Japanese Downstep Revisited

Shinichiro Ishihara

October 2008

Abstract This paper presents the results of a production experiment in which downsteps in Tokyo Japanese were re-examined. There are three major assumptions that have been widely accepted in the literature: 1) Major Phrase (MaP) is the domain of downstep; 2) Syntactic boundary blocks a MaP boundary, blocking downstep; and 3) Focus inserts a MaP boundary, blocking downstep. The results of the experiment raise questions about these basic assumptions, and call for new theoretical explanations of the data. There are two major findings in the results: 1) No complete reset is seen, and 2) phonetic differences between the effects of focus and intonation of syntactic boundary. The first finding raises questions as to whether Assumption 1 and 2 should be maintained or not, and how they should be modified to capture the results. Recursive downstep is proposed to account for the incomplete reset. The second finding particularly casts doubt on Assumption 3, because the focus effect lacks some of the properties of the boundary effect. The difference between focus and boundary needs to be explained by assuming that focus effect is independent of MaP-phrasing.

Keywords Japanese -downstep- prosodic phrasing -focus- register resetting -syntactic-prosody mapping

1 Introduction

Japanese is one of the languages that have been extensively studied in the intonation phonology (McCawley 1965; Poer 1984; Redman and Pierrehumbert 1986; Pierrehumbert and Beckman 1988; Selkirk and Tatsuni 1988, 1991; Kuroda 1989, 1993; Nagahara 1994; Trechsel 1995; Sugahara 2003; Venditti 1997, 2003, among many others). In many of these studies, downstep (a.k.a. cataphora)-F0 downstep triggered by lexical H1L pitch accents (i.e., Only accented words trigger downstep). Japanese lexical items can be divided into two groups, depending on whether the item bears a lexically assigned pitch accent or not. The (lexical) pitch accent is always marked by H1L tone. This H1L tone induces a lowering of F0-contour of the following material. This means, only lexically accented words induce downstep. Unaccented words, on the other hand, do not trigger downstep.

Let us look at the examples in (2) (In the examples below, lexical pitch accents are indicated by '-.') The first two words in (2a) are unaccented, and hence show no sharp F0-fall. As a result, the F0 peak of the following word word 'wine' stays high. In (2b), on the other hand, the first two words are accented. The pitch accents on these words triggers sharp F0-fall. Due to this sharp F0-fall, the F0-peak of the following words are strongly lowered. Because of this effect, the F0-peak of the third word is realized much lower than in (2a).

(2) a. Na-no na-ri na-wa ni 'Naomi's big sister's wine'
   b. Na-o na-ri na-wa ni 'Naoya's big brother's wine'

Fig. 1 Sample F0-contours of the U-A sequence (2a) and the A-A sequence (2b)

With this definition of downstep as the starting point, the following three assumptions will be reviewed in the remainder of this section.

1. Major Phrase (MaP) as the domain of downstep
2. Syntactic boundary blocks downstep (Syntactic-prosody mapping principle)
3. Focus blocks downstep (MaP rephrasing by focus)
2.1 Major Phrase (MaP) as the Domain of Downstep

The first assumption to be examined in this paper is the prosodic domain of downstep effect. Among the levels of the prosodic hierarchy, Major Phrase (MaP, a.k.a. intermediate phrase) has been considered to be the prosodic domain of downstep (Fosser 1984; Pantelidou and Beckman 1996; Kubosono 2003, among others).

(3) Assumption 1:
Major Phrase (MaP) is the domain of downstep.

As we saw in (2) above, downstep lowers the F0-realization of the material following each pitch accent. It has been claimed or assumed by many researchers that this effect is observed only within a MaP. This means that at the end of the MaP, any downstep effect is canceled. Accordingly, at the beginning of the following MaP, the F0-contour is realized in a full pitch span (Register resetting).

(4) Selkirk and Tateishi (1991: 535 (16))
a. Downstep: Within a Major Phrase, introduce downstep (i.e., lower the pitch register) after the first accent.
b. Register Resetting: At the beginning of a Major Phrase, reset the pitch register.

Downstep (1) and register resetting (1) can be illustrated in 2:

Fig. 2 Schematic illustration of downstep and register resetting

Now, here comes a question: If MaP is assumed to be the domain of downstep, what determines the MaP phrasing? Some researchers, e.g., Kubosono (2007), take a position that downstep is the defining criterion for specifying the size of MaP. That is, as long as downstep effect is observed between two words, they are contained in a single MaP. When a (complete) register resetting is observed, there is a MaP-boundary at this position. This analysis is clear about what defines a MaP, while does not make any prediction as to how the MaP structure of a sentence would be.

Other researchers, e.g., Selkirk and Tateishi (1991), adopt some version of syntax-prosody mapping principle, according to which MaPs are created. Under this analysis, it is clear where MaP-boundaries should appear. Since downstep is expected to be observed within each MaP created by this principle, one can predict where downstep and subsequent register resetting would take place. If, however, the downstep effect (and the register resetting) is observed across a MaP-boundary predicted by the mapping principle, this will be a problem for this analysis. One of the goals of this paper is to examine the domain of downstep. In the next section, we review the syntax-prosody mapping principle.

2.2 Syntax-Boundary Blocks Downstep

The second assumption examined in this paper has to do with the syntax-prosody mapping principle proposed by Selkirk and Tateishi (1991). They proposed that left boundaries of syntactic maximal projections (XP) correspond to left boundaries of MaP, blocking downstep. With this principle, the MaP structural constituency can be predicted from the syntactic structure.footnote{1}

1 This line of analysis has been widely accepted in the Optimality Theoretic (Prince and Smolensky 1993) accounts as one of the ALIGN constraints, e.g., Truckenbrodt (1996, 1999).
Although their claims make the same point in that the second phrase in the right-branching structure in (7b) has a higher tone than the one in the left-branching (7a), their explanations are different. In Safir and Tateishi’s (1981) syntax-phonology mapping principle (5), it is due to the absence of downstep on the second phrase in the right-branching structure. In Hurokawa’s (1993) analysis, on the other hand, it is due to the Matricl Boost on that phrase.

The experiment conducted in the current study is designed to reexamine the prosodic effect caused by the syntactic boundary, asking whether syntactic boundaries block downstep and induce a register resetting, and if they do, whether it is a complete reset or a partial reset.

2.3 Focus Blocks Downstep

The last assumption concerns the prosodic effect of focus. It is well-known that semantic focus affects phonetic realization of the sentence. In Japanese, when a sentence contains a focus (in the sense of Rooth (1992)), two phenomena can be observed. The F0-peak of the semantically focalized phrase is raised; and the pitch contour of the post-focal material is lowered (Pierrehumbert and Beckman 1988; Nagahara 1994; Ishihara 2003; Sugahara 2003). We will call these two phonetic phenomena focal F0-rise and post-focal downstep, respectively.

An example of the F0-rise and the subsequent post-focal downstep is shown in Figure 3. The black solid line shows a sample pitch contour for the sentence in (9) without any narrow focus. The dashed line is a sample pitch contour with focus on the second phrase (the indirect object anyonome-no ‘sister-in-law-dat’), superimposed onto the contour without a focus. The F0-peak of the focused phrase is realized higher than the non-focused counterpart, while the post-focal peaks are realized lower after the focus.

(9) Ayoyama-ni anyonome-no ni erimaki-o Anda
Ayoyama-NOM sister-in-law-DAT scarf-ACC to Anda
‘Ayoyama knitted a scarf for his sister-in-law.’

a. No focus (solid line)
   b. With focus (red dashed line)

(a) Ayoyama-ni anyonome-no ni erimaki-o Anda
Ayoyama-NOM sister-in-law-DAT scarf-ACC to Anda
‘Ayoyama knitted a scarf for his sister-in-law.’

b. With focus (red dashed line)

Fig. 3 Focus F0-rise and post-focal downstep

These prosodic effects of focus have been often explained in the literature by modification of MaP phrasing (Nagahara 1994; Trueschkebrot 1992; among others). For example, based on Pierrehumbert and Beckman’s (1988) observation that a focused word shows no downstep effect, Nagahara (1994) (cf. Pierrehumbert and Beckman 1988) extended the idea of the syntax-prosody mapping principle discussed above to the focus effect. He proposed that the left edge of a focused phrase corresponds to a left MaP-boundary, blocking downstep.

10. MaP rephrasing by Focus

a. Focus-left-boost (Pierrehumbert and Beckman 1988)
   b. Focus-to-end

The following example illustrates how the analysis works. In (11), according to the syntax-prosody mapping principle discussed above, a MaP-boundary would be created between the subject Ayoyama-no anyonome-go ‘Ayoyama’s sister-in-law-NOM’ and the VP erimaki-o Anda ‘knitted a scarf’, as shown in (12a). Let us call this phrasing the default MaP phrasing—the syntax-prosody mapping principle. With this default phrasing, the non-MaP-initial phrases, i.e., the nominative DP anyonome-go and the verb Anda, would show downstep effect. The MaP-boundary between the subject and the VP blocks downstep on the object erimaki-o ‘scarf-ACC’, inducing a register resetting on the object.

When a semantic focus is placed on anyonome-go ‘sister-in-law-NOM’, the default MaP phrasing will be modified according to (10), as shown in (12b). First, a MaP-boundary is inserted on the left of focus, in this case, anyonome-go ‘sister-in-law-NOM’. The insertion of this MaP-boundary would block the downstep effect that is expected in the default phrasing (12a). This effect would be observed as a F0-rise on the focused phrase. Next, the MaP-boundary between the subject and the VP will be deleted, due to (10b). This means that the focused phrase anyonome-go and the VP will confirm a single MaP. As a result, the object erimaki-o ‘scarf-ACC’, which is expected to show register resetting in the default phrasing, will show downstep effect in the post-focal domain.

(11) Ayoyama-ni anyonome-go ni erimaki-o Anda]
Ayoyama-NOM sister-in-law-NOM scarf-ACC to Anda
‘Ayoyama’s sister-in-law knitted a scarf.’

(a) No Focus
   b. Focus on anyonome-go ‘sister-in-law-NOM’

(a) Ayoyama-ni anyonome-go ni erimaki-o Anda]
Ayoyama-NOM sister-in-law-NOM scarf-ACC to Anda
‘Ayoyama’s sister-in-law knitted a scarf.’

(b) Ayoyama-ni anyonome-go ni erimaki-o Anda]
Ayoyama-NOM sister-in-law-NOM scarf-ACC to Anda
‘Ayoyama’s sister-in-law knitted a scarf.’

The gist of this analysis is that the post-focal downstep is explained as downstep, and the focal F0-rise as the lack thereof. The obligatory insertion of MaP-boundary at the left of focus blocks downstep, while the deletion of MaP-boundaries in the post-focal domain induces downstep throughout the domain. From this observation, we have the third assumption:


There are, however, pieces of evidence that contradict this assumption. For example, Shinya (1995) showed that a focused phrase exhibits a downstep effect. Using examples like (14a), where the sequence of accented words in a left-branching structure induces successive downsteps (14a), he examined whether the accumulated downstep effect would be completely canceled when the last word (14b) is focused. The result showed that although the focused phrase showed a higher F0-peak than a non-focused counterpart, it is still lower than its counterpart with no downstep effect (i.e., a phrase in a sequence of unaccented words, (14b)). He concluded that focus raises the F0-peak, but does not cancel the downstep effect, confirming the claim made by Poser (1984). Shinya’s (1999) results suggest that
downstep and focus F₁-boosting can co-occur on the same phrase, which is never expected under the assumption that focus blocks downstep.

(14) a. **Downstep case (with a sequence of accented words)**
   Q: sore-ge Osomori-no ann-ko mēgane-no funi-ni nite iru no desu ka?
   "Is it similar to Osomori’s brother’s glasses’ frame?"
   A: [[[ Osomori-GEN brother-GEN glasses-GEN color COP
   ‘(It is) the COLOR of Osomori’s brother’s glasses.’

   b. **No downstep case (with a sequence of unaccented words)**
   Q: sore-ge Osomori-no ake-ko yūnami-no yagoo-ni nite iru no desu ka?
   "Is it similar to Osomori’s sister’s tea cup’s pattern?"
   A: [[[ Osomori-GEN sister-GEN tea-cup-GEN color COP
   ‘(It is) the COLOR of Osomori’s sister’s tea cup.’

   (Shinya 1999: ex. (4))

   Kubonono (2007) showed that wh-phrase, which is inherently focused, is also affected by downstep. If focus always blocks downstep, it would mean that focused phrases such as wh-phrase would never exhibit downstep.⁶

   Given that there are contradicting analyses, it is worth re-examining Assumption 3. It will be examined whether focus cancels downstep, as well as whether it behaves prosodically in the same way as syntactic boundaries.

2.4 Goals

So far we reviewed three widely adopted assumptions on Japanese downstep: (i) MaP as the domain of downstep, (ii) syntactic boundary blocks downstep, and (iii) focus blocks downstep. We also saw that there are some questions and counterarguments against them. The goal of this study is to clarify these questions and to verify the counterarguments. The experiment conducted in this study was specifically designed to examine the following two questions:

1. Does a syntactic boundary or a focus really block downstep?
2. Does focus behave exactly like a syntactic boundary?

The first question concerns the validity of Assumption 2 (=6) and Assumption 3 (=12). If we find a downstep effect across a syntactic boundary or on a focused phrase, we would need to reconsider the definition of the MaP, namely, Assumption 1 (=3). The second question is whether it is appropriate to analyze the effect of syntactic boundary and that of focus in the same way. With these two questions in mind, we will take a look at the experimental design in the next section.

3 Experiment

3.1 Stimuli

All the stimuli sentences are constructed with five prosodic words: Three noun phrases followed by a postpositional phrase (PP) and a verb. We will call the first three noun phrases N₁, N₂, and N₃, respectively.

(15) **Stimulus template**

N₁ N₂ N₃ PP VP

* Although Kubonono does not mention this point, his data also cast doubt on Assumption 2, because the wh-phrase in his experiment stimuli are located at a syntactic boundary, where downstep is expected to be canceled, but still show the downstep effect.

3.2 Subjects

The subjects for the experiment are 9 volunteers. All the subjects are native speakers of Japanese.

3.3 Procedure

The experiment is conducted in a quiet and comfortable environment. The subjects are asked to listen to the stimuli and press a button when they hear a focus word. The stimuli are presented at a constant rate and the subjects are instructed to respond as quickly and accurately as possible.

4 Results

The results of the experiment are analyzed using a two-way ANOVA. The factors are (i) focus (F₁ or F₂) and (ii) syntactic boundary (within or across). The analysis shows a significant interaction between the two factors, indicating that focus and syntactic boundary have different effects on downstep.

5 Discussion

The results of the experiment support our hypothesis that focus blocks downstep, while syntactic boundary does not. This difference in behavior suggests that focus and syntactic boundary are distinct phenomena.

6 Conclusion

In conclusion, we have shown that focus blocks downstep, while syntactic boundary does not. This has implications for our understanding of the nature of focus and downstep in Japanese.

(16) 5 factors

1. **Lexical Pitch Accent on N1/N2: [+Accent]**
   [-Accent] Nōma-ko een-no wina ‘Nōma’s mother’s wine’
   [-Accent] Nōma-ko een-no wina ‘Nōma’s mother’s wine’
   (U-U-A)
   (U-A-A)

2. **Focus on N3: [+Focus]**
   [-Focus] wina ‘wine’
   [-Focus] wina ‘wine’
   (Nom phrase)
   (Wh-phrase)

3. **Syntactic Boundary before N3: [+Boundary]**
   [-Boundary] [N1 N2 N3]PP VP VP
   [-Boundary] [N1 N2 N3]PP VP VP
   (Long object)
   (Subject + Object)

As shown in (2), lexically unaccented words do not trigger downstep, while accented words do. A stimulus with unaccented N₁ and N₂ would show no downstep effect on N₃, while its counterpart stimulus with accented N₁ and N₂ would show a downstep effect on N₃, unless any blocking factor (i.e., a syntactic boundary or a focus) triggers a register resetting and cancels the downstep. This means that stimuli with unaccented N₁ and N₂ provide a baseline for the performance of the test stimuli with accented N₁ and N₂ with regard to the existence of downstep effect. If the F₁-peaks of N₃ in the [+Accent] stimulus is lower than that of the [-Accent] counterpart, it would mean that there is no downstep effect on N₃. If they are at the same height, it would mean that there is no downstep effect.

Next, the stimuli is alternated according to the focal status of N₃. For the [-Focus] stimuli, accented common nouns such as wina ‘wine’ are used for N₃. In their [+Focus] counterparts, N₃ is replaced by a wh-phrase such as nani ‘what’, and a question particle no is added to the end of the sentence. The wh-phrase in a wh-question in Japanese prosodically behaves as a focused phrase: it shows a focal F₁ rise and triggers a post-focal downstep of the following material (Misakawa 1991; Deguchi and Kitagawa 2002; Ishihara 2002, 2003, among others). Given Assumption 3, we expect that downstep on N₃ in the [+Accent, -Focus] stimulus would be blocked in the [+Accent, +Focus] condition.

Lastly, the stimuli are constructed with either one of the following two syntactic structures.

(17) a. **Long object structure: [-Boundary]**

b. **Subject + Object structure: [+Boundary]**
In the [-Boundary] structure, N1, N2, and N4 form a single DP and serves as the object of the sentence. The subject for this structure is in a null pronoun (pro). There is no syntactic boundary in front of N3. In the [+Boundary] structure, on the other hand, N1 and N2 form a DP and serves as the subject of the sentence, while N3 stands alone as the object of the sentence. In this structure, there is a VP-boundary between N2 and N3. Given Assumption 2, we expect downstep on N3 in the [+Boundary] structure, while the register resetting bluods downstep in the [-Boundary] structure.

This set of factors (2 x 2 x 2 design) makes 8 different conditions for the experiment. One stimulus set is given below.

(18) Sample Stimuli
a. –Accent, –Focus, –Boundary
[VP [N1 N2 pro-o wain-o] wainyrus-2m-de nmanda]
N-gen sister-gen wine-ACC wineglass-with drink
"(I) drank Naomi's big sister's wine with a wineglass."
b. +Accent, –Focus, –Boundary
[VP [N1 N2 pro-o wain-o] wainyrus-2m-de nmanda]
N-gen bro-gen wine-ACC wineglass-with drink
"(I) drank Naoya's big brother's wine with a wineglass."
c. –Accent, +Focus, –Boundary
[VP [N1 N2 pro-o wain-o] wainyrus-2m-de nmanda]
N-gen sis-NOM wine-ACC wineglass-with drink
"Naomi's big sister's what? did you drink that with a wineglass?"
d. +Accent, +Focus, –Boundary
[VP [N1 N2 pro-o wain-o] wainyrus-2m-de nmanda]
N-gen bro-NOM wine-ACC wineglass-with drink
"Naoya's big brother's what? did you drink that with a wineglass?"
e. –Accent, +Focus, +Boundary
[VP [N1 N2 pro-a] [VP wain-o wainyrus-2m-de nmanda]]
N-gen sis-NOM wine-ACC wineglass-with drink
"Naomi's big sister drank wine with a wineglass."
f. +Accent, +Focus, +Boundary
[VP [N1 N2 pro-a] [VP wain-o wainyrus-2m-de nmanda]]
N-gen bro-NOM wine-ACC wineglass-with drink
"Naoya's big brother drank wine with a wineglass."
g. –Accent, –Focus, +Boundary
[VP [N1 N2 pro-a] [VP wain-o wainyrus-2m-de nmanda]]
N-gen sis-NOM what-ACC wineglass-with drink
"What, did Naomi's big sister drink that with a wineglass?"
h. +Accent, –Focus, +Boundary
[VP [N1 N2 pro-a] [VP wain-o wainyrus-2m-de nmanda]]
N-gen bro-NOM what-ACC wineglass-with drink
"What, did Naoya's big brother drink that with a wineglass?"

There are several important comparisons to be made in order to examine the effects of pitch accent, focus, and syntactic boundary. First, the comparison of the two [+Focus, –Boundary] conditions ((18a) vs. (18b)) shows the pure effect of pitch accent, i.e., downstep. As we saw in (2), lexically accent words triggers downstep while unaccented word do not.

Whether focus cancels downstep or not can be checked by comparing the two [+Focus, –Boundary] conditions ((18c) vs. (18d)). In the same way, the boundary effect can be checked by comparing the [+Focus, +Boundary] conditions ((18e) vs. (18f)). Also, the comparison of the [+Focus, –Boundary] conditions ((18g) vs. (18h)) tells us whether the two effects can be cumulatively combined or not.

In addition, if we compare the [+Focus, –Boundary] conditions with the [+Focus, +Boundary] conditions, we can examine whether the focus effects and the boundary effects behave the same way or not.

(10) **Crucial comparisons**

- Accent effect (i.e., Downstep):
  - (18a) [+Acc, +Focus, –Boundary] vs. (18b) [+Acc, –Focus, –Boundary]
  - Focus effect:
    - (18c) [+Acc, +Focus, –Boundary] vs. (18d) [+Acc, +Focus, –Boundary]
    - Boundary effect:
      - (18e) [+Acc, –Focus, +Boundary] vs. (18f) [+Acc, –Focus, +Boundary]
    - Focus+Boundary effect:
      - (18g) [+Acc, +Focus, +Boundary] vs. (18h) [+Acc, +Focus, +Boundary]
- Focus effect vs. Boundary effect
  - (18c) [+Acc, +Focus, –Boundary] vs. (18d) [+Acc, +Focus, –Boundary]
  - (18e) [+Acc, +Focus, –Boundary] vs. (18f) [+Acc, +Focus, –Boundary]

3.2 Method

12 subjects (5 females and 7 males) participated in the experiment, who are all brought up in Tokyo and the surrounding areas. Six different stimulus sets of the 8 conditions shown above (48 sentences total, cf. §3.1 and Appendix Appendix A) were used. Stimuli were mixed with 66 filler sentences, provided in a pseudo-randomized order (so that two sentences from the same set are not presented in a row). The stimuli were separated into two groups ((18a), (18b), (18f), (18h) in one group, (18c), (18d), (18e), (18g) in the other), and recorded in two separate sessions, with a week long interval between. Each sentence was presented to the subject on a computer screen, one at a time. Each subject made 3 recordings of the entire set of stimuli. Each recording session used a different pseudo-randomized order of the sentences. Subjects were asked first to read the sentences (either aloud or quietly) to understand the meaning of the sentence, and then to read aloud for the recording.

For the analysis, three measuring points were taken for each of the three nouns (N1, N2, N3): the F0-maximum and the F0-minima before and after it. The data from all the 12 subjects were first analyzed individually, and then normalized to factor out the pitch range differences among speakers, using the linear regression formula in (20). Actual values x in each subject's data are converted to normalized values y relative to the two reference points (R1, R2) given in (21):

(20) **Normalization formula**

\[ y = \frac{x - R_2}{R_1 - R_2} \]

(21) **Reference points (R1, R2)**

a. \( R_1 = \text{Mean value of } F_0 \text{-peak of } N1 \)

b. \( R_2 = \text{Mean value of } F_0 \text{-valley after } N3 \)

This formula basically means that for each subject the mean value of F0-peak on N1 and that of F0-valley after N3 will be converted as 1 and 0, respectively, and all the measured points will be mapped onto this scale.

4 Results

There are two noteworthy findings in the results. First, we did not observe a complete register resetting at the focused phrases, nor across a syntactic boundary. In both cases, N3 in the [+Acc] condition is realized lower than the [+Acc] counterpart. This finding raises theoretical question regarding the definition of
Map (Assumption 1). Second, a few remarkable differences were observed between the prosodic effects of syntactic boundary and those of focus. While focus only affects the realization of the F0-peak on the focal phrase, syntactic boundary creates a lower F0-valley at the boundary, and/or induces a lowering of the preceding F0-peak. This finding suggest that they should be analyzed as independent phenomena.

4.1 Finding 1: No Complete Reset by Focus/Boundary

4.1.1 Accent

Before looking at the results from the [+Focus] and the [+Syntactic Boundary] data, we first need to confirm the effect of lexical pitch accent, i.e., downstep. Figure 4 shows the result from the [-Focus, -Boundary] conditions (18a and 18b) . In the following presentations of the results, the normalized F0-means of the 9 measuring points—F0-maxima on N1-N3 and the F0-minima before and after them—will be shown. See Appendix A for individual results.

(22) /...-

\( N_{/N\text{-GEN}} \) sis/\textit{beo\text{-GEN}} \textit{wine-ACC} \textit{wineglass-with} \textit{drink}

‘(I) drank Naomi’s big sister’s/Naoya’s big brother’s wine with a wineglass.’

Fig. 4 Accent effect (downstep)

In the [-Accent] condition (black line), no downstep is observed. The slight downturn on N2 and N3 is presumably declination (Pierrehumbert and Beckman 1988; Ladis 1990). In the [+Accent] condition (gray line), on the other hand, the F0-peaks of N2 and N3 show downstep. As a result, the F0-peak of N3 in [+Accent] condition becomes significantly lower than that of the [-Accent] counterpart. This large difference in F0-height between [-Accent] and [+Accent] can be attributed to the existence/absence of lexical pitch accents on N1 and N2. All the 12 subjects showed downstep on N3. Based on this contrast, we examine whether this difference on N3 will be canceled by focus and boundary.

4.1.2 Focus

Figure 5 shows the effect of focus. The [-Focus, -Boundary] conditions (18c, 18d, repeated below as (23)) have a full phrase on N3. If focus blocks downstep, the downstep effect in the [+Focus] condition would disappear, i.e., the F0-height differences on N3 is no longer expected between the two conditions. (The [+Focus, -Boundary] conditions are shown in dashed line, and superimposed onto the [-Focus, -Boundary] baseline results in Figure 4.)

(23) /...-

\( N_{/N\text{-GEN}} \) sis/\textit{beo\text{-GEN}} \textit{what-ACC} \textit{wineglass-with} \textit{drink} \textit{Q}

‘Naomi’s sister’s/Naoya’s brother’s where did you drink it with a wineglass?’

Fig. 5 Focus effect

Although the F0-peak of N3 is raised strongly by focus compared to [-Focus] conditions (solid lines), the difference between [+Accent] (gray line) and [-Accent] (black lines) remains on N3. This suggests that the effect of downstep is not completely canceled.

Some individual variations were found (see Appendix A). Out of the 12 subjects did not show any significant difference on N3 (IT, ET, NT, and TK). Among them, IT and NT, both made subjects, are also among the group of speakers who did not show a significant contrast on N5 (see fn. 3). These group of speakers have relative small downstep effect on N3 as well. They have relatively narrow pitch range for entire utterance. It may be the case that because their pitch range is relatively narrow, the F0-peak of the focused phrase reaches its highest possible point even in the [-Accent] condition. This does not mean, though, that all the speakers with relatively narrow pitch range lose the downstep effect on N3. The difference is observed in some other speakers with a narrow pitch range (see, for example, SM’s results).

In the downstep effect on the following F0-contour but also a boosting effect on the pitch accent itself, this boosting effect raise the starting point of the F0-fall (note that N1 is higher in [+Accent] than in [-Accent]). The lack of significant on N2 does not concern us here anyway, because we are more interested in the realization of N3, whereas the other two factors ([Focus], [Boundary]) are also controlled.
KT and TK have very strong boost on focus. For these speakers, it seems to be the case that the focus effect is so strong that the F0-peak reaches its highest pitch range and shows a complete reset.4

In sum, as a general tendency, the downstep effect remains on a focused phrase. Only some speakers showed a complete reset.

4.1.3 Boundary

Figure 6 shows the effect of syntactic boundary. In the [-Foc, +Boundary] conditions ((18e), (18f), repeated below as (24)) there is a syntactic boundary between N2 and N3 (dotted line). If a boundary blocks downstep, no difference is expected between the F0-height of N3 in the two conditions.

(24) AbAccent, -Focus, +Boundary: ((18e)/(18f))
[pt Naomin/Naoma-NO sari/fu/ga ] [pt wapi-o waingirasa-de nandia ]
N,N-GEN sis/bro-NOM wine-ACC wineglass-with drunk

'Naomi's sister/Naoya's brother drank wine with a wineglass.'

4.1.4 Focus + Boundary

Even when the focus effect and the boundary effect are both expected at the same time (i.e. [+Foc, +Boundary] conditions, (18g) and (18h)) the reset is still incomplete, although the difference is much smaller. This is shown in Figure 7. In the individual data, 9 of the 12 subjects showed this contrast. Three subjects (AK, IT, SM) did not show any significant effect.5

(25) AbAccent, +Focus, +Boundary: ((18g)/(18h))
[pt Naomin/Naoma-NO sari/fu/ga ] [pt wapi-o waingirasa-de nandia ] no?
N,N-GEN sis/bro-NOM what-ACC wineglass-with drunk G

'What, did Naomi's sister/Naoya's brother drink wine with a wineglass?'

Fig. 6 Boundary effect

Just like the results for focus effect in the previous figure, the F0-peak on N3 of the [-Acc] condition is significantly lower than that of the [+Acc] condition. Although the F0-peak of N3 is raised strongly by boundary (dotted line) in both conditions, the difference between [+Acc] (gray line) and [-Acc] (black line) remains. This suggests that downstep is not completely reset by a syntactic boundary.

As for the individual results, two subjects (IT, SM) showed no significant difference on N3. Note that IT also did not show any significant difference in the focus effect. Also, two other subjects (KY, NT) showed a significant difference in the opposite direction: N3 in the [-Acc] condition is significantly lower than in the [+Acc] condition, although the differences are rather minimal. For these speakers, we can conclude that the complete reset is observed. As a general tendency, however, most subject did not show complete reset at the syntactic boundary.

4.1.5 Finding 1: Summary

From the results presented above, we can conclude that downstep effect is not completely reset by focus or syntactic boundary. This conclusion raises questions regarding the three assumptions we reviewed in §2. If neither a focus nor a syntactic boundary blocks downstep, we need to reconsider the definition of

---

4 There is only one subject who showed a very weak focus effect, KY. Her data is the only data in which the [+Foc, +Foc] condition is realized lower than the baseline condition, i.e., [-Foc, -Foc]. That is, the downstep effect is stronger than the F0-peak (reset) caused by focus.

5 This result confirms Kusumoto's (2007) experimental results (see fn. 2).
4.2 Finding 2: Syntactic Boundary ≠ Focus

Next, we compare the details of the phonetic effects triggered by focus and those by boundary. As we will see, there are clear differences in the phonetic realizations of focus and those of syntactic boundary.

In order to compare the focus effects and the boundary effects, we compare the [+Foc, −Boundary] conditions and the [−Foc, +Boundary] conditions. Figure 8 and Figure 9 show this comparison, in unaccented word sequence ([−Acc]) and accented word sequence ([+Acc]), respectively.

![Figure 8: Focus vs. Boundary Effects in [−Acc] Condition](image)

There are two major differences between focus and boundary effects: (i) the amount of F0-rise at N3, and (ii) pre-focus/boundary F0-realisation. We discuss these one by one.

### 4.2.1 Amount of F0-rise at N3

The first major difference between the focus effect and the boundary effect is the F0-realisation of N3. In both figures, the F0-peak of N3 is significantly higher in the [+Foc] condition (dashed line) than in the [−Foc] condition (dotted line). This suggests that everything being equal, a focused phrase is realised higher than a phrase at the left edge of a syntactic boundary.

6. There is one possible interpretation of the results with which one can still consider that there is a complete reset in these cases. We will also briefly discuss such a possibility in the discussion section.

### 4.2.2 Pre-boundary F0-lowering

The second major difference appears on the pitch contours preceding N3. While focus only affects the F0-peak of the focused phrase (N3 in the current experiment), boundary insertion affects not only the F0-peak of the phrase at the boundary (N1), but also the F0-valley at the boundary (between N2 and N3) or the preceding F0-peak (N2).

First, in the [−Acc] condition (Figure 8), the F0-valley between N2 and N3 are significantly lowered in the [−Boundary] condition (dotted line) compared to the baseline [−Boundary] condition (solid line). That is to say, the existence of a syntactic boundary triggers an F0-dip. Second, in the [−Acc] condition (Figure 9), the F0-peak of the preceding phrase (N2) is significantly lowered in the [+Boundary] condition. That is, the F0-peak of the pre-boundary (i.e., MaP-final) phrase is lowered. (Accounts for these F0-lowering effects will be discussed in 8.2.1.)

Such large F0-lowering phenomena are not observed in the [+Acc] conditions. In both figures, the [−Foc] condition (dashed line) and the baseline [−Foc] condition (solid line) basically show the same contour until the F0-peak of the focused phrase N3. These results show that focus only affects the pitch contour of the focused phrase (as well as the post-focal area, which is not discussed in this paper). The pre-focal area, however, is not affected by focus.

7. In fact, the two F0-lowering effects (the F0-dip in Figure 8 and the N2-peak lowering in Figure 9) can be found in the [−Foc] conditions (dashed line) as well. They are also statistically significant in the normalised results. These effects, however, are not as strongly observed in individual results as the cases of the [−Boundary] condition (not to mention that the effects are much smaller). The F0-dip in Figure 8 is significant in 5 of the 12 subjects’ data in the [−Foc] condition, while in the [−Boundary] condition it is significant in 8 subjects’ data. The N2-peak lowering is significant only in 4 subjects’ data in the [−Foc] condition, while it is significant in 11 subjects’ data in the [−Boundary] condition.

8. Although phonetic effects in the post-focal area were not examined in this current experiment, post-focal downward has already been discussed in the literature (Segaars 2003, Liebhaber 2000). Segaars (2002) showed that there are two different kinds of downward involved in the post-focal area. One can be attributed to deletion of prosodic phrase boundary (which she calls the "structural effect") and the other one can be observed when the prosodic phrasing remains intact ("non-structural effect").
4.2.3 Finding 3: Summary

The phonetic effects found in the [+Foc] conditions and those in the [-Boundary] conditions are summarized below:

(27) Focus effect
   a. F0-rise on the focused phrase (N2)
   b. F0-rise on the focused phrase (N3)

(28) Boundary effects
   a. F0-rise on the focused phrase (N2)
   b. [-Avc]: F0-dip at the boundary (between N2 and N3)
   c. [+Avc]: F0-lowering of pre-boundary phrase (N2)

The results presented here indicate that the realization of focus and that of syntactic boundary are phonetically distinct. In both cases, F0-rise on N3 is observed. However, the amount of the rise is different: The rise caused by focus (77a) is stronger than the one by boundary (78a). The other two phenomena—the F0-dip after N2 (25b) and the N2-peak lowering (26c)—are peculiar to the [+Boundary] conditions, and not found in the [+Foc] conditions. This raises interesting questions regarding the theory of prosodic phrasing. The analysis which treat the boundary effect and the focus effects in the same way, would face a problem explaining the phonetic differences revealed in this study.

In addition to the results shown above, this conclusion is further supported by the results for the [+Foc, +Boundary] conditions, where N3 bears a focus and is located at the boundary. As shown in Figures 10 and 11, properties of the focus and boundary effects can be observed simultaneously in such conditions.

![Figure 10 & 11: Focus-Boundary effects (gray solid line) compared to pure Focus (dashed) and Boundary (dotted) effects](image)

In both [-Avc] and [+Avc] conditions, the [+Foc, +Boundary] condition (gray solid line) shows a strong F0-rise at N3, just like the [+Focus] condition (dashed line), while it also shows the pre-boundary F0-lowering phenomena, just like [-Boundary] condition (dotted line): an F0-dip at the boundary (between N2 and N3) in the [-Avc] condition (Figure 10), and the lower F0-peak at N2 in the [+Avc] condition (Figure 11). These results suggest that phonetic effects of focus and those of syntactic boundary are different, and affect phonetic output independently. Given the evidence presented so far, we reach the following conclusion.

(29) Finding 3: Boundary effect and Focus effect

Focus and syntactic boundary show different phonetic effects.

In the next section, we discuss the theoretical assumptions introduced in §2 and related issues, based on the results of the experiment.

5 Discussion

5.1 Finding 1: No Complete Reset by Focus/Boundary

The first finding in §4.1 raises questions regarding the three basic assumptions discussed in 2. Assumption 1 in (3) defines MxP as the domain of downstep. If we try to maintain it, would that mean that we have to conclude that neither focus nor syntactic boundary form a MxP, because there's no complete resetting? In other words, would we have to modify the other two assumptions in order to maintain the first one? Alternatively, should we consider the definition of MxP? In that case, what would be the definition of MxP and the true domain of downstep? How do we account for the partial reset found in the results?

In order to account for the partial reset, it will be proposed that downstep takes place in a recursive fashion. The basic idea will be presented below, together with several possible theoretical implementations.

In the discussion of the second finding in §4.2 (differences between the focus and the boundary effect), it will be proposed, for the reasons to be made clear below, that only the syntactic boundary, but not the focus, creates MxP-boundaries. Therefore, we will postpone the discussion of Assumption 3 (focus effect) until the next subsection (§5.2), and concentrate here on Assumption 1 (the definition of downstep domain) and Assumption 2 (the syntactic boundary effect).

5.1.1 Recursive downstep

The proposal here is to assume that downstep applies recursively, following many other researchers (Ladd 1989, 1995; van den Berg et al. 1992; Truckenbrodt 2002; Fité and Truckenbrodt 2005). There are more than one way to theoretically implement this idea. Here we will entertain three possible options:

1. Two levels of Downstep (van den Berg et al. 1992; Truckenbrodt 2002; Fité and Truckenbrodt 2005)
2. Recursive MxP structure (Wagner 2005; Ito and Mester 2006, 2007)
3. Downstep as utterance-level effect (Kubozono 1993; Tokuda 2005)

The first option is to assume two different levels of downstep. Following the downstep model for Dutch by van den Berg et al. (1992), Truckenbrodt (2002) discusses two levels of downstep in Southern German, namely between pitch accents and between two phraseal regions. Fité and Truckenbrodt 2005 propose that downstep takes place both at the intermediate phrase (ip, which corresponds to MxP in this paper) and at the intonation phrase (IP).

(30) Recursive downstep and partial reset (Truckenbrodt 2005: 87)

![Recursive downstep and partial reset](image)

This line of analysis can be applied to Japanese downstep as well. The first level is the one stated in Assumption 1 in (3), namely, downstep within a MxP: If there are two accentuated prosodic words (Pwds) within a single MxP, the F0-peak of the second Pwd is realized lower. The other level of downstep is the one across two MxPs. That is, when there are two MxPs (and the first one contains a lexical pitch accent), the second MxP is realized lower than the first one.

Such analysis can account for our experimental results for the syntactic boundary effect without any modification of the assumptions. When there is a syntactic boundary in front of N3 (cf. [+Foc, +Boundary] conditions (16e) and (16f)), a MxP-boundary is inserted at this position according to Assumption 2. With this phrasing, two different downstep phenomena are expected: one within MxP (i.e., between N1 and N2), and one across the two MxPs.
(31) MaP-phrasing of [\textit{r-Boundary}] conditions according to Assumption 2

\[ [\text{M1} \text{~N2} \text{~N3}] \rightarrow ([\text{M1} \text{~N2}]_{\text{MaP}} \rightarrow ([\text{N1} \text{~N2}]_{\text{MaP}} \rightarrow \text{N3})_{\text{MaP}} \text{MaP2}

\text{Downstep:}

a. within \([\text{N1} \text{~N2}]_{\text{MaP}}\)

b. between \([\text{N1} \text{~N2}]_{\text{MaP}}\) and \([\text{N3}]_{\text{MaP}}\)

The within-MaP downstep in (31a) lowers the F0-peak of N2, followed by a register resetting after N2, which is realized on N3. The across-MaP downstep in (31b) lowers the F0-realization of MaP2, which contains N3. This means that N3 is subject to two effects: a register resetting from the within-MaP downstep, and the lowering from the across-MaP downstep. As a result, N3 would show a register resetting, but not completely. Because of the across-MaP downstep, we only observe a partial resetting at the syntactic boundary.

Note that this analysis does not require modifications of the basic assumptions. Only an additional assumption is added that the across-MaP downstep exists. This analysis can be called a rather conservative option, as the basic assumptions on prosodic phrasing and downstep phenomena (e.g., Assumption 1 and 2 above, as well as in the Strict Layer Hypothesis (Naspor and Vogel 1986; Selkirk 1986)) are all maintained. Also, the additional assumption to be added is very restricted: it allows no more than two levels of downstep.

The second possibility is to allow MaP's to have recursive structure by abandoning the Non-recursivity of Strict Layer Hypothesis. The basic idea is to strengthen the syntax-prosody mapping principle, and to allow more direct mapping between syntactic boundaries and prosodic MaP's. If MaP's occur in a recursive fashion, i.e., if a MaP contains another MaP, then downstep is also expected to take place in a recursive fashion, because the embedded MaP would be subject to the downstep effect of the larger MaP.

Such recursive prosodic structures have been already proposed elsewhere. Lade (1988) discusses recursive domain of declination effect (which is Intonation Phrase, IP, one level higher than MaP discussed here). Wagner (2005) also propose a model in which prosodic domains are more directly mapped from syntax to a recursive fashion. Such recursive models of register effects are compatible with the experimental result presented in this paper. For Japanese, Oh and Mester (2006, 2007) propose a prosodic model in which the same level of prosodic phrasing can be adjusted recursively. In their model, prosodic phrases (\(\Phi\))—presumably the domain of downstep—can be embedded in a larger \(\Phi\), which would result in recursive application of downstep.

If we assume that every XP corresponds to a MaP, then the syntactic structure for the [r-Boundary] condition in the experiment would yield a MaP-phrasing below (cf. (31) above). Assuming that Assumption 1 in (3) is still intact, there are three different downstep domains:

(32) MaP-phrasing of [\textit{r-Boundary}] conditions according to the recursive MaP analysis

\[ [\text{M1} \text{~N2} \text{~N3}] \rightarrow ([\text{M1} \text{~N2}]_{\text{MaP}} \rightarrow ([\text{N1} \text{~N2}]_{\text{MaP}} \rightarrow \text{N3})_{\text{MaP}} \text{MaP2}

\text{Downstep:}

a. within MaP1: \([\text{N1} \text{~N2}]_{\text{MaP}}\)
b. within MaP2: \([\text{N1} \text{~N2}]_{\text{MaP}}\)
c. within MaP3: \([\text{N1} \text{~N2}]_{\text{MaP}} \rightarrow \text{N3})_{\text{MaP}}\)

(31b) can be ignored here because there is no F0-peak in MaP2 on which downstep is realized, given that MaP2 contains only F0. The other two downstep effects would be exactly the same as the one predicted by the previous proposal discussed in (31): a downstep from N1 to N2, and then another downstep from MaP1 to MaP3 (i.e., N3).

The crucial difference between the two-level downstep analysis in (31) and the recursive MaP analysis discussed here is that in the recursive MaP analysis, the levels of downstep is in principle unlimited. There can be an unlimited number of embedding of MaP's and hence an unlimited levels of downstep. The other analysis, on the other hand, would only allow two levels.

This line of analysis maintains the two assumptions (Assumption 1 and 2), while abandoning the non-recursivity of prosodic phrasing in the Strict Layer Hypothesis. One important consequence is that these models would make detailed predictions as to how the prosodic phrasings of a syntactic object would look like, because the prosodic structure would be computed more directly from the syntactic structure.

Once we allow recursive MaP structure, one might wonder if the distinction between MaP and other levels of phrasing (e.g., Minor Phrase, MP; Intonational Phrase, IP) becomes less clear. This is, however, not necessarily an unwelcome result. Oh and Mester (2006, 2007) in fact propose that MP and MaP can be defined uniformly by using a single level of prosodic hierarchy (\(\Phi\)). The difference between MaP and MP in this view can be explained by using a relational notion ("minimal" and "maximal" projection). In such an analysis, the most crucial aspect of the prosodic structure is that it allows unlimited number of recursion. The results of the experiment may be considered as one of the many cases where recursion seems to be involved.

The third option is to abandon the notion of MaP entirely, and assume that the domain of downstep is always the entire utterance. If we adopt a stronger correspondence between syntactic and prosodic boundaries as suggested in the previous analysis, one could even dispense with Assumption 1, and propose that the domain of downstep is the entire utterance. That is, all the lexical pitch accent triggers downstep of the following material, regardless of its location. Under this analysis, the register resetting at MaP-boundaries is independent of downstep, and taken to be an F0-boosting effect at each XP-boundary. Recall that, we found no case of complete reset. In other words, there are only different amounts of F0-raises in different conditions.

This difference in the amount of downstep can be calculated according to the syntactic structure. For example, Kohnoo (1993) proposed that at each syntactic left-boundary, there is an F0-boosting effect which he calls the Metrical Boost (MB). When more than one syntactic left-boundary coincides at the same position, then this position is subject to a stronger Metrical Boost effect. If a phrase follows a lexical pitch accent and there is no left-boundary in front of it, it would show no boosting, and hence exhibit the full amount of downstep effect.

Technically, this analysis is the same as the previous one, although they analyze the same phenomena from different perspectives. Downstep takes place beyond the prosodic domain that has been considered to be the domain of downstep (MaP, i.e., between two syntactic left-boundaries). The different amount of downstep-blocking effect is explained by the syntactic structure. This analysis has a conceptual advantage in the sense that it makes the syntactic hierarchy simpler. What remains to be seen is whether the amount of register resetting exactly corresponds to what the syntactic structure predicts.

So far we have considered three analyses. They all share a common property: downstep (or register resetting) takes place in a recursive fashion, according to the recursive prosodic structure. Recursive prosodic structure explains the partial reset found in the experimental results. It remains to be further investigated which implementation of the idea makes the correct predictions of other downstep phenomena.

5.1.2 Partial reset as undershooting

Although the proposal above can explain the partial reset, it is not the only possible interpretation of the results. The incomplete resetting could be due to undershooting as a result of phonetic optimization in the

\[10\] One might point out that Non-recursivity can be considered as a viable constraint in the Optimality Theoretic framework (Prince and Smolensky 1993), and hence the partial reset is the result of the constraint ranking in which Non-recursivity is ranked lower (cf. Trudgill 1999; Selkirk 2000). The proposal suggested here would be to assume that there is no such constraint at all, because the experimental results presented here suggest that a sign of recursivity can be found even in the case where no recursivity was assumed. Theoretical and empirical consequences of such a claim need further investigation.

11 Kohnoo (1993) himself does not abandon the notion of MaP. He proposes MB as an amount for the difference in the amount of downstep within a MaP. The analysis presented here is therefore an extension of his idea.

Takasaki (2005) also argued that the strength of the prosodic phrase boundary corresponds to the number of syntactic boundary.
The boundary effect can be explained by assuming an insertion of a MaP-boundary, i.e., by maintaining Assumption 2. Adopting some version of syntax-prosody mapping principle (e.g., (5)), a MaP-boundary is inserted at the left of each syntactic boundary. This MaP-boundary will block downstep, inducing an F0-rise after the boundary. As we discussed in §4.1.1, the incomplete reset can be explained by incomplete downstep.

The F0-dip found in the [+Ac]–[-Foc]–[Boundary] condition (see Figure 8) is due to the MP-final L-tone (Pierrrehumbert and Beckman 1988). If there is a MaP-boundary, it should coincide with every lower level of prosodic hierarchy. This means that there is a MaP-boundary between N2 and N3. The MaP-boundary is the source of the boundary lowering. In the [+Ac]–[-Foc]–[Boundary] condition, N1, N2, and N3 can (though not necessarily have to) be included in a single MIP. If this is the case, there is no MaP-boundary between N2 and N3. Hence no initial lowering is expected. In the [-Ac]–[-Foc]–[Boundary] condition, on the other hand, a MaP-boundary and a MIP-boundary are obligatorily inserted between N2 and N3. As a result, there is an initial lowering at this position. This initial lowering is realized as the F0-dip.

Also, the lowering of the N2-peak in the [+Ac]–[-Foc]–[Boundary] condition in Fig. 11 can be explained if we assume that there is a final lowering at the end of MaP (Pierrrehumbert and Beckman 1988, 1989, 1993, 2004). Traum and Traum (2004) showed that in Southern dialects of German the last element of the sequence of downstepped tones shows a final lowering effect. He claimed that downstep has not only the lowering effect of the downstepped tone, but also an effect of raising the phonetic realization of the preceding tone. The final lowering is explained as the lack of this boosting effect. The last element of the downstep sequence is not followed by another downstepped tone, and hence does not receive this boosting effect.

The lowered N2-peak found in the results here could be due to this final lowering effect. In Japanese, downsteap has been the only prosodic phenomenon that has been claimed to be associated with MaP. If this final lowering can be accounted consistently at the MaP-final position, this will be another phenomenon associated with MaP.16

5.2 Finding 2: Syntactic Boundary ≠ Focus

The second finding in §4.2—the difference between the prosodic effects of focus and those of syntactic boundary—has an important theoretical implications. Prosodic effects of focus should not be treated parallel to the those of syntactic boundaries. Also, the fact that the focus and the boundary effects appear in an accumulative fashion (cf. Figure 10 & 11) suggests that their effects are independently motivated. This point is particularly important because there have been many analyses in the literature that treat the focus effect and the boundary effect equally. They often assume that both syntactic boundary and focus insert a prosodic phrase boundary in the same way. Such analyses would not be able to account for the phonetic differences found in this study.

5.2.1 MaP boundary insertion at the syntactic boundary

Recall that focus has an effect only on the F0-peak of the focused phrase, whereas a syntactic boundary induces more than one effect in the F0-realization (see (27) & (28)). While the F0-peak after the syntactic boundary (N3 in the experiment) is raised, the F0-peak before the boundary (N2) is lowered. In addition, the F0-valley at the syntactic boundary is also lowered. In order to account for the difference between the focus effect and the boundary effect, the proposal is to assume different mechanisms responsible for the boundary effect and for the focus effect, respectively.

---

15 This possibility is suggested by Robert Traum (p.c.).
16 Pierrrehumbert and Beckman (1988) discuss final lowering, but they claim that the domain of this final lowering is larger than the MaP (i.e., their intermediate phrase). This phenomenon is presumably of a different kind from the one being discussed here.

17 Alternatively, we can also hypothesize that the amount of downstep varies depending on the length of the MaP. Note that in the [-Ac]–[-Foc]–[Boundary] condition, the MaP contains N1, N2, and N3, while in the [+Ac]–[-Foc]–[Boundary] condition, it only contains N2 and N3. If there is a critical target of downstep at the end of each MaP, it would have to be resolved earlier in the [+Ac]–[-Foc]–[Boundary] condition than in the [Ac]–[-Foc]–[Boundary] condition. This would result in the larger amount of downstep.


---
6 Conclusion

In this paper, the results of an experiment were presented in which downstep in Tokyo Japanese was reexamined. There are two important findings in the results: 1) incomplete register resetting by focus/boundary, and 2) phonetic difference between focus effect and boundary effect. These findings cast questions on the basic assumptions that have been often adopted in the study of downstep. Recursive downstep is proposed to account for the incomplete resetting. The difference between focus and boundary is explained by assuming that focus effect is independent of MaP-phrasing.

Appendix A Individual Results

For each subject, the mean F2-values of (i) +Focus -Boundary (Left), (ii) -Focus +Boundary (Center), and (iii) +Focus +Boundary (Right) are shown, with 95% CI. Solid lines in the figures show the F2-means of the baseline conditions (+Focus, +Boundary).

A.1 AK

A.2 AY

A.3 HT

A.4 IT

A.5 KT
Appendix B Stimuli Sets

(A0) Conditions
a. [-Accent, +Focus, -Boundary]
  - [a] [vp [np Naomi-GEN baer-GEN *wilím-o | wamíkana-da nódòa | Naomi-GEN big-sister-GEN wine-GEN drink-ACC] 'Naomi's big sister's wine with a wineglass.'
  - [b] [vp [np Naomi-GEN baer-GEN *wilím-o | wamíkana-da nódòa | Naomi-GEN big-brother-GEN wine-GEN drink-ACC] 'Naomi's big brother's wine with a wineglass.'

  b. [vp [np Naomi-GEN baer-GEN *wilím-o | wamíkana-da nódòa | Naomi-GEN big-brother-GEN wine-GEN drink-ACC] 'Naomi's big brother's wine with a wineglass.'
  d. [vp [np Naomi-GEN baer-GEN *wilím-o | wamíkana-da nódòa | Naomi-GEN big-brother-GEN what-GEN drink-ACC] 'Naomi's big brother's wine, what did you drink it with a wineglass?'

  b. [vp [np Naomi-GEN baer-GEN *wilím-o | wamíkana-da nódòa | Naomi-GEN big-brother-GEN how-GEN drink-ACC] 'Naomi's big brother's wine with a wineglass.'

A3. a. [vp [np Yuiti-GEN owa-GEN *audiyil-e | xiiiiwana-da urit̡al̡ita | Yuiti-GEN father-GEN Audi-ACC drink-ACC] 'Yuiti's father's Audi with a drink.'
  b. [vp [np Yuiti-GEN owa-GEN *audiyil-e | xiiiiwana-da urit̡al̡ita | Yuiti-GEN big-brother-GEN Audi-ACC drink-ACC] 'Yuiti's big brother's Audi with a drink.'

A4. a. [vp [np Yuiti-GEN owa-GEN *audiyil-e | xiiiiwana-da urit̡al̡ita | Yuiti-GEN father-GEN Audi-ACC drink-ACC] 'Yuiti's father's Audi with a drink.'
  b. [vp [np Yuiti-GEN owa-GEN *audiyil-e | xiiiiwana-da urit̡al̡ita | Yuiti-GEN big-brother-GEN Audi-ACC drink-ACC] 'Yuiti's big brother's Audi with a drink.'

(1) lost Naomi's sister-in-law's scarf at the travel abroad.

A. [vp [np Oyamo-GEN owa-GEN *fírmak·ilo | kagguyliyikoo-da nákunita | Oyamo-GEN little-sister-GEN scarf-ACC what-ACC travel-ACC lost] 'Oyamo's little sister's scarf at the travel abroad.'

B. [vp [np Adyama-GEN owa-GEN *fírmak·ilo | kagguyliyikoo-da nákunita | Adyama-GEN little-sister-GEN scarf-ACC what-ACC travel-ACC lost] 'Adyama's little sister's scarf at the travel abroad.'

C. [vp [np Oyamo-GEN owa-GEN *fírmak·ilo | kagguyliyikoo-da nákunita | Oyamo-GEN little-sister-GEN what-ACC travel-ACC lost] 'Oyamo's little sister lost her scarf at the travel abroad.'

D. [vp [np Adyama-GEN owa-GEN *fírmak·ilo | kagguyliyikoo-da nákunita | Adyama-GEN little-sister-GEN what-ACC travel-ACC lost] 'Adyama's little sister lost her scarf at the travel abroad.'

E. [vp [np Oyamo-GEN owa-GEN *fírmak·ilo | kagguyliyikoo-da nákunita | Oyamo-GEN little-sister-GEN where-ACC travel-ACC lost] 'Oyamo's little sister lost her scarf at the travel abroad.'

F. [vp [np Adyama-GEN owa-GEN *fírmak·ilo | kagguyliyikoo-da nákunita | Adyama-GEN little-sister-GEN where-ACC travel-ACC lost] 'Adyama's little sister lost her scarf at the travel abroad.'
(A5) a. [Vp [Vp Yamamoto-no rinin-no domogyo-o] yosokurasa-tei tatamencotai] Y-gen neigh-gen is.house-ACC new.look-to rebuilt
'(Yamamoto's neighbor) rebuilt Yamamoto's house to a new look.'

b. [Vp [Vp Yamaguchi-no doya-no domo-o] yosokurasa-tei tatamencotai] Y-gen land-gen is.house-ACC new.look-to rebuilt
'(Yamaguchi's neighbor) rebuilt Yamaguchi's house to a new look.'

'Yamamoto's neighbor's what? did you rebuild it to a new look?'

'Yamaguchi's house's what? did you rebuild it to a new look?'

e. [Vp [Vp Yamamoto-no rinin-ga] [Vp domogyo-o yosokurasa-tei tatamencotai] Y-gen neigh-gen is.house-ACC new.look-to rebuilt
Yamamoto's neighbor rebuilt the main house to a new look.'

f. [Vp [Vp Yamaguchi-no doya-ga] [Vp domoy-0 yosokurasa-tei tatamencotai] Y-gen land-gen is.house-ACC new.look-to rebuilt
'Yamaguchi's house rebuilt the main house to a new look.'

g. [Vp [Vp Yamamoto-no rinin-ga] [Vp domo-no yosokurasa-tei tatamencotai] Y-gen land-gen is.house-ACC new.look-to rebuilt Q
'Yamamoto's what? did you rebuild it to a new look?'

h. [Vp [Vp Yamaguchi-no doya-ga] [Vp domo-no yosokurasa-tei tatamencotai] Y-gen land-gen is.house-ACC new.look-to rebuilt Q
'Yamaguchi's house's what? did you rebuild it to a new look?'

References:


