Pitch Pattern Alternation in Goshogawara Japanese: Evidence for a Prosodic Phrase above the domain for Downstep

Yosuke Igarashi [1]

[1] Japan Society for the Promotion of Science, JPN, National Institute for Japanese Language, JPN

yosuke.igarashi@kokken.go.jp

Abstract

The lexically accented words in Goshogawara Japanese can be realized in either of the two surface pitch patterns. The pitch pattern is said to alternate regularly, depending on the phrasing structure of an utterance. The organization of prosodic phrasing of this dialect, however, has been little investigated and thus it remains unclear what prosodic phrase functions as the domain for the alternation.

This work determines the level of the phrase serving as the domain for the pitch pattern alternation. Is it hierarchically higher or lower than the domain for downstep? The experimental results reveal that the alternation does not take place at the prosodic boundary where downstep effect is blocked. The results provide evidence for the phrasing one-level above the downstep domain, whose existence was not evident in the model proposed for Tokyo Japanese.

Index Terms: Japanese dialects, Goshogawara, phrasing

1. Introduction

Goshogawara Japanese is a regional dialect spoken in Tsugaru region of Aomori Prefecture. Just as Tokyo (standard) Japanese, Goshogawara has words without pitch accent (unaccented words) and those with it (accented words). The accentedness is lexically determined. Fig. 1 illustrates fundamental frequency (F0) contours for an unaccented prosodic word *kimono-kara* 'from kimono' and accented word *uramí-kara* 'due to grudge' pronounced in isolation. (The accented syllable will be designated by acute accent.)

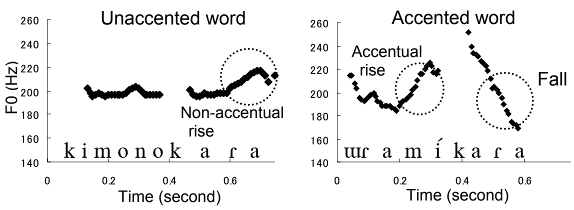

Figure 1: *F0 contours for unaccented and accented words in Goshogawara: 'kimono-kara' and 'uramí-kara'. The accented word is in non-connective form.*

Unaccented words always have a small rise in the word-final syllable (Fig. 1, left). The pitch pattern of accented words is more complicated. In *uramí-kara* (Fig. 1, right), the accented syllable *-mí-* exhibits a rise. The rise marks the accented syllable, and thus this dialect is known to have rising pitch accent (aka 'ascending kernel') [3, 11]. Notice that there is a *fall* in the final syllable. This fall is relevant to 'pitch pattern alternation' which will be discussed below.

Pitch pattern alternation is a phenomenon where a surface pitch pattern of accented words regularly alternates between two forms, depending on the *phrasing* structure of the utterance. One of the pitch patterns is *non-connective form* which has a fall in the word-final syllable (Fig. 1, right) and the other is *connective form* which has no final fall [1] [3,11,12,13]. Fig. 2 shows an F0 contour for *uramí-kara* in connective form. It can be seen that after the accentual rise, F0 does not fall in the word-final syllable. Instead, a flat high (or slightly rising) F0 prevails until the end of the word[2].

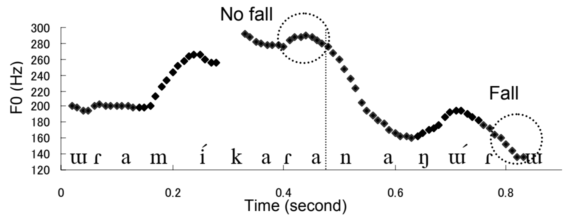

Figure 2: *F0 contour for an utterance 'Uramí-kara nagúru' ((He) punches due to grudge.). The word 'uramí-kara' is in connective form.*

As mentioned above, phrasing structure is claimed to determine the pitch patterns. Specifically, the connective form occurs when the word is at the non-final position of the 'phrase', while the non-connective form appears when the word is at the phrase-final position [12,13]. But what does the 'phrase' exactly mean? While pragmatic factors that affect the phrasing have been described in the past studies (continuation vs. finality) [12], the *phonological* organization of the prosodic phrasing remains unclear. If we assume, following the autosegmental-metrical theory of intonational phonology [2,7], that the prosodic phrasing is hierarchically organized, then we can ask a question: what level of prosodic phrase serves as the domain for pitch pattern alternation?

In Pierrehumbert and Beckman's model, Tokyo Japanese is proposed to have *three* levels of phrasing above prosodic word [9]. The smallest is *accentual phrase (α)* which contains at most one pitch accent. The constituent one-level above ***α*** is called *intermediate phrase (ι)*, which is the domain for pitch range specification. The pitch range compression effect caused by pitch accent, which is called catathesis or *downstep*, is blocked at the ***ι***-boundary. The largest phrase is *utterance*, which is the domain for the final boundary tone.

Let us assume that Goshogawara has a prosodic hierarchy similar to Tokyo. Since the prosodic phrase for pitch pattern alternation is larger than ***α***, it should be either ***ι*** or above it.

[1] Pitch pattern alternation was firstly reported by Uwano for Hirosaki Japanese (Aomori Pref.) [11] and Shizukuishi Japanese (Iwate Pref.) [12,13]. Hirosaki and Goshogawara belong to the Tsugaru dialectal group).

[2] It must be noted that the fall in non-connective form can not be analyzed simply as an utterance-final boundary tone (L%). One argument against this analysis is the fact that unaccented words never exhibit the word-final fall. Besides, there is good evidence to prove the fall is a property of the accented words *per se* [3].

The present work determines the level of the prosodic phrase for pitch pattern alternation. A production experiment will be performed to examine whether connective form appears immediately before the ***ι***-boundary. Since ***ι*** is the domain for downstep (which was proven to exist in Goshogawara as well [3]), the occurrence of connective form at the ***ι***-final position serves as evidence showing that the prosodic phrase for pitch pattern alternation is above ***ι***.

2. Methods

2.1. Speech materials

Three sets of sentences (1)–(3) were designed. They were originally written in standard Japanese, and then translated into Goshogawara Japanese by two native speakers (the participants). The sentences contained two factors which are reported to introduce an ***ι***-boundary for several Japanese dialects, i.e., right-branching syntactic boundary and focus[3].

Datasets (1) and (2) contained the sentences with different branching structures. It is known for Japanese that a right-branching syntactic boundary tends to block downstep effect, while a left-branching boundary does not [5,8]. For each Dataset, (b) had a right-branching syntactic boundary between the first and second words, and thus an ***ι***-boundary was expected to be inserted here.

(1a) [[[Midóri-no yané-no] hóteru] mi-éru].
green-GEN roof-GEN hotel see-POTENTIAL
'I see a hotel with a green roof.'

(1b) [[Midóri-no [okkína hóteru]] mi-éru]
green-GEN big hotel see-POTENTIAL
'I see a green big hotel.'

(2a) [[Aómori-no nama-bíːru] [tádade nom-éru]]
Aomori-GEN draft-beer free drink-POTENTIAL
'You can drink Aomori's draft beer free.'

(2b) [Aómori-de [nama-bíːru [tádade nom-éru]]]
Aomori-LOC draft-beer free drink-POTENTIAL
'In Aomori you can drink draft beer free.'

Datasets (3) contained the sentences with broad focus (3a), with narrow focus on the first word (3b), and with narrow focus on the second word (3b). The reported effects of narrow focus in Japanese are a pitch range reset in the focused word and a strong pitch range compression in the post-focal words [8,9]. The test sentences were in short dialogues, in which the focus fell on the appropriate word. Specifically, (3a), (3b) and (3c), respectively, were preceded by "What happened with him?", "Which local gang was he punched by?" and "Who in Inagaki was he punched by?". Since (3c) had focus on the second word, an ***ι***-boundary was expected to be inserted between the first and second word.

(3a) Inágaki-no inaka-yákuza-sa futak-áe-ta
Inagaki-GEN local-gang-DAT punch-PASSIVE-PAST
'(He) was punched by a local gang in Inagaki'

(3b) INÁGAKI-no inaka-yákuza-sa futak-áe-ta
Inagaki-GEN local-gang-DAT punch-PASSIVE-PAST

(3c) Inágaki-no INAKA-YÁKUZA-sa futak-áe-ta
Inagaki-GEN local-gang-DAT punch-PASSIVE-PAST

[3] The view that an F0 boost caused by these factors as an indicator of downstep reset in Japanese is recently questioned by some researchers [4,6]. Still, alternative accounts for the boost are far from uncontroversial. For a detailed discussion, see [15].

2.2. Measurements

In order to confirm that an ***ι***-boundary was inserted between the first and second prosodic words, Relative Peak Height (RPH) of the second word was calculated.

RPH = the peak F0 value of the second word / the peak F0 value of the first word

When RPH is around 1, the second peak is as high as the first, suggesting that pitch range is reset and hence ***ι***-boundary is inserted between the first and second words.

Also, Lowering Ratio (LR) was computed so as to decide whether a given prosodic word appeared in connective form or non-connective form. LR is the measure for F0 difference between the penultimate and final syllables of the word.

LR= the mean F0 of the final vowel / the mean F0 of the penultimate vowel

The value far below 1 indicates that there is a considerable lowering in the word-final syllable, meaning that the word appears in non-connective form.

2.3. Subjects and analysis procedures

Two female native speakers (K and M) of Goshogawara Japanese participated in the experiment. Both were 22 years old and spent their life in Goshogawara (0-14), Hirosaki (15-17) and Tokyo (18-22). Speakers read entire set of the translated sentences five times. The recordings were made using Marantz PMD 660 and saved onto a Compact Flash memory card at a 44.1 kHz sampling rate. Recorded materials were analyzed using the Praat software [1].

3. Results

3.1. Branching

Fig. 3 shows the F0 contours for Dataset (1). We see iterative downstep effects throughout the left-branching utterance (a): the peak of each word was lower than its preceding word. In right-branching utterance (b), on the other hand, the peak of the second word was as high as the preceding one.

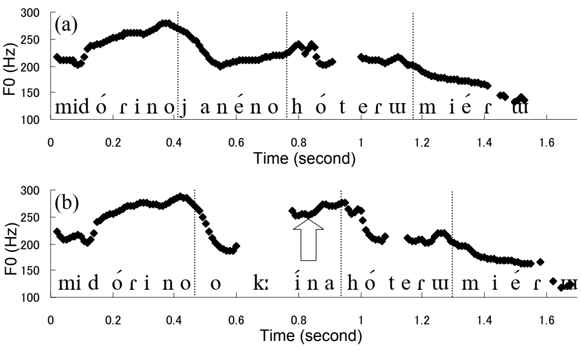

Figure 3: *F0 contours for Dataset (1). Produced by M.*

Figure 4 represents RPH. It was over 0.9 in (b) for both speakers, meaning that the height of second peak approached the first one. While there was a speaker-specific difference in RPH, general tendency was the same: for both speakers, RPH was larger in (b) than (a) and the second peak of (b) was as high as the first. A significant effect of branching on RPH was revealed by *t*-tests [for K, $t = -21.93$, $p < 0.001$, $df = 8$; for M, $t = -9.23$, $p < 0.001$, $df = 8$]. It was thus confirmed that in (b) pitch range was reset and hence an ***ι***-boundary was inserted at the beginning of the second word.

As for pitch pattern alternation, we see from Fig. 3 that all the words but the final appeared in *connective* form both in

(a) and (b). Importantly, the first word in (b) was realized in connective form, though it was at the final position of ***ι***.

Fig. 5 demonstrates LR of each word in (b). It can be seen that the LR of the final word was far below 1, indicating that there was a considerable lowering. The one-way analyses of variance (ANOVAs) with LR as the dependent variable and with word position (first, second, third, final) as the independent variable, for each speaker separately, revealed that there was a significant effect of word position on LR [for K, F (3, 16) = 27.98, p < 0.001; for M, F (3, 16) = 81.47, p < 0.001]. The results of *post hoc* Bonferroni tests are reported in Fig. 5. The results confirmed that pitch pattern alternation did not occur at the utterance-medial ***ι***-boundary.

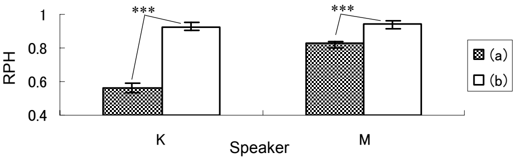

Figure 4: *Relative Peak Height (RPH) for Dataset (1). Error bars indicate SD. *, **, ***, respectively designate significance at 0.05, 0.01, 0.001 levels.*

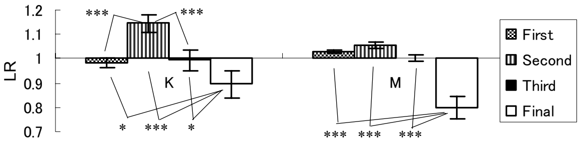

Figure 5: *Lowering Ratio (LR) for Dataset (1b).*

Fig. 6 shows the F0 contours for Dataset (2). We see that the peak of the second word was downstepped in (a), while downstep effect appeared to be blocked at the beginning of the second word, where there was a right-branching boundary.

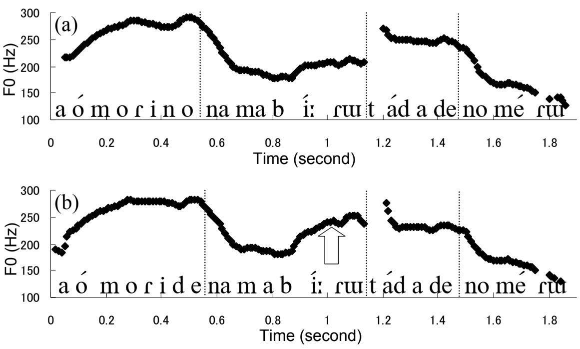

Figure 6: *F0 contours for Dataset (2). Produced by M.*

Means for RPH are presented in Fig. 7. I can be seen that a pitch range reset in (b) was not as clear as Dataset (1): mean RPH was 0.74 for K and 0.85 for M. Nevertheless, t-test revealed a significant effect of branching on RPH for both speakers [for K, t = -5.80, p < 0.001, df = 8; for M, t = -6.37, p < 0.001, df = 8]. While a question remained as to whether pitch range was indeed *reset*, it was significantly expanded in the second word in (b).

Visual inspection of the F0 contours tells that, just as in Dataset (1), only the final word was realized in non-connective form. From Fig. 8, showing LR for (b), we see that the final word exhibited a considerable lowering. One way ANOVAs with LR as the dependent variable and with word position as the independent variable revealed that there was a significant effect of word position for both speakers [for K, F (3, 16) = 44.32, p<0.001; for M, F (3, 16) = 25.27, p<0.001]. Thus, it was confirmed that the word realized in non-connective form was only the final one.

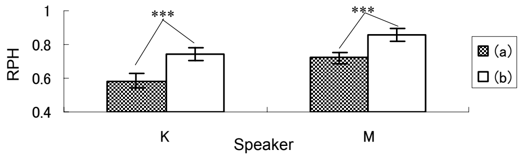

Figure 7: *Relative Peak Height (RPH) for Dataset (2).*

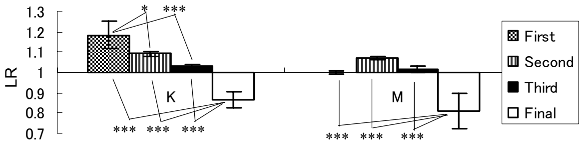

Figure 8: *Lowering Ratio (LR) for Dataset (2b).*

In summary, a right-branching syntactic boundary generally caused a pitch range reset and thus introduced an ***ι***-boundary. The word immediately before the ***ι***-boundary was realized in *connective* form, meaning that pitch pattern does not alternate at the final position of the utterance-medial ***ι***. The results for Datasets (1)-(2) revealed that the prosodic phrase for pitch pattern alternation is above ***ι***.

3.2. Focus

Fig. 9 depicts the F0 contours for Dataset (3). We observe iterative downstep effects in the utterance with broad focus (a), a pitch range compression of post-focal words in the utterance with narrow focus on the first word (b). In the utterance with narrow focus on the second word (c), we see a range reset in the second word followed by a post-focal range compression in the final word.

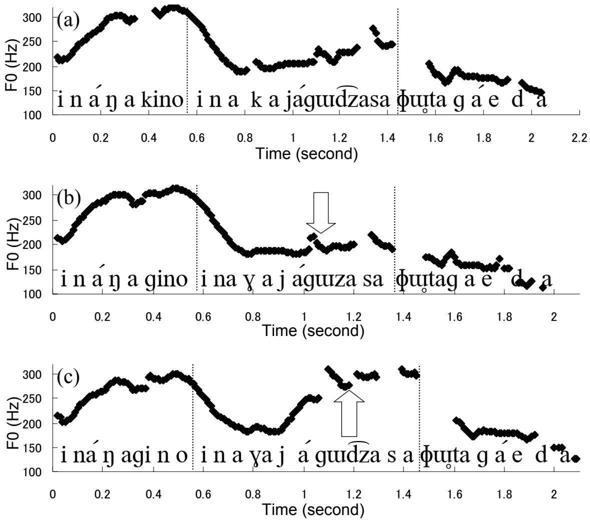

Figure 9: *F0 contours for Dataset (3). Produced by M.*

It is shown in Fig. 10 that RPH was the smallest in (b) and the largest in (c), and that mean RPH in (c) was above 1. One-way ANOVAs, with RPH with the dependent variable and with focus (broad, first, second) as the independent variable, revealed that there was a significant effect of focus for both speakers [for K, F (2, 12) = 75.29, p<0.001; for M, F (2, 12) = 82.09, p<0.001]. The results of *post hoc* Bonferroni tests are reported in Fig. 10. It was thus confirmed that in (c) pitch range was reset and hence an ***ι***-boundary was inserted at the beginning of the second word.

The results regarding pitch pattern alternation were quite similar to those in Datasets (1)-(2): all the prosodic words but the final were realized in connective form. Fig. 11 shows LI in (b). One-way ANOVAs, with LR with the dependent variable and with word position as the independent variable, revealed that there was a significant effect of word position. [for K, F (2, 12) = 35.95, $p < 0.001$; for M, F (2, 12) = 60.05, $p < 0.001$].

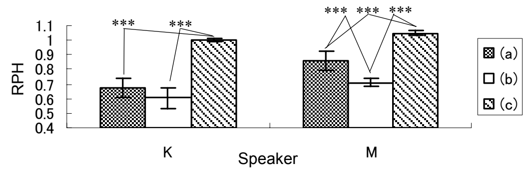

Figure 10: *Relative Peak Height (RPH) for Dataset(3).*

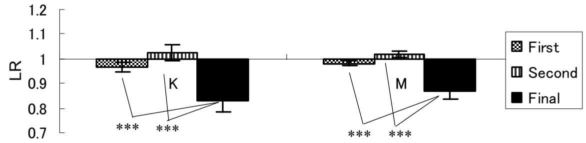

Figure 11: *Lowering Ratio (LR) for Dataset (3c).*

In summary, while a focused word introduced an ***ι***-boundary, the word immediately before the boundary was realized in *connective* form. Again, it was shown that the prosodic phrase for pitch pattern alternation is above ***ι***.

4. Discussion and Conclusion

The results of the present experiment revealed that pitch pattern alternation does not take place at the utterance-medial intermediate phrase (***ι***) boundary. They indicate that the prosodic phrase that serves as the domain for the alternation is one-level *above* ***ι*** (the domain for downstep).

Let us propose a tentative prosodic tree for Goshogawara, which is similar to the one developed for Tokyo [9]. The hierarchical model being proposed fundamentally relies on the Strict Layer Hypothesis [10], though alternative descriptions may be possible. This needs to be explored in detail in the future.

The domain for pitch pattern alternation is referred to as ‘intonational phrase’ (***IP***). An accented word is realized in non-connective form when the word is ***IP***-final. Thus, Goshogawara is proposed to have *three-level* prosodic phrasing above prosodic word (***ω***): that is, accentual phrase (***α***), intermediate phrase (***ι***) and intonational phrase (***IP***). The prosodic organizations for the utterances of Dataset (1) (with left- or right-branching syntactic structure) can be represented by the prosodic trees illustrated in Fig. 12.

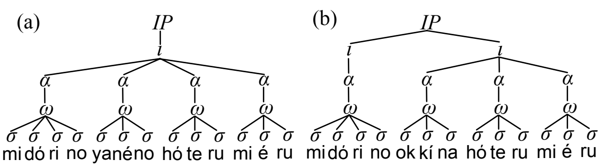

Figure 12: *Proposed prosodic trees of Goshogawara.*

In both utterances each ***ω*** constitutes a single ***α***, because it has pitch accent. Since the utterance (a) consists of a single ***ι***, downstep occurs interactively throughout the utterance. In (b), there is an ***ι***-boundary between the first and second ***ω***, and hence downstep is blocked at the beginning of second ***ω***. Note that in both utterances, pitch pattern alternation occurs only once, because they consist of a single ***IP***. The ***ω*** located at the final position of ***IP*** is only the utterance-final ***ω*** *miéru*, and thus it is the ***ω*** that is realized in non-connective form.

The present study, showing the existence of the prosodic phrase one-level above intermediate phrase in Goshogawara Japanese, will contribute to our understanding of nature of the organization of prosodic phrasing in general. In Pierrehumbert and Beckman’s model for Tokyo Japanese [9], the empirical foundation for postulating ‘utterance’ (the prosodic phrase above intermediate phrase) seems not solid enough. Indeed, in Japanese Tone and Break Indices (ToBI) (intonation labeling scheme for Tokyo) [14], the ‘utterance’ and intermediate phrase are merged into a single prosodic phrase called ‘intonation phrase’. Thus the prosodic organization above words becomes double-layered.

Goshogawara Japanese, on the contrary, has clear phonological evidence (i.e. the domain for pitch pattern alternation) for postulating the prosodic phrase above intermediate phrase. The triple-layered prosodic hierarchy model is viable at least for Goshogawara.

5. References

[1] Boersma, P. and Weenink, D., Praat: doing phonetics by computer, ver. 4.5.04 (Computer program), 2005.

[2] Gussenhoven, C., The Phonology of Tone and Intonation, Cambridge Univ. Press, 2004.

[3] Igarashi, Y., "Lexical pitch accent in Goshogawara Japanese: Rising or falling?", Proc. 16th ICPhS, 2007.

[4] Ishihara, S., "Japanese downstep revisited", Poster presentation at Word Accents and Tones in Sentence Perspective, Lund University, January 10, 2007.

[5] Kubozono, H., The Organization of Japanese Prosody, Ph.D. dissertation, Edinburgh University, 1988.

[6] Kubozono, H., "Focus and Intonation in Japanese: Does focus trigger pitch reset?", To appear in: Ishihara, S. (ed.) Interdisciplinary Studies on Information Structure, Universitätsverlag Potsdam, Potsdam.

[7] Ladd, D. R., Intonational Phonology, Cambridge University Press, Cambridge, 1996.

[8] Maekawa, K., "Muakusentohogen no intoneeshon (shiron)", Kinki Society for Phonetics (ed.) Onsei Gengo IV, 87-110, Osaka, 1990.

[9] Pierrehumbert, J and Beckman, M., Japanese tone structure, MIT Press, 1988.

[10] Selkirk, E.O., Phonology and Syntax: The Relation between Sound and Structure, MIT Press, 1984.

[11] Uwano, Z., "Nihongo no akusento", Shibata, T. (ed.) Iwanamikoza Nihongo 5, 281-321, Iwanami, Tokyo, 1977.

[12] Uwano, Z., "Nihongo no akusento", Sugito, M. (ed.) Koza Nihongo to Nihongokyoiku 2: Nihongo no Onsei On’in, vol. 1, 178-205, Meiji Shoin, Tokyo, 1989.

[13] Uwano, Z., "Classification of Japanese accent systems", S. Kaji ed., Proc. Symposium Cross-Linguistic Studies of Tonal Phenomena, 151-178, Tokyo, 1998.

[14] Venditti, J., "The J_ToBI model of Japanese intonation", Jun, S.-A. (ed.) Prosodic typology, 172-200, Oxford University Press, New York, 2005.

[15] Venditti, J., Maekawa, K. and Beckman, M., "Prominence marking in the Japanese intonation system", To appear in: Miyagawa, S. and Saito, M. (eds.) Handbook of Japanese Linguistics. Oxford University Press.