LENGTH-DEPENDENT PROSODIC PHRASING IN JAPANESE SENTENCES

Yoko Mori

Doshisha University

sk109695@mail.doshisha.ac.jp

ABSTRACT

This study examines the effect of length of prosodic words (PWs) on the formation of prosodic phrases (PPs) in Japanese sentences. 15 Tokyo-Japanese speakers read test sentences comprised of 3 nonsense prosodic words (PW1-PW3) followed by a verb phrase (PW4), all initially accented. PW1-PW3 varied in length (2-5 moras). The results showed that the sentence-final two PWs were consistently tonally united to form a single PP. When PW2 was 2 or 3 moras, in 36-65% of the utterances, PW3 showed significant upwards pitch resetting, dividing the four PWs into 2PP; downstep was observed in 31-32% of utterances. PW2 with 4 or 5 moras showed more downstep (35-44%) and grouping of the last three PWs into one PP (29-33%). These varied frequencies of prosodic phrasing patterns suggest that the length of all component PWs in a sentence interacts one another and affects the phrasing patterns of PWs.

Keywords: prosodic phrasing, length effect, grouping of prosodic words

1. INTRODUCTION

Kubozono [1] reports that in uniformly left-branching phrases consisting of four accented words, "the third minor phrase is realized as high as (or even higher than) the second minor phrase", dividing the four constituents into a pair of sub-groups. He claims that the principle of rhythmic alternation underlies this phenomenon. Selkirk et al. [3] report that in Japanese sentences the degree of initial pitch rise of minor phrases is affected by the phonological length (number of moras) of the target phrases, as well as the syntactic structure.

Although previous studies have shown syntactic, semantic, phonological effects on intonational phrasing in Japanese, the current study attempts to examine exclusively length effects on the prosodic phrasing of Japanese sentences from the empirical data gained by using nonsense words. Specifically, I attempt to provide evidence for the formation of prosodic phrases with about four-to-ten moras, in which two or more PWs are combined and the initial pitch rise of the following phrases is either attenuated or deleted.

In this paper the unit I attempt to propose is referred to as a "prosodic phrase (PP)" in the sense that it is an upper-level unit above a prosodic word (PW) in the prosodic structure of Japanese. Prosodic phrasing may be realized not only by pitch but also pausing or final lengthening. However, in Japanese read sentences with no punctuation marks, phrase-final lengthening is not consistent, and pausing is also arbitrary. Thus, in this study, the patterns of prosodic phrasing is examined solely by pitch movements, using only initially-accented words and asking subjects not to insert a pause in a test sentence. The phonetic features of the PP observed here were similar to those of major phrases: it is the domain where (total) downstep or tonal merger takes place in the following components whereas the phrase-initial PW shows a substantial initial pitch rise. We hope to see how the varying numbers of moras in the three PWs interact each other and affect the prosodic phrase formation of the test sentences.

2. EXPERIMENT

2.1. Experimental materials

In this study, nonsense words were used since in sentences made up of real words, speakers may place focus on specific words without knowing. 18 test sentences were prepared as shown in (1). The three nonsense nouns (N1, N2, and N3) are repetitions of /ne/, /na/, and /ma/, respectively, all initially accented. N1 and N2 are followed by the particle /no/ that indicates possession and turns the nouns into modifiers. N3 is followed by subject marker /ga/, forming a subject phrase with the preceding phrases. All three phrases varied in length from two to five moras. The subject phrase is followed by a three-mora verb phrase (VP) *mi'eta* 'could be seen'. The three phrases and VP are called PW1, PW2, PW3, and PW4, respectively. The second PW (PW2) is a target PW whose length ranges from 2 to 5 moras.

(1) Test sentences: (the symbol ' denotes the boundary between H and L tones.)

e.g. *ne'ne-no na'nana-no ma'ma-ga mi'eta*

(N1-no)	(N2-no)	(N3-ga)	(Verb-phrase)
(2, 3, 5μ)	(2 to 5 μ)	(2 to 4 μ)	(3 μ)
PW1	PW2	PW3	PW4

2.2. Speakers and procedures

15 Japanese (8 males and 7 females), who speak Tokyo Japanese, read the sentences. The ages of speakers ranged from 20 to 60.

Each test sentence was written on a card; nonsense words in *katakana*, the others in *hiragana* and *kanji*. The test sentences were mixed with filler sentences made up of nonsense words and presented to each speaker in random order. Speakers were asked to read each sentence so that the sentence would sound natural and comfortable to him or her. They were also told not to insert pauses in each sentence. Since the speakers may have difficulties in reading nonsense words that consisted of the same syllable, each speaker was asked to read each sentence three times in succession. If they misread any of the sentences, they were allowed to read it again and again until they could read it successfully at least two times. The total number of utterances analyzed would be 540 (18 sentences x 15 subjects x 2 utterances).

Recordings were analyzed with Sugi Speech Analyzer and the waveform, f0 contour, and spectrogram of each utterance were extracted. Since this experiment examines the intonation patterns of each sentence produced by each speaker, when there was a variation in intonation pattern in any sentence produced by a single speaker, each of the different intonation patterns was analyzed and used for further analyses. If one of the repetitions of each sentence had disfluencies or mistakes, the other two readings were analyzed. When the three productions of each sentence were almost identical, the second and third readings were analyzed. The maximum and minimum f0 values of each PW were measured for all the utterances. Some speakers could not produce some of the longer nonsense words correctly. This led to a loss of 29 tokens in total for further analyses.

2.3. Data analyses and classification of phrasing patterns

The maximum (Peak) and minimum (Valley) f0 values of the PWs were referred to as P1, P2, P3, P4, and V2, V3, V4, respectively. The f0 of V1 was excluded since some of the utterances produced by some speakers hardly showed any initial lowering for the sentence-initial PW. In order to eliminate interspeaker differences in voice pitch, these f0 values were normalized and represented in semitone using the f0 of P1 of each utterance by each speaker as a reference point. Next, the differences between each pair of adjacent peaks (P1-P2, P2-P3 and P3-P4) were calculated to examine upwards pitch resetting in PWs. Also, in order to see the degree of initial pitch

rise for each PW, the difference between each peak and its preceding valley was calculated (P2-V2, P3-V3, and P4-V4).

Based on these data, phrasing patterns were classified as shown in (2). As for sentence-final PW4 *mi'eta*, tonal merger or (total) downstep with the preceding PW3 was observed in almost all the utterances produced by all speakers, forming a single PP with PW3. Phrasing pattern (2)a is represented as [2 + 2] in which four PWs are grouped into two PPs, each of which consists of two PWs. This pattern was identified when P2 was lower than P1 by more than 1 semitone, and P3 was as high as or higher than P2 (Figure 1). The second pattern (2)b is "Downstep" where P2 was lower than P1, and P3 than P2, by 1 or more semitones and gradual lowering of the initial pitch rise over the utterance took place (Figure 2). When P2 was almost as high as or higher than P1, and P3 was lower than P2 by more than one semitone, the last three PWs were assumed to form a single PP [1 + 3]. When the difference between P2 and P3 was within 1 semitone and P2 was as high as or higher than P1, the utterance was assumed to be divided into [1 + 1 + 2].

(2) 4 patterns of prosodic phrasing
a. [2 + 2] : P1>P2, P2≦P3 (P1-P2>1, P2-P3≦1)
   (PW1 + PW2) + (PW3 + PW4)

b. Downstep : P1 > P2 > P3   (P1-P2>1, P2-P3>1)
   PW1 + PW2 + PW3 + PW4

c. [1 + 3] : P1≦P2, P2 > P3 (P1-P2≦1, P2-P3>1)
   PW1 + (PW2 + PW3 + PW4)

d. [1 + 1 + 2] : P1≦P2, P2≦P3
                  (P1-P2≦1, P2-P3≦1)
   PW1 + (PW2) + (PW3 + PW4)

Figures 1 to 4 show each of the above phrasing patterns. F0 values in semitone are collapsed across all the utterances produced by all the speakers. Error bars indicate standard errors above and below the mean.

3. RESULTS AND DISCUSSIONS

3.1. Tonal merger of the final two PWs

In the total of 511 utterances produced by all 15 speakers, the sentence-final VP *mi'eta* tended to be merged with the preceding subject-final PW3, forming a single PP with PW3, ranging from 5 to 7 moras. The last PW showed almost no initial pitch rise; the average pitch of 511 utterances is -8.77 for

V4 and -8.30 for P4 with standard deviation of 2.96 and 2.66, respectively.

Figure 1: Mean maximum and minimum f0 in semitone of the 115 utterances that showed [2 + 2] phrasing for 2-mora PW2
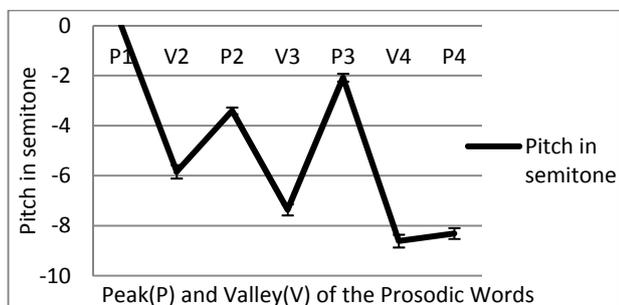

Figure 2: Mean maximum and minimum f0 in semitone of the 47 utterances that showed Downstep for 4-mora PW2
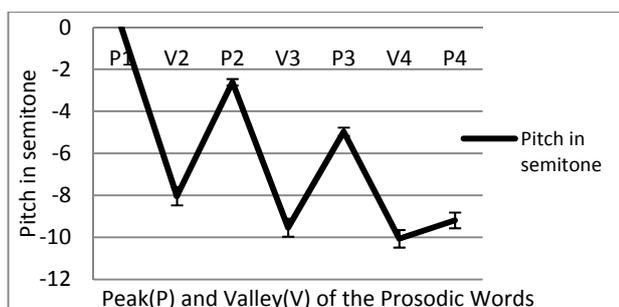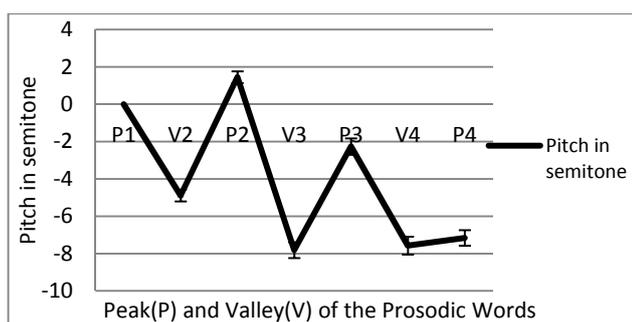

Figure 3: Mean maximum and minimum f0 in semitone of the 36 utterances that showed [1 + 3] phrasing for 5-mora PW2

3.2. PW2 with 2 or 3 moras

Table 1 shows the number of occurrences of each prosodic phrasing pattern of each sentence by the 15 speakers. In the case of PW2 with 2 moras, 64.6% of the four PWs are divided into two PPs, each consisting of two PWs [2 + 2]. 30.9% of the utterances show downstep. The other two patterns of phrasing are scarce.

Figure 4: Mean maximum and minimum f0 in semitone of the 14 utterances that showed [1 + 1 + 2] phrasing for 3-mora PW2
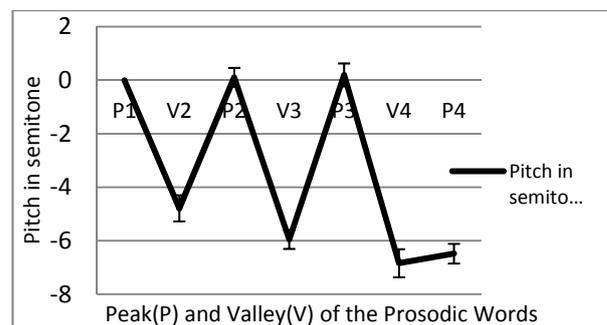
Looking closer at Table 1, two more interesting tendencies are observed. First, PW3s with 3 or 4 moras tend to show an initial pitch rise and form another PP with PW4s. This tendency leads to division into 2PWs and 2PWs (2+2) as in sentences (2+2+3+3), (2+2+4+3) and (5+2+4+3). In the sentence of (5+2+4+3) moras, PW1 is the longest. However, perhaps the total of 12 moras may be a little too long for the favorable PP size. Thus, it is mostly divided into 2PPs. The same phenomenon is observed in the sentence of (5+3+4+3) with 3-mora PW2, where it tends to be divided into (5+3) and (4+3). However, in sentences made up of (3+2+2+3) and (5+2+2+3) moras, the first longest PW initiates a PP and downstep takes place in the following PWs. Interestingly, when the first three PWs have only two moras (2+2+2+3), the majority of the utterances show [2 + 2] phrasing.

These findings suggest that not only the length of the target PW but also the length of all the component PWs affect the formation of PPs in the sentences.

Table 1. Mora counts of component PWs of each sentence with 2-to-3 mora PW2 and frequencies of each prosodic phrasing pattern

No. of moras of PWs	[2 + 2]	Downstep	[1 + 3]	[1 + 1 + 2]
2+2+2+3	**21**	8		1
2+2+3+3	**27**	3		
2+2+4+3	**29**	1		
3+2+2+3	6	**20**	4	
5+2+2+3	9	**18**	1	2
5+2+4+3	**23**	5		
Sum (%)	**115 (64.6)**	55 (30.9)	5 (2.8)	3 (1.7)
2+3+2+3	2	**12**	**14**	2
2+3+4+3	**14**	1	5	10
5+3+2+3	4	**19**	4	1
5+3+4+3	**21**	5		1
Sum (%)	**41 (35.7)**	37 (32.2)	23 (20)	14 (12.2)

Similar tendencies are more obviously observed for 3-mora PW2. In a sentence of (2+3+2+3), 3-mora PW2, which is longer than the preceding and following PWs, shows an initial pitch rise and forms

a PP with the following PWs in 14 out of 30 utterances [1 + 3]. When PW3 is 4-mora long, PW3 initiates another PP, resulting in more [2 + 2] phrasing patterns as in (2+3+4+3) and (5+3+4+3). Downstep is most frequent in (5+3+2+3) in which the PW1 is the longest of all the 4 PWs.

3.3. PW2 with 4 or 5 moras

When PW2 is 4 or 5 moras in length, the percentage of [2 + 2] division decreases, whereas that of [1 + 3] increases. As can be seen in Table 2, the typical phrasing patterns are quite similar both for 4 and 5 mora PW2. When PW2 is the longest in a sentence (e.g. 2+4+2+3, 2+5+4+3), PW2 initiates PPs dividing 4 PWs into 1PW and 3 PWs [1 + 3]. When the longest PW is at the head of the sentence, the sentence shows downstep as in (5+4+2+3) (5+5+2+3) or division into two PPs when PW3 is long enough as in (5+4+4+3) and (5+5+4+3).

Table 2. Mora counts of component PWs of each sentence with 4-to-5 mora PW2 and frequencies of each prosodic phrasing pattern

No. of moras of PWs	[2 + 2]	Downstep	[1 + 3]	[1 + 1 + 2]
2+4+2+3	4	9	15	
2+4+4+3	4	8	11	2
5+4+2+3	7	18	4	
5+4+4+3	13	12	1	
Sum (%)	28 (25.9)	47 (43.5)	31 (28.7)	2 (1.9)
2+5+2+3	4	4	19	2
2+5+4+3	4	5	15	6
5+5+2+3	3	17	2	2
5+5+4+3	13	12	0	2
Sum (%)	24 (21.8)	38 (34.5)	36 (32.7)	12 (10.9)

The initial pitch rise of the target PW2 with four or five moras may be partly due to the "global, lookahead effect" as Selkirk et al.[3] points out in Japanese words. That is, the longer a PW, the higher the initial pitch rise since speakers expect the pitch lowering over the following syllables/moras. However, the results from the current experiment show that the frequencies of the initial pitch rise of the target PW2 depend on not only on the length of the target PWs but also their environment. That is, the relative length of all the component PWs in the same sentence interacts with one another and affects the presence or absence of the initial pitch rise of the target PW2.

4. SUMMARY AND CONCLUSIONS

The findings from the current experiment can be summarized as follows:

- The last two prosodic words (PWs) in the test sentences have a strong tendency to be tonally merged to form a single prosodic phrase (PP) of 5 to 7 moras. This finding is worth noticing since the last PW is an initially accented verb phrase and generally predicted to form a minor phrase on its own with an initial pitch rise.
- When the second PW is 2 or 3 moras, the most frequent phrasing pattern is grouping into two PPs [2 + 2], where PW3 shows upward pitch resetting whereas PW2 indicates downstep in relation to PW1. The resulting PPs range from 4 to 8 moras.

    (PW1+PW2) + (PW3+PW4)

    However, when PW1 is the longest in a sentence and both PW2 and PW3 are 2 or 3 moras, the majority of the utterances show downstep over the sentence (Table 1).
- When the second PW is 4 or 5 moras, downstep is most frequently observed (43.5% and 34.5%, respectively). However, when PW1 is only 2 moras, PW2 tends to initiate a PP and PW3 shows downstep in relation to PW2. That is, the last three PWs group together to form a single PP [1 + 3]. This phrasing pattern takes place in about 30% of the utterances with 4- or 5-mora PW2 (Table 2).

    PW1 + (PW2+PW3+PW4)

Despite these varied distributions of prosodic phrasing patterns, one obvious finding from this study is that the last two PWs always combine with each other tonally. It was also found that not only the number of moras of the target PW but also those of the other PWs in the same sentence have a great effect on the prosodic phasing patterns in Japanese sentences. More specifically, a relatively long PW in a sentence tends to initiate a PP with the following PWs. This tendency leads to downstep or the prosodic phrasing of [1 + 3] when the first or second PW is the longer PW. However, when there are several long PWs and the resulting PP becomes too long, the PWs may divide into more PPs ([2 + 2] or [1 + 1 + 2]). When the third PW is a relatively long one, it initiates another PP, dividing the four PWs into a pair of 2PWs [2 + 2]. When there are no relatively long PWs as in a sentence of (2+2+2+3) moras, the sentence still tends to be divided into 2PPs.

All of these prosodic phrasings of Japanese sentences may result in nearly systematic occurrences of initial pitch rise in the test sentences. Mori [2] shows that Japanese words tend to group together to form PWs ranging from 2 to at most 5 moras. In this study, it was found that these PWs could further combine with one another to form PPs depending on their phonological length as well as the length of the other constituents of the sentence.

5. REFERENCES

[1]Kubozono, H. 1993. *Organization of Japanese Prosody* Tokyo: Kurosio Publishers

[2]Mori, Y. 2002. Lengthening of Japanese monomoraic nouns. *Journal of Phonetics*, 30(4), 689-708.

[2]Selkirk, E.; Shinya, T.; Kawahara, S. 2004. Phonological and phonetic effects of minor phrase length on f0 in Japanese. *Proc. Speech Prosody 2004*, Nara, Japan, 183-187.