the next verb phrase (relative clause), after which the curve gradually descends toward the end.

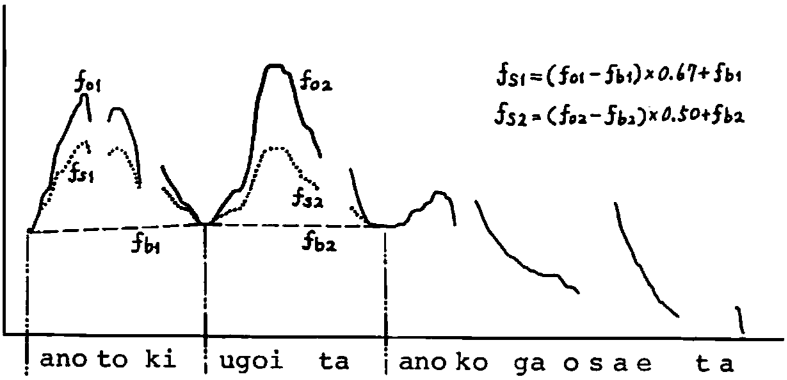
In order to synthesize the pitch contours in which the height of two peaks are independently changed, the beginning and end of the pitch pattern in the portion of the adverbial phrase at the beginning of the sentence (hereafter, A) are connected by a straight line, which is considered as the base line. Then, the synthetic pitch curve is determined as follows:

$$f_s = (f_o - f_b) \times C + f_b$$

where f_b is the pitch frequency for base line, f_o is the original pitch frequency and f_s is the synthesized pitch frequency. C is the coefficient to change the height of the peak in the pitch contour. The procedure is illustrated in Fig. 7.

The height of the peak of the pitch pattern in the portion of the verb phrase, which constitutes a relative clause (hereafter, B), is likewise changed by multiplying the difference between the base line and the original pitch pattern by a coefficient. In the case of test sentence 1, the pitch pattern is changed to the portion near the end of the sentence, because of the smallness of the peak in the portion of the verb phrase.

Fig. 7 Pitch variations in A and B given to test sentence 2, where f_{o1} and f_{o2} are original pitch contours, f_b is the pitch frequency for base line and f_s is the synthesized pitch frequency.



Using the above method, the pitch contours were changed 2 steps in A and 4 steps in B in test sentence 1; and 4 steps in both test sentences 2 and 3. As a result, 40 stimulus sentences with synthesized pitch contours were prepared. In test sentence 1, when the peak in A is made higher than that of the original (i.e., the coefficient is 1.0 and above), the synthesized sentences become unnatural and sound as if the word accents were changed.

These sentences were then eliminated from the stimulus sentences. Accordingly, the number of synthesized stimuli is 8 in test sentence 1, while it is 16 both in test sentences 2 and 3. The synthesized pitch contours are shown in Figs. 8-17.

In Table 1, all the coefficients applied to the A-portion and B-portion in the 40 respective stimulus sentences, and the ratio of the height (Hz) at the peak in B to that in A (B/A) are shown. The larger portion indicates the higher pitch in B, and the smaller portion indicates the higher pitch in A. The dotted and broken lines in the table will be referred to in section 2.5.

Table 1. Ratio of the pitch in B to the pitch in A.
The figures in the upper row are the coefficients applied to the A-portion, and the figures in the left column are the coefficients applied to the B-portion in the respective test sentences.

TEST SENTENCE 1

B \ A		0.5		1.0
0.0	(1)	0.97	(2)	0.82
0.5	(3)	1.06	(4)	0.91
1.0	(5)	1.15	(6)	0.98
1.3	(7)	1.21	(8)	1.03

TEST SENTENCE 2

B \ A		0.33		0.67		1.0		1.33
0.25	(9)	0.99	(10)	0.88	(11)	0.78	(13)	0.71
0.5	(13)	1.11	(14)	0.98	(15)	0.87	(16)	0.79
0.75	(17)	1.24	(18)	1.10	(19)	0.98	(20)	0.88
1.0	(21)	1.38	(22)	1.22	(23)	1.08	(24)	0.99

TEST SENTENCE 3

B \ A		0.5		1.0		1.5		2.0
0.25	(25)	0.93	(26)	0.85	(27)	0.78	(28)	0.72
0.5	(29)	1.03	(30)	0.94	(31)	0.87	(32)	0.78
0.75	(33)	1.14	(34)	1.04	(35)	0.96	(36)	0.88
1.0	(37)	1.25	(38)	1.14	(39)	1.05	(40)	0.96

2.2 Subjects

Thirty-four Japanese native adults (24 females, 10 males, mean age 26) participated in the experiment. The factor of dialects was neglected though it might affect the comprehension of the sentences having differently accented pitch patterns.

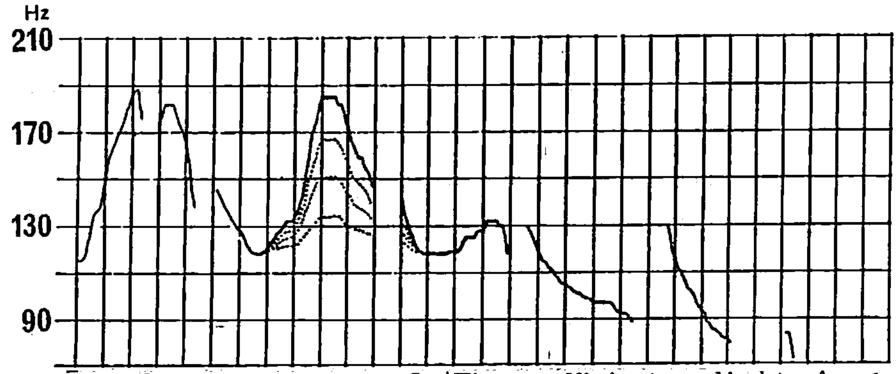


Fig. 8 Test sentence 2 (The coefficient applied to A = 1.33)

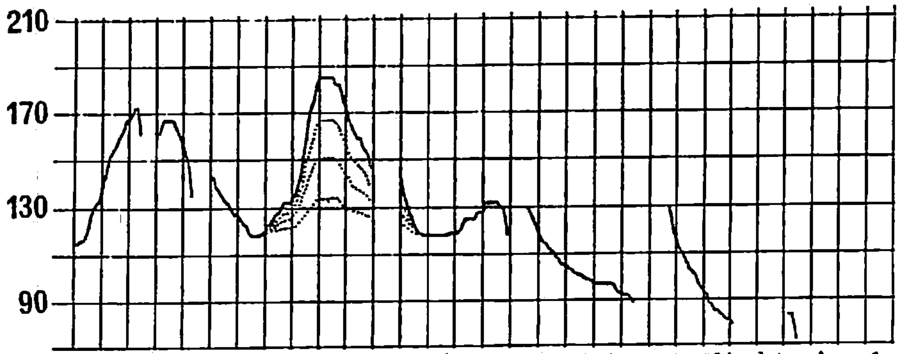


Fig. 9 Test sentence 2 (The coefficient applied to A = 1.0)

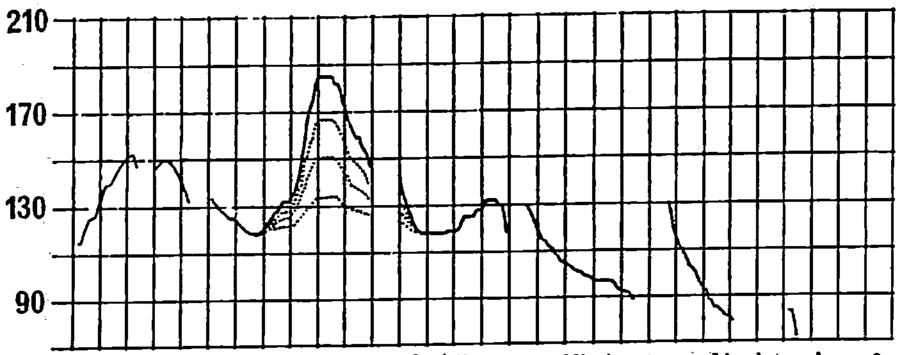


Fig. 10 Test sentence 2 (The coefficient applied to A = 0.67)

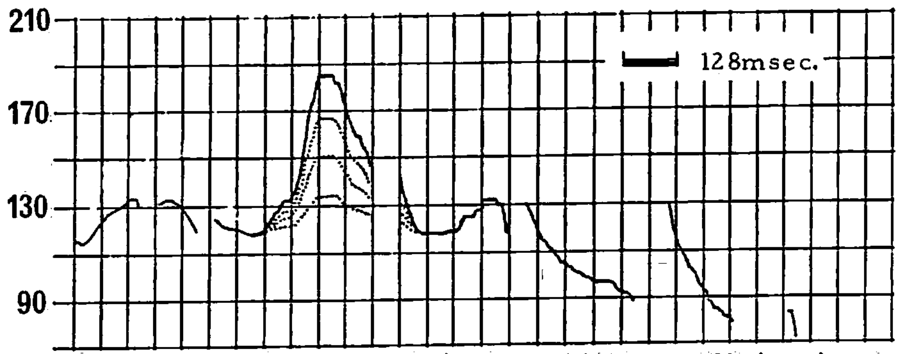


Fig. 11 Test sentence 2 (The coefficient applied to A = 0.33)

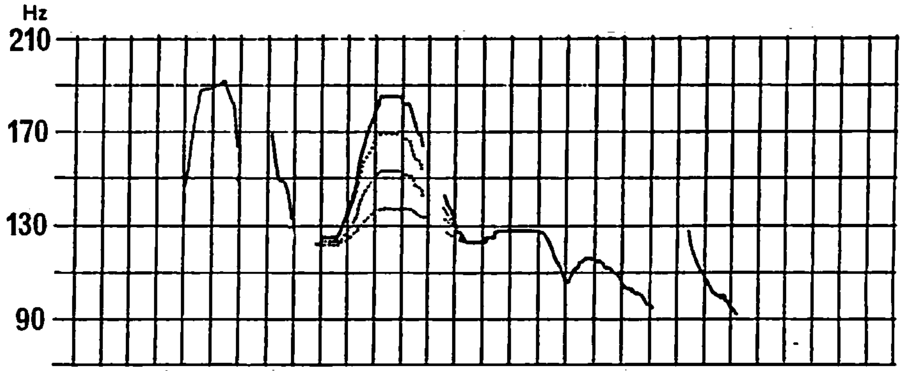


Fig. 12 Test sentence 3 (The coefficient applied to A = 2.0)

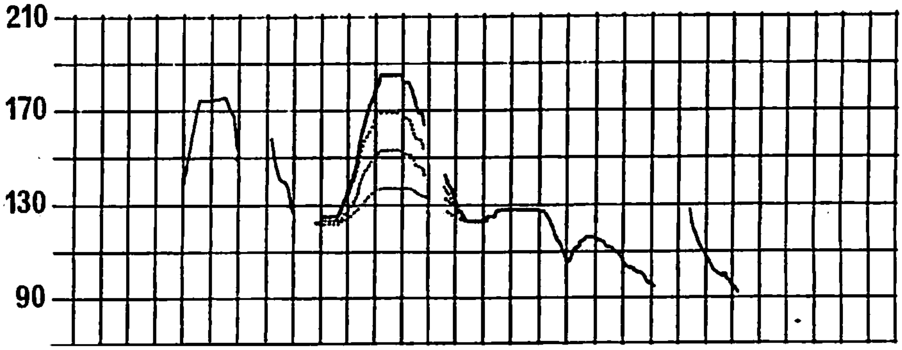


Fig. 13 Test sentence 3 (The coefficient applied to A = 1.5)

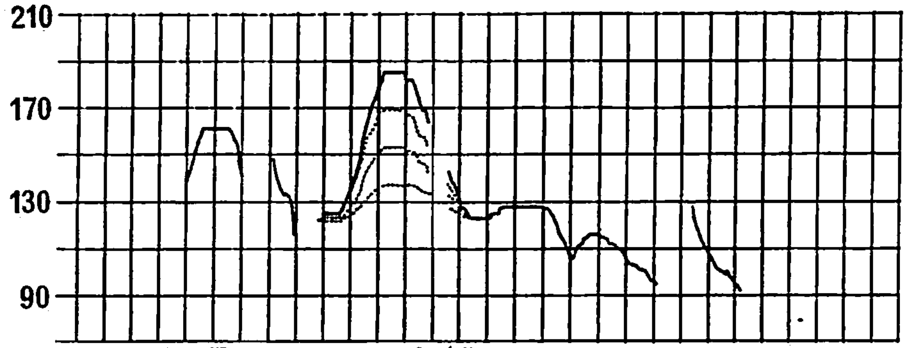


Fig. 14 Test sentence 3 (The coefficient applied to A = 1.0)

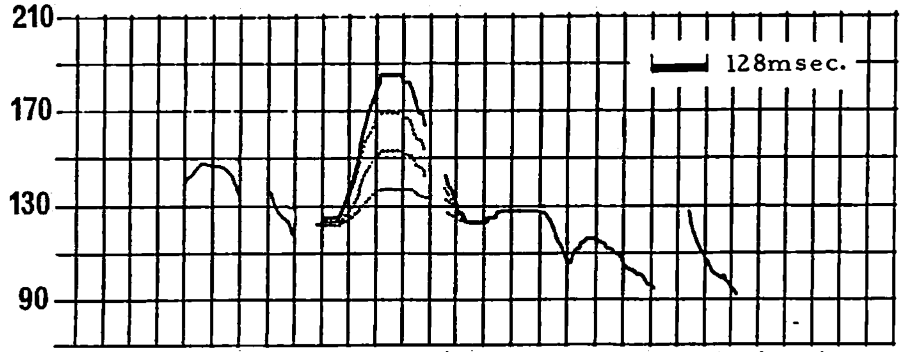


Fig. 15 Test sentence 3 (The coefficient applied to A = 0.5)

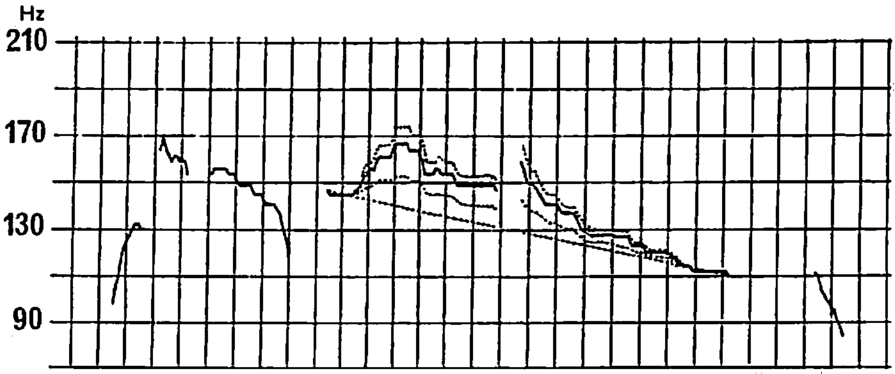


Fig. 16 Test sentence 1 (The coefficient applied to A = 1.0)

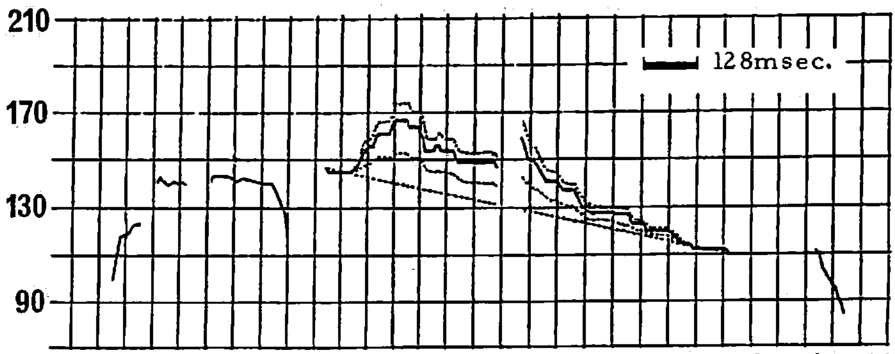


Fig. 17 Test sentence 1 (The coefficient applied to A = 0.5)

2.3 Procedure

The experiment was carried out to all subjects simultaneously. Prior to the session, a pre-test was given to the subjects in order to let them understand the purpose of the experiment. They tried to understand the sample sentence:

kinoo	moratta	ringo o	tabeta
yesterday	received	apple	ate

I ate the apple that I received yesterday.

or: Yesterday, I ate the apple I received.

which could be interpreted two ways, i.e., left-branching and center-embedding interpretations, depending on the difference in pitch contours.

The subjects had to choose one of two types of modification as shown in print-outs like that below:

☐ kinoo moratta
yesterday received
I received (it) yesterday.

☐ kinoo tabeta
yesterday ate
I ate (it) yesterday.

If a subject had understood the sentence as left-branching, he would mark the former, and if center-embedding, the latter.

The subjects were instructed not to look at the printed items while they were listening to the stimulus sentences, because they were expected to listen to the sentences as naturally as possible.

Stimulus sentences were arranged randomly and tape-recorded with 3 seconds of inter-stimulus interval. Three sets of differently randomized lists were prepared. They were presented twice; i. e., each stimulus sentence was presented a total of 6 times.

2.4 Results

The ratio of the interpretation of LB and CE for each stimulus is shown in Table 2. On the basis of χ^2 test, items which obtained significantly more LB-interpretation ($p < 0.01$) are asterisked, and items which yielded significantly more CE-interpretation are marked with an exclamation mark. LB-interpretation is significantly higher in the items above the dotted lines, while CE-interpretation is significantly higher in the items below the broken lines.

From these results, it is clear that an increase in B/A value tends to yield CE-interpretation, and a decrease in B/A value tends to yield LB-interpretation. In detail, the following interpretations are obtained in respective B/A value.

Test sentence 1: LB in $B/A < 0.82$, CE in $B/A > 0.91$

Test sentence 2: LB in $B/A < 0.88$, CE in $B/A > 1.08$

Test sentence 3: LB in $B/A < 0.85$, CE in $B/A > 0.96$

When the absolute value of A and B are extremely low, such as in (1), (9) and (25), however, the cases are slightly different. In (1), B/A is 0.97 but no significant difference is found between CE and LB interpretations, while in (4) where B/A is 0.91 or in (6) where B/A is 0.98, CE-interpretation is significantly higher. If the interpretation of CE and LB are assumed to be determined by B/A value, these results are contradictory. Similar results are also found in (9) and (25). It is probably natural to regard these as atypical cases with extremely low values in A and B.

3. Concluding Remarks

On the basis of the abovementioned results, the following points were made clear:

(1) In ambiguous sentences such as used in the experiment, pitch contours function as the decisive factor in the interpretation of the sentence structure.

(2) In the relative clause constructions, the interpretation differs by the relative height in pitch assigned to the relative clause. When the pitch assigned to the relative clause is the same or higher than that of the preceding portion of the sentence, it tends to be interpreted as center-embedded, otherwise as left-branching.

Table 2. Ratio of the interpretation of LB and CE for each stimulus.

TEST SENTENCE 1

B \ A		0.5			1.0	
		LB	CE		LB	CE
0.0	(1)	46.1	53.9	(2)*	72.4	27.6
0.5	(3)!	17.2	82.8	(4)!	32.0	68.0
1.0	(5)!	10.3	89.7	(6)!	14.2	85.8
1.3	(7)!	12.7	87.3	(8)!	12.7	87.3

TEST SENTENCE 2

B \ A		0.33			0.67			1.0			1.33	
		LB	CE		LB	CE		LB	CE		LB	CE
0.25	(9)*	67.2	32.8	(10)*	66.5	33.5	(11)*	85.2	14.8	(12)*	82.3	17.7
0.5	(13)	43.1	56.9	(14)	47.8	52.2	(15)*	69.1	30.9	(16)*	78.9	21.1
0.75	(17)!	34.5	65.5	(18)!	33.3	66.7	(19)	50.0	50.0	(20)*	70.1	29.9
1.0	(21)!	28.2	71.8	(22)!	25.6	74.4	(23)!	29.1	70.9	(24)	46.6	53.4

TEST SENTENCE 3

B \ A		0.5			1.0			1.5			2.0	
		LB	CE		LB	CE		LB	CE		LB	CE
0.25	(25)*	67.0	33.0	(26)*	67.2	32.8	(27)*	72.1	27.9	(28)*	61.1	38.9
0.5	(29)!	30.9	69.1	(30)	43.8	56.2	(31)	54.4	44.6	(32)	53.9	46.1
0.75	(33)!	16.7	83.3	(34)!	24.0	76.0	(35)!	30.4	69.6	(36)	46.6	53.4
1.0	(37)!	15.2	84.8	(38)!	19.1	80.9	(39)!	27.0	73.0	(40)!	30.4	69.6

(3) Slight differences are seen among the values in B/A in the three test sentences, which affect the distinction of CE-interpretation and LB-interpretation. These differences are probably due to the pitch accents of the words that constitute the three test sentences.

(4) Atypical interpretations were obtained when the values of A and B were extremely low, such as in simuli (1), (9) and (25). In the experiment, when an extremely low value was used as the coefficient in B, then the height of B decreases, while the base line stays the same. The rise of the contour in B, therefore, becomes extremely limited. This seems to be the cause of the atypical interpretations.

More detailed examinations are necessary for (3) and (4) above.

The problem of pauses in relation to the comprehension of complex sentences will be left for further studies.

Acknowledgment

We wish to express our appreciation to Professor Shuzo Saito of the Research Institute of Logopedics and Phoniatrics, University of Tokyo, for his valuable advice in the preparation of the stimulus sentences.

References

Hadding, K. and M. Studdert-Kennedy (1973) Are You Asking Me, Telling Me, or Talking to Yourself? HASKINS LABORATORIES: Status Report on Speech Research SR-33.

Hakoda, K. and H Sato (1978) The Investigation of Prosodic Rules in Connected Speech, Transactions of Committee on Speech Research, Acoustical Society of Japan, No. S78-07.

Harada, S. I., T. Uyeno, H. Hayashibe and H. Yamada (1976) On the Development of Perceptual Strategies in Children: A Case Study on the Japanese Child's Comprehension of the Relative Clause Constructions, Ann. Bull. RILP, No. 10, 199-224.

Hirose, H., H. Fujisaki and K. Ohta (1979) Nihongo Heizyobun no Kihon Syuuhasuu Patan no Onkyooteki Tokutyoo (Acoustic Features of the Fundamental Frequency Contours of Declarative Sentences in Japanese), Transactions of Committee on Speech Research, Acoustical Society of Japan, No. S79-03.

Liberman, M. and Prince, A. (1977) On Stress and Linguistic Rhythm Linguistic Inquiry 8. 2.

Lieberman, P. (1967) Intonation, Perception, and Language. M. I T. Press

Pierrehumbert, J. (1979 a) The Perception of Fundamental Frequency Declination J. Acoust. Soc. Am. 66(2)

Pierrehumbert, J. (1979 b) Intonation Synthesis Based on Metrical Grids Speech Communication Papers Presented at the 97th Meeting of the ASA.

Uyeno, T., Hayashibe, H., and Imai, K. (1979) On Pitch Contours of Declarative, Complex Sentences in Japanese Annual Bulletin, RILP, 13.

Uyeno T., H. Hayashibe, K. Imai, H. Imagawa and S. Kiritani (1979) Hatuwa to Rikai ni okero Fukubun no Intoneesyon (Intonation of Complex Sentences in Production and Comprehension). Special Research Project of "Language," Ministry of Education, The Sawashima Group, Report No. 54-6.