Ann. Bull. RILP (1988) No. 22, 173-181

SOME REMARKS ON PROSODY IN READING A STORY IN JAPANESE*

Tazuko Uyeno** and Hiroshi Imagawa

1. Introduction

This study aims to grasp how pitch patterns and pauses are used in a discourse, specifically in reading a story in Japanese. It is assumed that pitch patterns, which are basically the realization of prosodic units in a sentence, may also play important roles in a discourse along with intrasentential and intersentential pauses.

Intrasententially, pitch patterns may function as indicators of phrases and clauses, and pauses may contribute to marking of their boundaries. In a discourse, where utterances require longer time span, it is natural to expect more pauses to be used for rebreathing. Pauses are thus inevitable for the continuation of utterances and, at the same time, they may differ in length along with the roles they play.

In this study, the first portion of an old Japanese story known as *Momotaroo*, "The Peach Boy," (hereafter, the story) was taken as the material to be examined. The story was read three times in a natural speech style by a professional male announcer who is a speaker of the Tokyo dialect, and recorded in a sound-proof room. The last reading was used for the extraction of pitch patterns and the identification of pauses employed.***

2. Materials and Methods

2.1. Recorded Materials

The story recorded is the following, which consists of 15 segments, J-1) ~ J-15). Each segment contains one or more sentences. In each segment, the locations of the pauses given are marked with parentheses in which the length of pauses employed are indicated in milliseconds, while the total length of time each segment required in reading is shown in square brackets in seconds.

---

* This study was in part supported by a Grant-in-Aid for scientific research (No. 410207) from the Japanese Ministry of Education, Science and Culture.

** The National Language Research Institute

*** The story was taken from the book *Nihon Mukashibanashi (5)*, Tsubota, Joji, Kaiseisha Bunko, (1978) pp.11-12.

English equivalents of these segments follow with the corresponding segment numbers E-1) ~ E-15), below, for reference.

J- 1) Mukasi, mukasi, aru tokoro ni (320 msec) oziisan to obaasan to ga arimasita.[4.7 sec] (2,110 msec)
J- 2) Oziisan wa, yama e takigi o tori ni, (760) obaasan wa, kawa e sentaku ni ikimasita.[5.5] (2,110)
J- 3) Obaasan ga, sentaku o site imasu to, (416) kawakami no hoo kara, (192) momo ga, (480) tunbura, tunbura to nagarete kimasita. [6.8] (2,870)
J- 4) Obaasan ga, sore o hirotte, hitokuti, kazitte mimasu to, (384) nantomo ienai kurai umai node, [5.3] (820)
J- 5) "Kore wa, umai momo da. (832) Hitotu, uti no oziisan ni mo motte kaette, tabesasete ageyoo. (1,280) Umai momo wa kotti e koi. (640) Nigai momo wa, atti e ike." [11.0] (1,500)
J- 6) Obaasan ga, soo yobikakemasu to, (512) hitotu no ookina momo ga. (320) tunbura, tunbura, nagareyotte kimasita.[6.4](2,510)
J- 7) "Kore wa mata umasoona momo da." (512) to, obaasan wa, ookina momo o hirotte, (128) motte kaerimasita. [5.8] (3,890)
J- 8) Bangata ni naru to, oziisan ga, (160) takigi o seotte, kaette kimasita. [4.1] (2,200)
J- 9) "Obaasan, obaasan, (320) ima, modotta zo." [2.3] (1,220)
J-10) "Oziisan desu ka. (384) Okaerinasai. (1,280) Kyoo wa, kawa de, ookina momo o hirotte, (96) todana ni simatte arimasu kara." [7.5] (2,640)
J-11) Obaasan wa soo iinagara, (192) todana kara momo o dasite, (320) manaita no ue ni okimasita. [5.1] (2,660)
J-12) Sate momo o waroo to site, (128) obaasan ga, hootyoo o atemasita tokoro, (576) momo ga, hitorideni, pokatto warete, (448) naka kara, kawaii otokonoko ga dete kimasita. [9.5] (2,090)
J-13) Sosite, hogyaa, hogyaa to nakimasita. [2.3] (2,250)
J-14) Oziisan to obaasan wa, bikkurisite, (640) "Ya, ya, (320) kore wa taihen da." (416) to sibaraku, (192) ooawate ni awate-masita. [7.1] (2,740)
J-15) Sawagi mo sizumaru to, (64) oziisan wa, (960) "Momo kara umareta kodomo da kara, (380) namae wa Momotaroo to sinakeryaa." (1,664) Soo itte, (192) Momotaroo to na o tukemasita. [11.5]

E- 1) Long, long ago, there lived an old man and old woman.
E- 2) The old man went into the mountains to gather firewood, and the old woman went to the stream to do her washing.
E- 3) When the old woman was washing at the stream, she saw a peach come floating down toward her.
E- 4) She picked it out of the water, took a bite of it, and found it indescribably delicious.
E- 5) "Oh, this peach is so delicious! I wish I had one to take home for my husband. Sweet peach, come this way! Sour

peach go away!"
E- 6) When the old woman began calling so, a big peach came floating down the stream.
E- 7) "How sweet it looks!" she said, and she took the peach home.
E- 8) In the evening, the old man came home carrying firewood on his back.
E- 9) "I'm back," (he said.)
E-10) "Is that you? Welcome home. Today, I found a big peach in the river. It's in the cupboard."
E-11) As she spoke, the old woman took the peach from the cupboard and put it on the cutting board.
E-12) Just when she was about to put a nife to the peach, the peach suddenly split open and a cute little boy appeared.
E-13) Then he started to cry.
E-14) "Oh, my gosh," said the old man and the old woman in surprise, and they did not know what to do.
E-15) After the initial excitement abated, the old man said, "We'll call him 'Momotaroo (Peach Boy)' because he was born from a peach." So saying, he named the boy 'Momotaroo.'

As the English equivalents suggest, most of the recorded segments consist of one sentence even though their length may vary. Exceptions are segments J-4) and J-5), which together constitute one sentence, segment J-10), which is a quotation containing three sentences, and segment J-15), which can be regarded to contain two sentences.

In terms of contents of the story, the segments can be divided into two groups. The first group contains the segments J-1) ~ J-7), and refers to the introduction of the old man and old woman, and what happened at the stream when the old woman was doing her washing. The second group includes the rest of the segments, J-8) ~ J-15), and refers to what happened after the old man came home in the evening.

2.2. Methods

Using the recorded materials described above, the fundamental frequency of the speech signal for each segment was analyzed using the interactive laboratory system. The speech signals were lowpass-filtered to 5 kHz, sampled at 10 kHz and digitized in 10 bits. PARCOR analysis was performed 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 a period of 38.4 msec. In the case where it was judged 'voiced', the pitch period was determined by detecting the location of the maximum peak in a range of 2.5 msec to 16 msec, in the autocorrelation function of the residual wave.

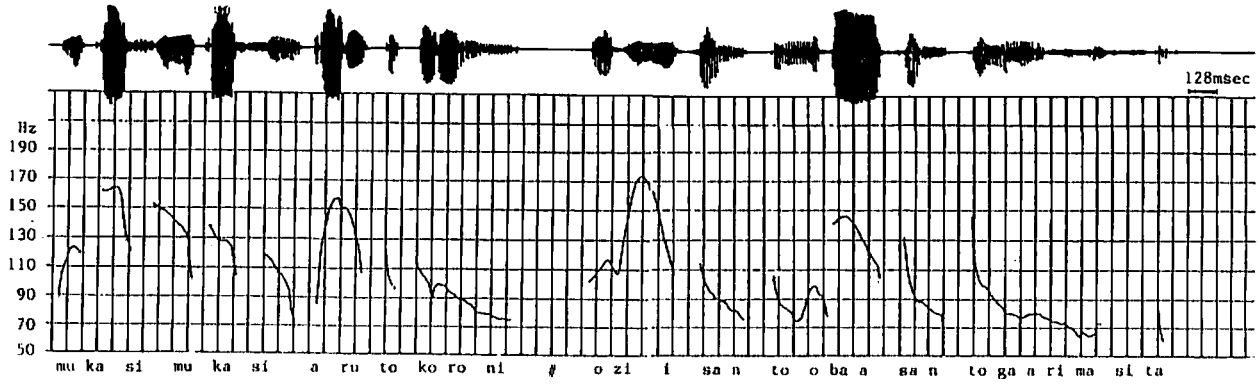

Fig. 1. Extracted pitch curves for measurement of fundamental frequency of segment J-1). # indicates where pauses are employed intrasententially.

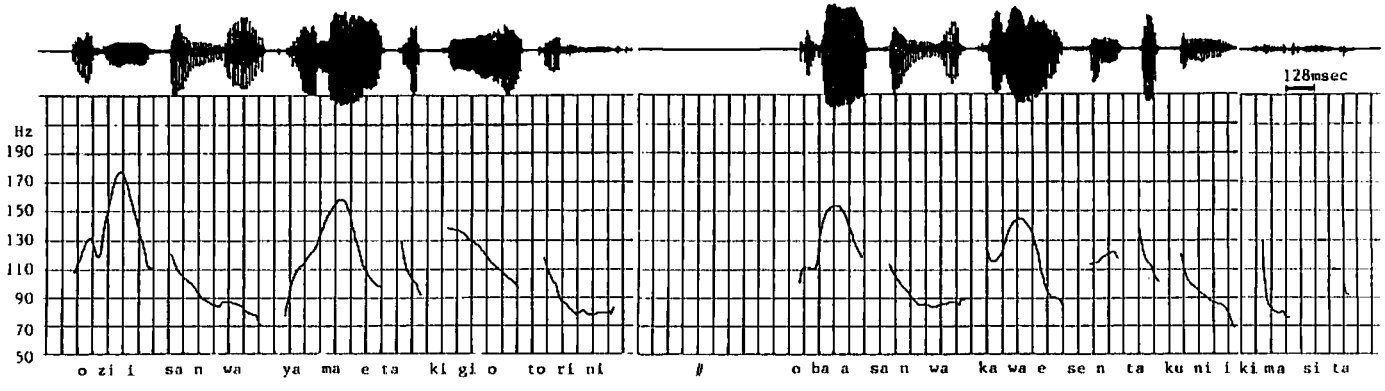

Fig. 2. Extracted pitch curves for measurement of fundamental frequency of segment J-2). # indicates where pauses are employed intrasententially.

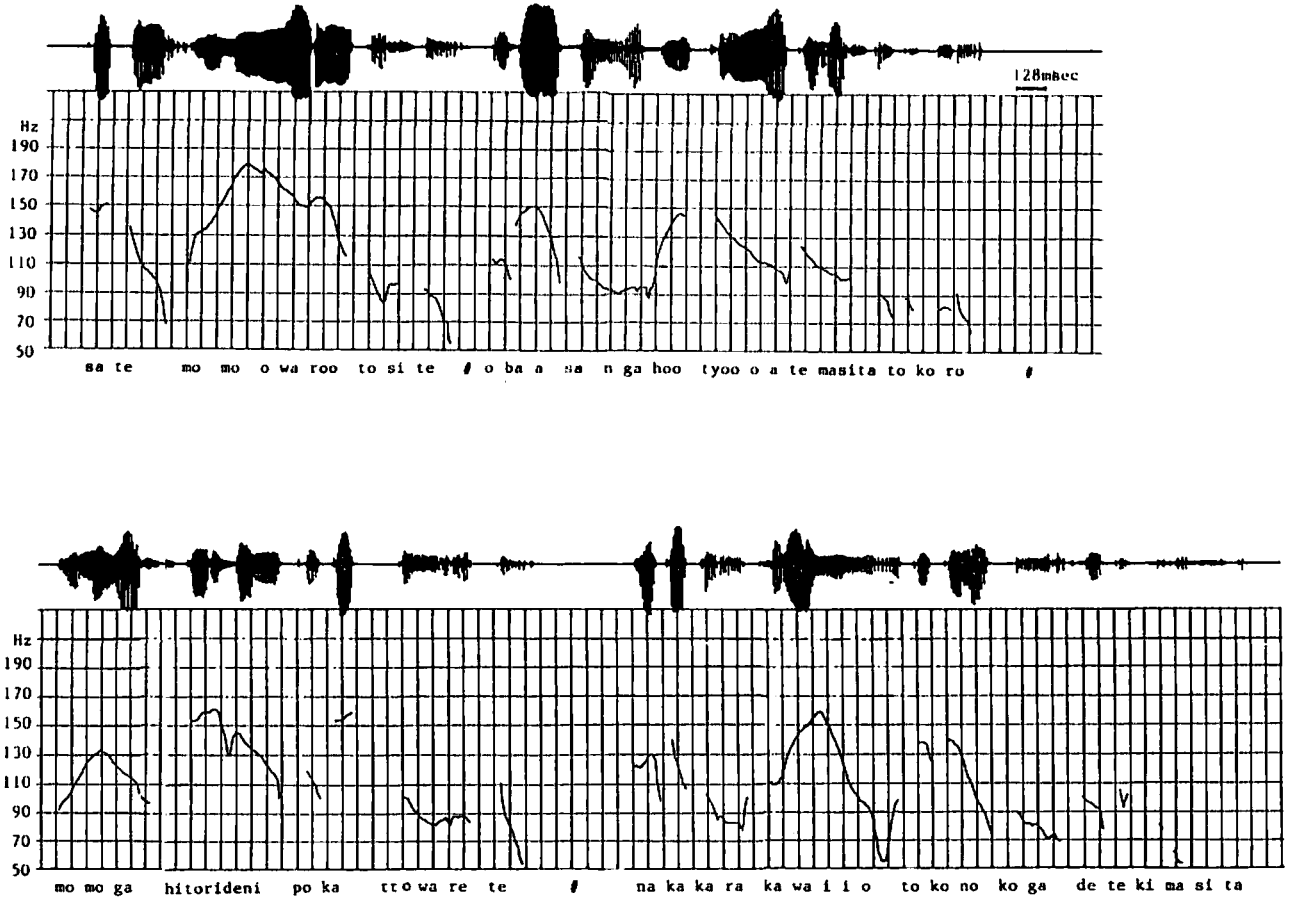

Fig. 3. Extracted pitch curves for measurement of fundamental frequency of segment J-12). # indicates where pauses are employed intrasententially.

In examining the pitch patterns of each segment, the pitch contours extracted through the analysis of fundamental frequency, as exemplified in Figs. 1 ~ 3, were used. In measuring the length of intrasentential and intersentential pauses, both the pitch contours and the corresponding speech envelopes were used.

3. Results and Discussion

3.1. Pitch Patterns

The extracted pitch contours for each segment were examined intrasententially and intersententially. As the pitch contours in Figs. 1 and 2 show, pitch curves peaking higher at the onset of an utterance and gradually moving lower toward the end of the utterance which is followed by a pause, constitute a unit. In Fig. 1, the pitch contours of the utterance *mukasi mukasi aru tokoro ni* (long, long ago) comprise the first unit, and those of the succeeding utterance *oziisan to obaasan to ga arimasita* (there lived an old man and an old woman) constitute the second unit. In Fig. 2, likewise, the pitch contours of the utterance *oziisan wa yama e takigi o tori ni* (the old man went into the mountains to gather firewood) constitute the first unit, while those of the succeeding utterance *obaasan wa kawa e sentaku ni ikimasita* (the old woman went to the stream to do her washing), following a pause, form the second unit. Within a unit, pitch curves correspond to the comprising phrases or clauses. In the first unit in Fig. 1, the first curve corresponds to the phrase *mukasi mukasi*, and the second curve to the phrase *aru tokoro ni*.

The peak of the top-most curves may reach as high as 190 Hz, while the lowest curves decline close to 50 Hz, and the rest stays in between. It is common to find that the first curve in a unit peaks the highest and the succeeding curves in turn move lower toward the end. (cf. Figs. 1 and 2) There are also cases where the succeeding curves stays as high as or higher than the preceding ones at peaks. In segment J-12) of Fig. 3, the curve corresponding to the clause *momo o waroo to site* (she was about to put a knife to the peach) in the first unit, peakes higher than the first curve corresponding to the phrase *sate* (when). The same is true of the curves corresponding to the phrases *hitorideni pokatto warete* (suddenly split open) in the third unit and *kawaii otokonoko ga* (a cute little boy) in the fourth unit. They peak higher than their respective preceding phrases, *momo ga* (the peach) and *naka kara* (from inside). These can be taken as cases where the lower pitch curves are assigned to an utterance of initial phrases due to either their grammatical or discoursal roles they play, so that the succeeding phrases with higher pitch curves are treated conspicuously in conveying the semantic contents each unit carries.

In the case of *sate*, the assignment of the lower pich curve may be due to its grammatical role being as a conjuntive. In the case of *momo ga*, which is a noun phrase carrying a new information, was followed by the adverbial phrases *hitorideni pokatto* which contains the most surprising information to provide. In the case of utterance initial adverbial phrase *naka kara*, it was followed by a noun phrase *kawaii otokonoko ga*, which is the subject of the clause.

Similar cases found in other segments are the following italicized phrases:

J- 5) hitotu, *uti no oziisan ni mo* motte kaette tabesasete ageyoo.
J- 7) obaasan wa, *ookina momo o hirotte*
J-10) kyoo wa kawa de *ookina momo o hirotte*
J-13) sosite, *hogyaa, hogyaa to* nakimasita
J-14) oziisan to obaasan wa *bikkurisite*
J-15) namae wa *Momotaroo to* sinakeryaa

In J-7), J-14), and J-15), initial phrases are topic noun phrases and the following italicized phrases provid comments. The initial phrases in J-5) and J-10) are adverbials, while that in J-13) is a conjunction.

A unit of pitch curves, as explained above, are marked with an initial and final pauses. In the materials examined, pauses are employed definitely for sentences, for many of clauses that constitute matrix sentences, and for some noun phrases and adverbials which also constitute matrix sentences. High rise of the initial peak of the pitch contours, shown in Figs. 1 ~ 2 is most commonly observed pattern. Within a unit of pitch curves some phrases which are not necessariry placed initially get high peaks due to their grammatical and discoursal roles, as obserbed in Fig.3. In some cases, phrases alone are marked by pauses, as in the cases of *kawakami no hoo kara* and *momo ga* in J-3), and pitch curves with high peaks are assinged to them. In these cases, phrases alone constitute units and by these arrangements, they may function as carriers of inportant semantic entity wihin a given discourse.

3.2. Pauses

According to the length of pauses obtained, pauses are divided into five groups: A (less than 300 msec), B (over 300 and less than 1,000 msec), C (over 1,000 and less than 2,000 msec), D (over 2,000 and less than 3,000 msec), and E (over 3,000 and less than 4,000 msec).

Most intersentential pauses belong to Group D. One exception, which belongs to E, was employed at the end of segment J-7), where

the first half of the story ends. Other exceptions are those which belong to C which were found at the end of quotations (in J-5, J-9 and J-15), and those which belong to B and were found intersententially within quotations (such as in J-5, J-7 and J-14). Length of pauses in narration parts and quotation parts are thus differentiated.

Many intrasentential pauses belong to B, which were employed at the end of coordinately conjoined clauses (in J-2, J-11 and J- 14) and adverbial clauses (such as in J-3, J-4, J-6 and J-12).

Short pauses which belong to A are used after a subject noun phrase (in J-8) or adverbial phrase (in J-3). They are also found after clauses such as coordinately conjoined clauses (in J-7, J-10 and J-12), and adverbial clauses (in J-11).

In some cases, relatively long pauses are used after noun phrases. See the pauses in J-3) after the subject noun phrase *momo ga* (a peach), and in J-15) after the topic noun phrase *oziisan wa* (the old man). In J-3), verb phrases describing how a peach floated down toward the old woman follow the noun phrase. In J-15), a quotation follows. Here, the longer pauses indicate clearer boundaries and as a results they function in drawing more attention to what follows. In J-3), even a higher pitch curve than that of the preceding noun phrase is assigned to the onset of verb phrases folloing the pause.

The distribution of pauses and variations in length described above suggest that intersentential boundaries in narrations are normally marked by longer pauses which belong to D. When extra-long pauses which belong to E are used, boundaries become more conspicuous, and they can function as signs of changes in a discourse. Intrasententially, pauses which belong to B are normally used in clause boundaries of matrix sentences, and shorter pauses which belong to A are used after noun phrases, adverbials, and embedded clauses when they are needed.

4. Concluding Remarks

The study covered only a very limited portion of a story in reading. The pitch contours and pauses obtained from the materials could clarify to an extent how pitch countours and pauses interplay their roles in conveying the unity of utterances, marking of important semantic information in a discourse, and presenting grammatical boundaries by all of which the given information through reading would be reconstructed and comprehended by listerners. Needless to say, further studies are necessary using various data in discourse to elucidate roles pitch contours and pauses play.

References

1) Sugito, M: Kutooten to. hatuwa ni okeru renzoku to kugiri - tenki yohoo no roodoku ni kanshite - . Osaka Shoin Daigaku Ronshu, 22, 1-7, 1985.
2) Sugito, M.: Gengo koodoo ni okeru onsei - kookateki na roodoku e no apurooti - . Nihongo Kyooshi yoo Sankoosyo I, Gengo Koodoo to Nihongo Kyooiku, Bonjin Co., Ltd., 41-261, 1585.
3) Sugito, M: Akusento no toraekata. Hayashi, S (Ed.) Oyo Gengogaku Koza 2, Gaikokugo to Nihongo. Meiji Shoin, 76-96, 1987.
4) Sugito, M: Pause and intonation. Aspects of Text and Context - An Analysis of Conversational Texts. The National Language Research Institute, Report 92. The National Language Research Institute, 107-138, 1987.
5) Uyeno, T., H. Hayashibe and K. Imai: On pitch contours of declaratiave, complex sentences in Japanese. Ann. Bull. RILP, 13, 175-187, 1979.
6) Uyeno, T., H. Hayashibe, K. Imai, H. Imagawa and S. Kiritani: Comprehension of relative clause construction and pitch contours in Japanese. Ann. Bull. RILP, 14, 225-236, 1980.
7) Uyeno, T., H. Hayashibe, K. Imai, H. Imagawa and S. Kiritani: Syntactic structures and prosody in Japanese - a study on pitch contours and the pauses at phrase boundaries -. Ann. Bull. RILP, 15, 91-101, 1981.