

《 Article 》

Prosody-Scope Match and Mismatch in Tokyo Japanese Wh-questions

Shinichiro Ishihara

Name: Shinichiro Ishihara
Affiliation: Universität Potsdam
Address: SFB 632 „Informationsstruktur“
Universität Potsdam
Postfach 601553
14415 Potsdam
Germany (Work)
Phone: +49 331-977-2979
E-mail: ishihara@rz.uni-potsdam.de

*I would like to thank the following people for their help, discussion and comments: Noam Chomsky, Gisbert Fanselow, Ingo Feldhausen, Caroline Féry, Danny Fox, Masako Hirotsu, Yoshihisa Kitagawa, Elizabeth Medvedovsky, David Pesetsky, and two anonymous reviewers. Some parts of this paper were presented at the Workshop on Prosody, Syntax, and Information Structure (WPSI) held at Indiana University, Bloomington on April 30–May 1, 2004. I would like to thank the participants of WPSI, as well as those of the Symposium on Prosody and Syntax at the 22nd Conference of the ELSJ, for their comments and discussion. All the remaining errors are my own.

Abstract

It has been claimed that Tokyo Japanese wh-questions exhibit a prosody-scope correspondence between focus intonation, which is always observed in wh-questions, and scope of those wh-questions (Deguchi and Kitagawa (2002), Ishihara (2002)). Based on experimental data, it is shown that when the wh-phrase is moved out of the scope of the wh-question by scrambling, focus intonation and wh-scope no longer show a correspondence. It will be argued that the Multiple Spell-Out model proposed in Ishihara (2003), which accounts for the Focus Intonation–Wh-scope Correspondence phenomena, also accounts for this apparently exceptional mismatch case.*

PROSODY-SCOPE MATCH AND MISMATCH IN
TOKYO JAPANESE WH-QUESTIONS

Keywords: focus intonation, wh-question, Tokyo Japanese, phase, Multiple Spell-Out

1 Introduction

In this paper, I will extend the discussion of Focus Intonation (FI)–Wh-scope Correspondence observed in Tokyo Japanese wh-questions (cf. Deguchi and Kitagawa (2002), Ishihara (2002, 2003), Kitagawa and Fodor (2003), Hirotsu (2004)).¹ Using experimental data, I will show that when the wh-phrase is scrambled out of the scope of the wh-question, the above mentioned correspondence collapses, a phenomenon which I will call FI–Wh-scope Mismatch. I will then introduce the analysis proposed in Ishihara (2003, 2004), and claim that this analysis accounts for both the Correspondence case and the Mismatch case. The proposed analysis is based on the notion of phase and Multiple Spell-Out in the recent Minimalist framework (Chomsky (2000, 2001a,b)). The main claim in this analysis is that prosody is computed cyclically, in a ‘phase-by-phase’ fashion.

The outline of the paper is as follows. First of all, as background, I will briefly introduce the FI–Wh-scope Correspondence phenomenon in Tokyo Japanese wh-questions, as well as the FI–Wh-scope Mismatch phenomenon, and present the main question of the paper (§2). Secondly, I will present details of the experiment that tests the intonation of the Mismatch sentence (§3). I will then introduce the Multiple Spell-Out model proposed in Ishihara (2003, 2004), and illustrate how this model accounts for both the FI–Wh-scope Correspondence and the Mismatch (§4). Lastly I will mention a few remaining questions (§5), and conclude the paper (§6).

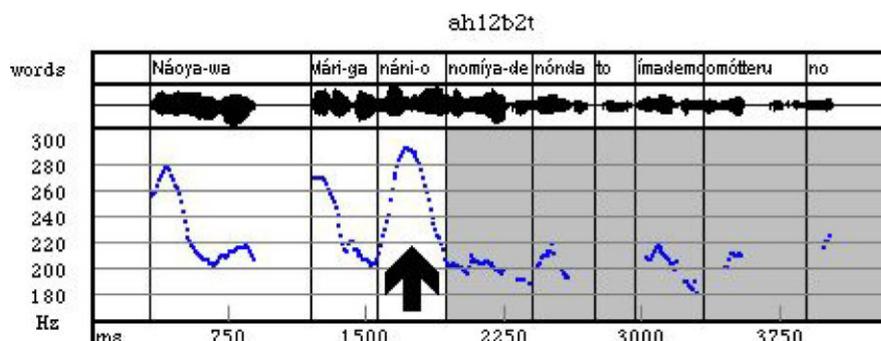
2 Background

2.1 Focus Intonation–Wh-Scope Correspondence

It is claimed that Tokyo Japanese, a wh-question sentence obligatorily exhibits Focus Intonation (FI): The F_0 -peak of the wh-phrase is boosted (focus F_0 -boosting), while the F_0 -peaks of the post-wh-phrases are significantly reduced (post-focus F_0 -reduction) (cf. Maekawa (1991, 1997)). Furthermore, it is claimed that the phonological domain of the FI (FI domain) and the semantic scope of the wh-question show a correspondence (Deguchi and Kitagawa (2002), Ishihara (2002, 2003)).² A post-focus F_0 -reduction in a wh-question continues until the end of the scope of the wh-question, where the question particle (Q-particle) that binds the wh-phrase appears. For example, in a matrix wh-question like (1a), the post-focus reduction continues until the end of the matrix clause, where the matrix Q-particle no appears, while in an indirect wh-question like (1b), the post-focus reduction stops at the end of the embedded clause, where the Q-particle ka appears, and a pitch reset is observed thereafter. This essentially means that the domain of the post-focus reduction indicates the scope of the wh-question.³

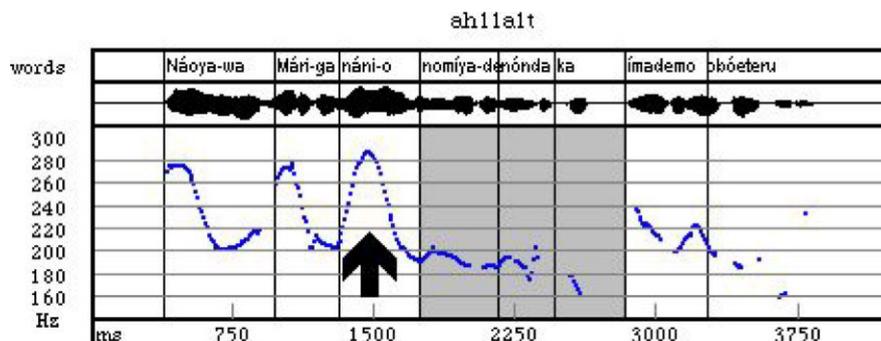
(1) a. Matrix wh-question

Náoya-wa [Mári-ga **náni-o** nomíya-de nónda to] ímademo
 Naoya-Top Mari-Nom what-Acc bar-Loc drank that even.now
 omótteru **no**?
 think Q
 ‘What_i did Naoya still think that Mari drank t_i at the bar?’



b. Indirect wh-question

Náoya-wa [Mári-ga **náni-o** nomíya-de nónda **ka**] ímademo
 Naoya-Top Mari-Nom what-Acc bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’



2.2 FI-Wh-scope Mismatch

There is a case, however, where the FI domain and the wh-scope do not correspond to each other properly. That is when the wh-phrase of an indirect wh-question is scrambled out of its scope, as in (2b).

(2) a. Indirect wh-question (= (1b))

Náoya-wa [Mári-ga **náni-o** nomíya-de nónda **ka**] ímademo
 Naoya-Top Mari-Nom what-Acc bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’

b. Indirect wh-question with wh-scrambling

náni_i-o Náoya-wa [Mári-ga t_i nomíya-de nónda **ka**] ímademo
 what-Acc Naoya-Top Mari-Nom bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’

In (2b), the wh-phrase is scrambled out of the embedded clause. The scope of the wh-question, however, is still the embedded clause, due to the radical reconstruction effect of long-distance scrambling (Saito (1989)). If the FI starts at the wh-phrase, just like in the non-scrambled case (1b), the FI–Wh-scope Correspondence would no longer be observed, because the FI would necessarily include the matrix material preceding the embedded clause (i.e. the matrix subject in (2b)). Therefore this case may be called FI–Wh-scope Mismatch.

Another point of interest in the FI–Wh-scope Mismatch case is the end point of the FI. Supposing that the FI starts from the scrambled wh-phrase, where does it end? Does it continue until the end of the matrix clause, just like a matrix wh-question (1a), or stop at the end of the embedded clause and show a pitch reset after that, just like a (non-scrambled) indirect wh-question (1b)?

I (Ishihara (2002)) and Kitagawa and Fodor (2003) have made similar observations about the intonation of the wh-scrambling sentence like (2b). Putting all the details aside, both Ishihara (2002) and Kitagawa and Fodor (2003) claimed that wh-scrambling sentences have the following FI:

- (3) In wh-scrambling examples, F_0 -reduction appears after the scrambled wh-phrase until the embedded Q-particle ka. A pitch reset occurs after the embedded clause.

$$[\text{CP } \boxed{\text{WH}} \text{ [TP } \dots \text{ [CP [TP } \dots \text{ } \underline{t_{\text{WH}}} \text{ } \dots \text{] } \mathbf{ka} \text{] } \alpha \dots \text{]}$$

↑

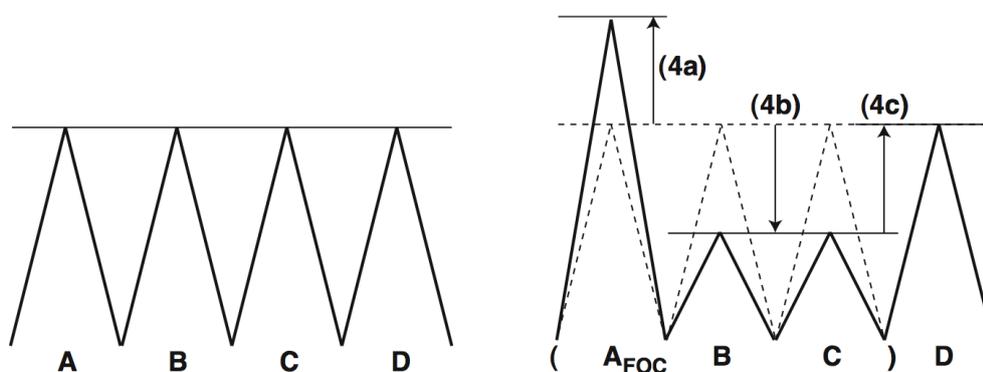
Pitch reset

Neither Ishihara (2002) nor Kitagawa and Fodor (2003), however, ran an experiment to confirm this observation. In this paper, therefore, I address particularly this question, i.e., how people actually pronounce wh-scrambling sentences like (2b). In §3, I will present details of the experiment specifically designed to examine this question.

2.3 Definitions

Before going into the details of the experiment, let us make clear the definitions of the phonetic phenomena to be discussed. FI can be detected by the three phonetic phenomena listed in (4). They are schematically illustrated in (5):

- (4) a. F_0 -boosting on the focused phrase (e.g. wh-phrase)
 b. post-focus F_0 -reduction
 c. pitch reset after FI domain.
- (5) a. Default contour (No FI) b. FI (Focus on A; FI domain A–C)



I will assume that focus F_0 -boosting (4a) is a phonetic effect that raises the F_0 -peak of the phrase bearing (semantic) narrow focus, and post-focus F_0 -reduction (4b) is a phonetic effect which compresses the pitch range of the post-focus material. In other words, an FI is created as a result of direct manipulation of pitch height or pitch range. In the schematic illustration in (5b), the F_0 -peak of the focused phrase (A) is raised, while the pitch range of post-focus elements (B and C) is compressed, resulting in lower F_0 -peaks for these phrases. I will call the phonological domain in which (4a) and (4b) apply FI domain.

The assumptions taken here depart from the standard analyses of FI in Tokyo Japanese (e.g. Pierrehumbert and Beckman (1988), Nagahara (1994), Truckenbrodt (1995), Selkirk (2003), Sugahara (2003)), in which FI is analyzed as a manipulation of Major Phrase (MaP) boundaries. Under these analyses, focus boosting is explained as an insertion of

MaP boundary on the left of focused phrase,⁴ and post-focus reduction as downstep as a result of MaP boundary deletion at the post-focus area. In other words, in the standard analyses, MaP always behaves as an FI domain.

In the assumption adopted in this paper, on the other hand, FI is a phonetic phenomenon independent of any prosodic phrasing or downstep. This means that a MaP phrase may appear within an FI, and that downstep may take place independently of the phonetic effects of FI listed in (4). In other words, the domain of downstep (MaP) and the domain of the FI (FI domain) are not identical. See Ishihara (2003) for arguments for this assumption about FI. We will see later (§4) how the FI domain is determined.⁵

Pitch reset (4c) is a phenomenon which cancels the effect of post-focus reduction after the FI domain. In (5b), where the FI domain is assumed to be (A B C), the compressed pitch range of the post-focus material (B and C) is reset to the original pitch range (horizontal dotted line) at the end of the FI domain. As a result, the phrase outside the FI domain, i.e. D, has the non-compressed pitch height.

This means that a pitch reset after the post-focus reduction will indicate the end point of the FI domain. In the indirect wh-question in (1b) above, for example, an FI is observed in the embedded clause: Focus boosting raises the F_0 -peak of the wh-phrase náni-o; the post-focus reduction compresses the pitch contour after the wh-phrase until the end of the embedded clause, where the Q-particle ka appears; and the pitch range is reset to the original height after the Q-particle. The FI domain in this case is between the wh-phrase and the Q-particle.

It should be mentioned here that downstep also shows pitch reset when its domain (i.e. MaP) ends. As mentioned above, I assume here that downstep and post-focus reduction are independent phenomena. Therefore, pitch resets for downstep and pitch reset for post-focus reduction can also occur independently. In the discussion below, I will use the term ‘pitch reset’ only for FI pitch reset effect, unless specifically mentioned otherwise. It is however very important to keep in mind that the two pitch reset phenomena can occur independently, because such a case will be found in the results of the experiment.

Lastly, it should also be made clear that in the experiment discussed below, all the phonetic effects are examined paradigmatically, not syntagmatically. That is, the existence/absence of phonetic effects is examined by comparing pitch contours, one in which these effects are expected and the other in which they are not. This distinction should be made clear because it could affect the interpretation of the results.⁶

In the next section, we examine the pitch contour of the wh-scrambling sentence.

3 Experiment

3.1 Stimuli

The following four sentence types are compared in the experiment. Below is one of the seven stimulus sets used in the experiment:⁷

(6) Stimulus set example

A. No scrambling, Non-wh-sentence

Náoya-wa [Mári-ga **rámu**-o nomíya-de nónda to] ímademo
 Naoya-Top Mari-Nom rum-Acc bar-Loc drank that even.now
 omótteru
 think
 ‘Naoya still thinks that Mari drank rum at the bar.’

B. No scrambling, Indirect wh-question

Náoya-wa [Mári-ga **náni**-o nomíya-de nónda **ka**] ímademo
 Naoya-Top Mari-Nom what-Acc bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’

C. Scrambling, Non-wh-sentence

rámu_i-o Náoya-wa [Mári-ga t_i nomíya-de nónda to] ímademo
 rum-Acc Naoya-Top Mari-Nom bar-Loc drank that even.now
 omótteru
 think
 ‘Naoya still thinks that Mari drank rum at the bar.’

D. Scrambling, Indirect wh-question

náni_i-o Náoya-wa [Mári-ga t_i nomíya-de nónda **ka**] ímademo
 what-Acc Naoya-Top Mari-Nom bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’

A and B are sentences with a canonical word order (i.e. no scrambling). B is an indirect wh-question, containing a wh-phrase and a Q-particle in the embedded clause. C and D are the scrambled versions of A and B, respectively. D is the wh-scrambling example, where the embedded wh-phrase is scrambled to the beginning of the matrix clause.

3.2 Predictions

Among the F₀ peaks in the sentences, those of the following two phrases are measured to examine the FI effects. They are labeled P1 and P2, respectively.

(7) Labels of the relevant F₀ peaks

[_{CP} ((Non-)WH) ... [_{CP} ... ((Non-)WH) ... Verb Q] α ...]
P1 P2

P1: Embedded clause verb

P2: Material immediately following the embedded clause

Expected FIs for the non-scrambled sentences (A and B) and for the scrambled sentences (C and D) are illustrated below, together with the predictions to be tested in the experiment. (Box indicates F₀-boosting, and underline indicates F₀-reduction.)

3.2.1 Non-scrambling Sentences (A and B)

We expect an FI only in B (i.e. indirect wh-question), and only in the embedded clause, as shown in (8).

(8) Non-scrambling sentences: A vs. B

A.	[_{CP} [_{TP} ... [_{CP} [_{TP} ... Non-WH ... P1] C]	P2 ...]
	↑	↑
	No reduction	No reduction
B.	[_{CP} [_{TP} ... [_{CP} [_{TP} ... WH ... P1] ka]	P2 ...]
	↑	↑
	Reduction	No reduction

P1 is reduced in B due to the F₀-reduction, but not in A. Therefore we expect a difference in the height of P1 between A and B. P2, on the other hand, would not show any difference between A and B, since no material in the matrix clause is affected by the FI in the embedded clause, given the FI-Wh-scope Correspondence. Accordingly, we have the following two predictions:

(9) Predictions for A and B

- a. P1: A > B
- b. P2: A = B

3.2.2 Scrambled Sentences (C and D)

There are two possible expected results for the scrambling sentences (C and D), depending on how the FI domain in D is realized.

(i) If the FI domain = the embedded CP: First of all, in the non-wh-scrambling sentence C, no FI is expected, just like its non-scrambled counterpart (=A). Therefore the F₀-peaks of P1 and P2 are not reduced. If the FI domain is the embedded clause,

as I (Ishihara (2002)) and Kitagawa and Fodor (2003) claimed, the F_0 -reduction ends at the end of the embedded clause, and the pitch reset takes place thereafter. In this case, we expect a difference only in P1. P2 would be the same between C and D.

(10) Scrambling sentences: C vs. D (FI domain = embedded CP)

C.	[_{CP} Non-WH _i [TP ... [CP [TP ... <u>t_i</u> ... P1] C]	P2 ...]]
	↑	↑
	No reduction	No reduction
D.	[_{CP} WH_i [TP ... [CP [TP ... <u>t_i</u> ... P1] ka]	P2 ...]]
	↑	↑
	Reduction	No reduction

(ii) If the FI domain = the matrix CP: If the FI domain is the matrix clause, the F_0 -reduction continues until the end of the sentence. Accordingly, both P1 and P2 will be lowered in D. In this case, therefore, we expect that both P1 and P2 of C are higher than those of D.

(11) Scrambling sentences: C vs. D (FI domain = matrix CP)

C.	[_{CP} Non-WH _i [TP ... [CP [TP ... <u>t_i</u> ... P1] C]	P2 ...]]
	↑	↑
	No reduction	No reduction
D.	[_{CP} WH_i [TP ... [CP [TP ... <u>t_i</u> ... P1] ka]	P2 ...]]
	↑	↑
	Reduction	Reduction

Accordingly, we have the following predictions:

(12) Predictions for C and D

- a. P1: C > D
- b. P2: (i) C = D, if the FI domain is the embedded CP.
- (ii) C > D, if the FI domain is the matrix CP.

3.3 Method

Subjects: Four females, AH, CS, CK, NM, and a male, YY, all non-linguists brought up in Tokyo or surrounding areas.

Presentation of the stimuli: Stimuli (7 sets of 4 sentence types, total of 28 sentences, see §3.1) are mixed with 108 filler sentences, provided in a pseudo-randomized order (so that two sentences from the same example set are not presented in a row). Each sentence is presented to the subject on a computer screen, one at a time. Each subject makes 3 recordings of the entire set of stimuli. Each recording session uses a different pseudo-randomized order of the sentences.

Task: Subjects are asked first to read the sentence (either aloud or quietly) to understand the meaning of the sentence, and then to read aloud for the recording.

Data exclusion: The results are first analyzed for each subject. After the examination of the data, one of the five subject's (NM) data is excluded in the final analysis. In NM's data, not only the expected contrasts, but also other syntax/semantics-related phenomena expected in an utterance (e.g. downstep, utterance final rising intonation for questions) were not attested. The data only showed the time-dependent declination effect. This fact suggests that the subject did not pay sufficient attention to the syntax/semantics of the sentences, and read them merely as sequences of words. Such data would not tell us anything important for our purpose.

Data normalization: The data from four of the five subjects (excluding NM's data) are normalized to see if the embedded FI can be observed as a general property among these speakers. All the measured values are transformed according to the following linear transformation:

$$\text{transformed_value} = (\text{original_value} - \text{Av}_S(\text{P1})) / (\text{Av}_S(\text{P2}) - \text{Av}_S(\text{P1}))$$

where $\text{Av}_S(\text{P}_n)$ is the speaker-specific mean F_0 -value of reference point P_n . This formula rescales for each speaker the mean of P1 measurements to 0 and the mean of P2 measurements to 1.

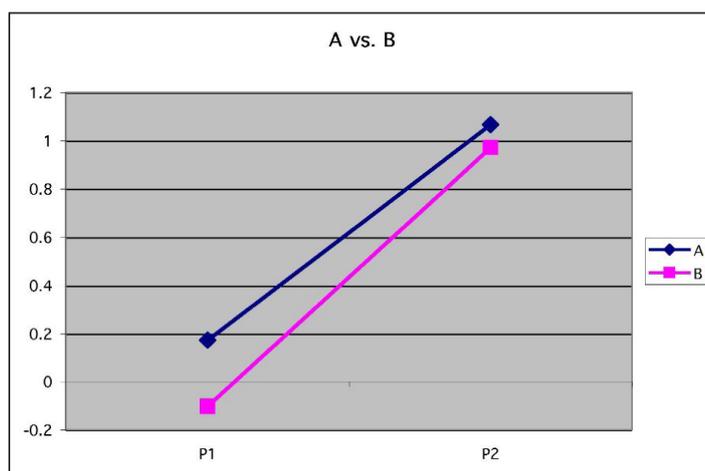
3.4 Results

3.4.1 Non-scrambling Sentences A and B

For the non-scrambled sentences A and B, the predictions in (9) are borne out. P1 (the embedded verb) is lower in B than in A. This difference is statistically significant (two sided t-test, $t(71) = 3.604$, $p < 0.001$). This difference indicates that in B, post-focus reduction takes place after the wh-phrase, and compresses the pitch range of the material thereafter, including P1. Since the post-focus reduction effect is not expected in A, we observe a difference between A and B.

On P2 (the post-embedded-CP phrase), although there is still a difference between A and B, it is much smaller compared to the one on P1. In fact, this difference is statistically not significant (2 sided t-test, $t(71) = 1.143$, $p = 0.2567$). This indicates that the post-focus reduction effect observed in the embedded clause in B is no longer in effect outside the embedded clause. That is, a pitch reset takes place after the embedded clause, and the pitch range of P2 is set back to the non-reduced value. The non-reduced value is namely the value of P2 in sentence A, where no FI is observed.

(13) A vs. B

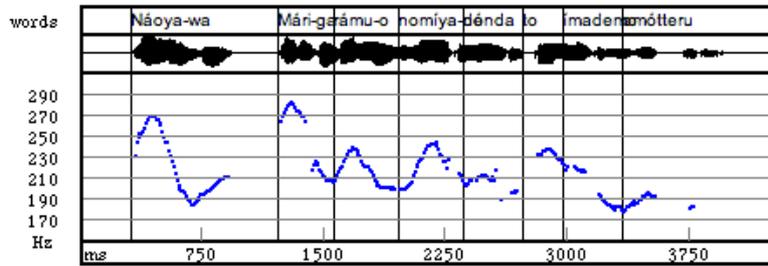


As one can see, there is an F_0 -rise from P1 to P2 in both A and B, regardless of the sentence type. This rise is due to downstep on the embedded verb (P1), followed by the downstep-related pitch reset on the following phrase (P2). According to the syntax-phonology mapping principle of Selkirk and Tateishi (1991) and others, no MaP boundary is expected before the transitive verb P1. Accordingly, P1 is subject to downstep, and hence realized with a low F_0 . On the other hand, a MaP boundary is expected before P2, which cancels the effect of downstep. Therefore a (downstep-related) pitch reset takes place, and boosts the F_0 -peak on P2 up to the normal (i.e. non-downstepped) height. Because of these effects, we expect a certain amount of F_0 -rise from P1 to P2, regardless of the existence/absence of FI.

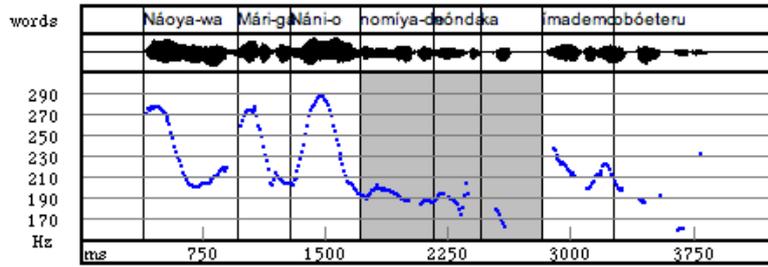
As explicated in §2.3, I assume here that downstep and FI are independent phenomena. Downstep on the embedded verb and the subsequent (downstep-related) pitch reset is part of the default intonation pattern that is observed in all the stimuli (A–D), and has nothing to do with the contrast that we are interested in, namely, the contrast caused by post-focus reduction and the FI-related pitch reset.

Example pitch contours for A and B are given in (14) below (shading indicates the domain of post-focus reduction). It is clear that in B the wh-phrase náni-o is strongly boosted while the pitch range of the following phrases is compressed. Pitch reset after the embedded clause is also clearly observed.

(14) a. Example pitch contour of A



b. Example pitch contour of B

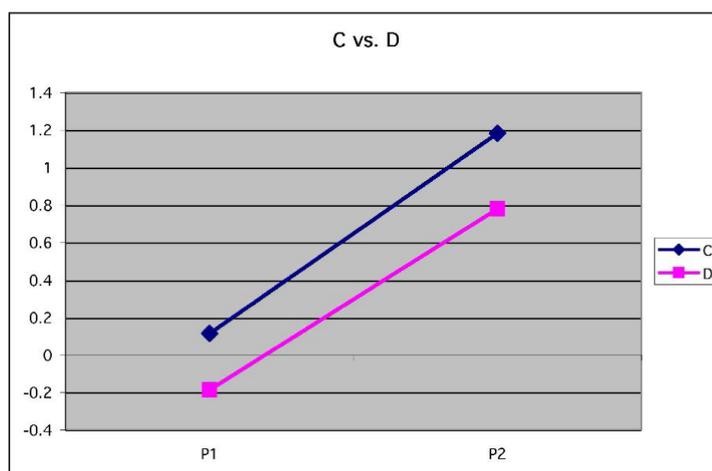


3.4.2 Scrambled Sentences C and D

In the scrambled sentences C and D, the results show that the FI domain in D is the matrix clause. First of all, P1 is significantly lower in D than in C due to the post-focus F_0 -reduction (2 sided t-test, $t(71)=3.444$, $p<0.001$), just like in the non-scrambled cases above, confirming (12a).

On P2, the differences between C and D are not reduced at all. In fact, they are still statistically significant (2 sided t-test, $t(71)=4.399616382$, $p<0.0001$), confirming (ii) in (12b). This means that the F_0 -reduction effect continues to the matrix material, reducing not only the F_0 -peak of the embedded verb (P1), but also that of the following phrase in the matrix clause (P2). This result is different from that of the non-scrambled version, where the post-focus reduction ends at the embedded clause.

(15) C vs. D

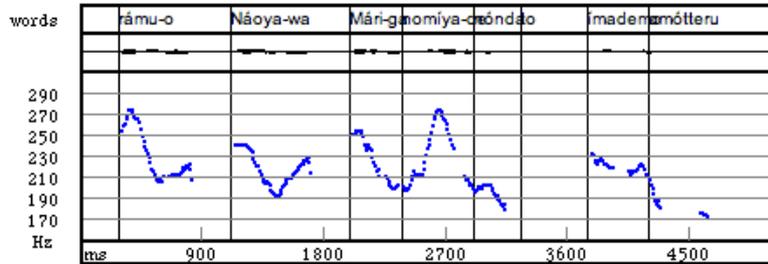


Note, again, that in both C and D there is a F_0 -rise from P1 to P2. As shown above, however, this F_0 -rise is due to the downstep-related pitch reset, which is motivated independently of FI. As we discussed in §2.3, our data should be examined paradigmatically, not syntagmatically. Looking at the results for D only, one might conclude that the post-focus reduction in D ends at the embedded clause, because there is an F_0 -rise from P1 to P2, which could be interpreted as a pitch reset in a syntagmatic interpretation of the data here. Such a conclusion is not tenable in a paradigmatic examination of the data. Since F_0 -rise is observed in both C and D, i.e., the F_0 -rise is observed with or without post-focus reduction, we should not attribute this F_0 -rise to FI. The pure effect of post-focus reduction in fact continues to P2, where the difference between C and D is still statistically significant. Therefore, we conclude that both P1 and P2 are lower in D than in C, confirming the predictions (12a) and (ii) of (12b). This means that the FI domain in D continues until the end of the matrix clause.

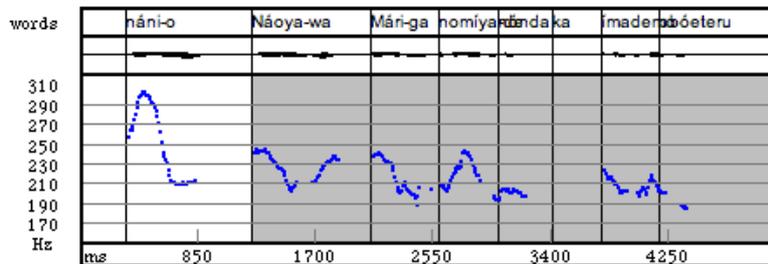
Example pitch contours for C and D are given in (16). In this pair of contours, the difference are not as clear as in the case of non-scrambled pair (A and B). It is observable, however, that the scrambled wh-phrase in D bears the higher F_0 than the scrambled non-wh-phrase in C. It is also generally the case that the rest of the sentence has a slightly narrower pitch range than in non-wh-scrambling case. Note that the F_0 -peak of the matrix adverbial phrase in D ímademo is not as high as in the non-wh-scrambling case

in C, and hence stays as low as embedded clause phrases (Mári-ga or nomíya-de). This indicates that there is no (FI-related) pitch reset effect.

(16) a. Example pitch contour of C



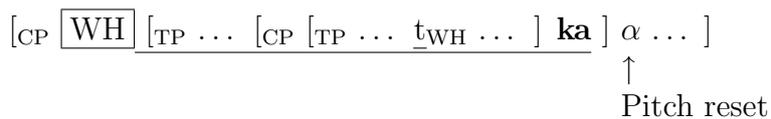
b. Example pitch contour of D



3.5 Discussion

The results presented here contradict the earlier observation in Ishihara (2002) and Kitagawa and Fodor (2003) that the FI domain of the wh-scrambling sentence is between the scrambled wh-phrase and the Q-particle at the end of the embedded clause (cf. §2.2). Contrary to their claim, the F₀-reduction continues until the end of the sentence, suggesting that the FI domain of the wh-scrambling sentence is the matrix clause.

(17) a. *Observation in Ishihara (2002), Ishihara (2003) (=3)



b. Actually attested pitch contour

$$\begin{array}{c}
[\text{CP } \boxed{\text{WH}} [\text{TP } \dots [\text{CP } [\text{TP } \dots \underline{t}_{\text{WH}} \dots] \mathbf{ka}] \alpha \dots] \\
\uparrow \\
\text{F}_0\text{-reduction}
\end{array}$$

This means that wh-scrambling is an exceptional case in terms of the FI–Wh-scope Correspondence observed elsewhere. In the wh-scrambling sentence, its FI domain is the matrix clause, to which the wh-phrase is scrambled, while the semantic scope of the wh-question remains in the embedded clause. This results in FI–Wh-scope Mismatch. This fact is particularly important because it suggests that FI–Wh-scope Correspondence is not a result of the direct phonology-semantics interaction. If it were the case, we would expect the prosody-scope correspondence regardless of the existence/absence of scrambling, which has been claimed to be semantically vacuous (Saito (1989)).

In the next section, I propose my analysis that explains both the FI–Wh-scope Match and Mismatch cases.

4 Multiple Spell-Out Account

In Ishihara (2003, 2004), I proposed an analysis that explains both cases where the FI domain and the wh-scope match and those where they mismatch. Adopting the notion of phase and Multiple Spell-Out in the recent Minimalist framework (Chomsky (2000, 2001a,b)), I claimed that the FI–Wh-scope Correspondence is a result of the cyclic computation, which, in normal cases, computes the domain of FI and the wh-scope at the same cycle.⁸ This analysis also correctly predicts the FI–Wh-scope Mismatch case we discussed above. In this analysis it is predicted that the syntactic movement creates a mismatch between the computation cycle at which the FI is created and the one at which the semantic wh-scope is fixed.

4.1 Proposal: Phase-by-Phase Prosody Computation

The main proposal of this analysis is that prosody is computed phase-by-phase. When a FOC-feature is involved in a derivation, an FI is created at a particular Spell-Out cycle. The size of the FI domain depends on at which Spell-Out cycle the FOC-feature is transferred to Φ . Below I will introduce the key notions and assumptions for the phase-by-phase prosodic computation.

Phase, Multiple Spell-Out: First of all, the analysis assumes phase and Multiple Spell-Out, as follows:

(18) Multiple Spell-Out (Chomsky (2001a))

- a. CPs and *v*Ps are phases.
- b. When a syntactic derivation reaches a phase (*v*P/CP) in the narrow syntax (NS), the complement of the phase head (i.e. VP/TP) is transferred to the interface levels (Φ/Σ). The phonological part of the Transfer (NS \rightarrow Φ) is called Spell-Out.

$[\text{CP (Spec) C}$	$[\text{TP (Spec) T}$	$[\text{vP (Spec) } v$	$[\text{VP ... }]]]]$
↑	↑	↑	↑
phase	Spell-Out	phase	Spell-Out

- c. A phrase in a phase ‘edge’ position (i.e. Spec,*v*P/CP) is outside the Spell-Out domain, and hence accessible to later syntactic operations (which means, in reverse, that it is inaccessible to the interface operations).

FOC-features on wh-phrase/Q-particle: I assume two FOC-features, one on the Q-particle, the other on the wh-phrase (cf. Deguchi and Kitagawa’s (2002) E-agreement analysis).

(19) FOC-features on wh-phrase/Q-particle

- a. An uninterpretable FOC-feature on Q-particle (FOC_Q)
- b. An interpretable FOC-feature on WH (FOC_{WH})

Agree: These two features establish an Agree relation within a phase.⁹ The FOC_Q-feature deletes on Agree.

$$(20) \quad \underline{\text{Agree}}$$

$$[\text{CP} \dots \text{WH}_{\text{FOC}} \dots \text{Q}_{\text{FOC}}]$$

$$\underbrace{\hspace{10em}}_{\underline{\text{Agree}}}$$

After establishing an Agree relation with FOC_Q, FOC_{WH} enters into Φ via Spell-Out operation.

$$(21) \quad \underline{\text{Spell-Out}}$$

$$[\text{CP} \underbrace{[\text{TP} \dots \text{WH}_{\text{FOC}} \dots]}_{\underline{\text{Spell-Out}}} \text{Q}]$$

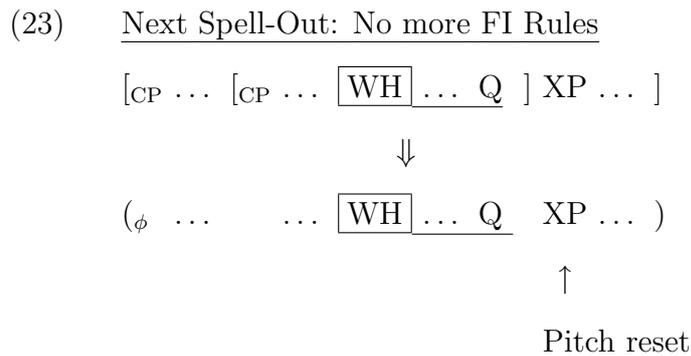
FI Rules: At each Spell-Out, if any FOC is found in Φ , an FI is assigned to the prosodic domain ϕ (see below) as a phonological realization of FOC. FIs are created according the following FI Rules:¹⁰

$$(22) \quad \underline{\text{FI Rules}}$$

- a. Focus F₀-boosting: Boost the F₀ of any phrase with a FOC-feature in ϕ .
- b. Post-focus F₀-reduction: Reduce the F₀ thereafter.

$$[\text{TP} \dots \text{WH}_{\text{FOC}} \dots] \implies (\phi \dots \boxed{\text{WH}} \dots)$$

After FI Rules apply to ϕ and an FI is created, FOC-feature deletes.¹¹ Once deleted, it will not induce any more FI at a later Spell-Out. Any material that is introduced to the derivation at a later Spell-Out is not affected by any previously created FI.



Prosodic domain ϕ : I assume here that the phonological material that is transferred into Φ forms a prosodic domain ϕ , which contains all the phonological material in Φ , i.e., including everything that is transferred into Φ at earlier Spell-Out cycles. This means that ϕ incrementally becomes bigger and bigger as Spell-Out takes place cyclically. This assumption appears to go against the spirit of Minimalist framework, especially with notions like phase and Multiple Spell-Out, since these notions are introduced to limit the narrow syntax (NS) computation to reduce the computational complexity. In NS, the computation access is limited to the material within each phase. No access is allowed to the material that is already transferred to interface levels at an earlier phase (Phase Impenetrability Condition, Chomsky (2000, 2001a,b)). In contrast to the computation at NS, however, this incremental computation appears to be an inevitable aspect of the interface level Φ , given that all the material has to be combined at Φ for actual pronunciation. Moreover, it is also inevitable that all the material at Φ be compiled in order to compute suprasegmental prosodic information such as intonation, because it has to deal with phonological content bigger than each Spell-Out domain.¹²

It is important, however, to note that at each Spell-Out, the content of ϕ is updated. Therefore the timing of the application of FI Rules (22) becomes crucial. If they apply at an early Spell-Out cycle, the prosodic domain ϕ contains a relatively small amount of phonological material. Therefore the domain of FI would be small. On the other hand, if (22) applies at a later Spell-Out cycle, the domain of the created FI would be larger.

Q-particle: Note that Q-particles are phase heads, and hence, are outside the Spell-Out domain. Phonologically, however, they seem to be included in the F_0 -reduction domain.

$$(24) \quad \text{Q-particle: outside the Spell-Out domain, but inside the FI}$$

$$[\text{CP } \underbrace{[\text{TP } \dots \text{WH}_{\text{FOC}} \dots]}_{\text{Q outside the Spell-Out domain}} \text{ Q}] \implies (\phi \dots \boxed{\text{WH}} \underbrace{\dots \text{Q}}_{\text{Q inside the } F_0\text{-reduction domain}})$$

I speculate that this is because Q-particles do not behave as Prosodic Words (PWd) by themselves and have no ability to create a new prosodic boundary at any level (Minor Phrase, Major Phrase, or Intonation Phrase). Hence, I assume that they always cliticize to the preceding phrase (i.e. verb). Their F_0 is therefore always dependent on that of the preceding phrase.¹³

Root Spell-Out: If there is any phrase at the Spec or the head of the matrix CP, these phrases are not sent to Φ at the Spell-Out of the matrix CP phase, because the Spell-Out domain of this phase is the matrix TP. Therefore I assume that there is another Spell-Out process, which I will call Root Spell-Out, at the very last stage of the derivation.

$$(25) \quad \text{Root Spell-Out}$$

- a. Matrix CP Spell-Out
 $[\text{CP } \text{XP } [\text{TP } \dots] \text{ Comp}] \implies (\phi \dots)$
- b. Root Spell-Out
 $[\text{CP } \text{XP } [\text{TP } \dots] \text{ Comp}] \implies (\phi \text{ XP } \dots \text{ Comp})$

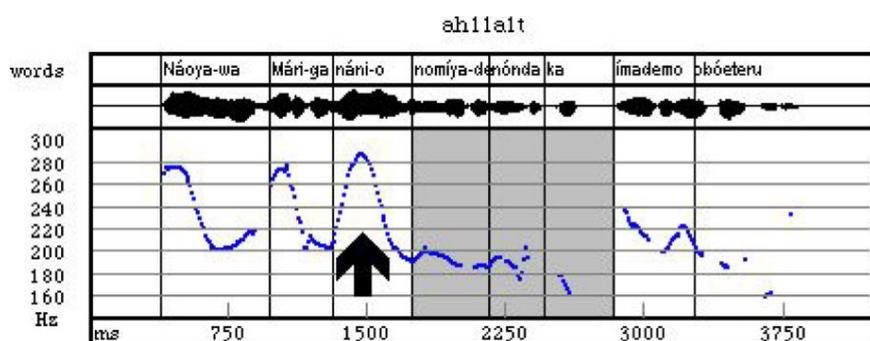
4.2 Examples

This analysis can account for the pitch contours of both the FI-Wh-scope Match case and Mismatch case. In this subsection I will illustrate how the analysis works.

4.2.1 Indirect Wh-questions

The first case is the indirect wh-question (without wh-scrambling) as in (1b). As the pitch contour below shows, the FI starts at the wh-phrase and ends at the end of the embedded clause.¹⁴ Let us look at the computation process phase-by-phase.

- (1b) [CP₂ Náoya-wa [vP₂ [CP₁ Mári-ga [vP₁ **náni**_{FOC-o} nónda] **ka**_{FOC}]
 Naoya-Top Mari-Nom what-Acc drank Q
 obóeteru]]
 remember
 ‘Naoya remembers what Mari drank.’



Embedded vP phase (vP1): At this phase, the wh-phrase moves to the Spec,vP by invisible wh-movement (see fn. 9). (In the example hereafter, we assume FOC-feature movement for expository purpose.) The Spell-Out domain (VP1) contains no FOC-feature that has Agreed. Accordingly, no FI is created at this Spell-Out. The intonation of the prosodic domain ϕ is created according to the default syntax-prosody mapping mechanism (cf. Selkirk and Tateishi (1991)).

- (26) vP1 phase (Spell-Out: VP1)—No FI created

[vP₁ FOC_{WH} [VP₁ **náni**_{t_{FOC}}-o nónda] v]
 ↑
 Invisible wh-movement (FOC_{WH} not yet Agreed with FOC_Q)
 ⇒ (ϕ **náni**-o nónda)

(29) CP2 phase (Spell-Out: TP2)—No FI created

[_{CP2} [_{TP2} Náoya-wa [_{CP1} Mári-ga náni-o nónda ka] obóeteru]]
 \Rightarrow (ϕ Náoya-wa Mári-ga náni-o nónda ka obóeteru)

(30) Root phase (Spell-Out: CP2)—No FI created

\Rightarrow (ϕ Náoya-wa Mári-ga náni-o nónda ka obóeteru) (Final output)

The resultant representation correctly illustrates the expected pitch contour—the F_0 -boosting on the wh-phrase; the post-focus F_0 -reduction up to the end of the embedded clause; and the pitch reset at the matrix material thereafter.

4.2.2 Wh-scrambling

Let us now turn to the wh-scrambling case. Recall that the experiment showed that the FI induced by the scrambled wh-phrase does not stop at the end of the embedded clause, but continues until the end of the matrix clause. The Multiple Spell-Out analysis predicts the correct intonation for a wh-scrambling case like (2b).

(2b) [_{CP2} náni_i-o [Náoya-wa [_{CP1} t_i [Mári-ga t_i nónda] ka] obóeteru]]
 what-Acc Naoya-Top Mari-Nom drank Q remember
 ‘Naoya still remembers what_i Mari drank t_i.’

Before we examine the derivation for (2b), I will assume, following Mahajan (1994), Miyagawa (1997, 2001, and others), that the landing site of A' -scrambling (including all instances of long-distance scrambling) is Spec,CP.¹⁵

The expected pitch contour is created as follows. (In the example below, I omit the vP phases for brevity. Although the wh-movements to Spec, vP take place at these phases, these movements do not directly affect the output in Φ in the relevant sense, because FOC-features never establish an Agree relation at these phases.)

Embedded CP phase (CP1): The FOC_{WH} -feature in $\text{Spec},v\text{P1}$ (after the wh-phrase moving to this position at the previous $v\text{P}$ phase) first establishes an Agree relation with FOC_Q . In principle, this feature is ready to enter into Φ and induce FI at this point. In the wh-scrambling case, however, the wh-phrase escapes from the Spell-Out domain (TP1) by scrambling to the Spec,CP . Accordingly, ϕ contains no FOC-feature at this Spell-Out. As a result, no FI is assigned at this point.

(31) CP1 phase (Spell-Out: TP1)—No FI created

$$\begin{array}{c}
 [_{\text{C}'} [_{\text{TP1}} \text{Mári-ga} [_{v\text{P1}} \text{náni}_{i\text{FOC-O}} [_{\text{VP}} \underline{t}_i \text{ nónda}]]]] \mathbf{ka}_{\text{FOC}}] \\
 \underbrace{\hspace{15em}}_{\text{Agree}} \\
 [_{\text{CP1}} \text{náni}_{i\text{FOC-O}} [_{\text{TP1}} \text{Mári-ga} [_{v\text{P1}} \underline{t}_i [_{\text{VP}} \underline{t}_i \text{ nónda}]]]] \mathbf{ka}] \\
 \underbrace{\hspace{10em}}_{\text{Wh-scrambling}} \\
 \implies (\phi \text{ Mári-ga } \underline{t}_i \text{ nónda})
 \end{array}$$

Matrix CP phase (CP2): The scrambled wh-phrase (and its FOC_{WH} -feature) keeps escaping from the Spell-Out domain by moving to each phase edge position successively, and reaches $\text{Spec},\text{CP2}$. Note that the FOC_{WH} -feature is still outside the Spell-Out domain of this phase (i.e. TP2). Therefore no FI has been created yet.

(32) CP2 phase (Spell-Out: TP2)—No FI created

$$\begin{array}{c}
 [_{\text{CP2}} \text{náni}_{i\text{FOC-O}} [_{\text{TP2}} \text{Náoya-wa} [_{\text{CP1}} \underline{t}_i [_{\text{TP1}} \text{Mári-ga } \underline{t}_i \text{ nónda}]] \mathbf{ka}] \text{ obóeteru}] \\
 \implies (\phi \text{ Náoya-wa} \text{ Mári-ga } \underline{t}_i \text{ nónda} \mathbf{ka} \text{ obóeteru})
 \end{array}$$

Root Spell-Out: The required FI is created at the Root Spell-Out. As a result, the whole sentence becomes the domain of the FI. Therefore no pitch reset is expected. This is exactly what we found in the result of the experiment.

(33) Root Spell-Out (CP2)—FI created

$$[_{\text{CP2}} \text{náni}_{i\text{FOC-O}} [_{\text{TP2}} \text{Náoya-wa} [_{\text{CP1}} \underline{t}_i [_{\text{TP1}} \text{Mári-ga } \underline{t}_i \text{ nónda}]] \mathbf{ka}] \text{ obóeteru}]$$

\Rightarrow (ϕ **náni**_i-o Náoya-wa Mári-ga t_i nónda **ka** obóeteru) (Final output)

4.3 Derivational Model vs. Representational Model

So far I have illustrated how the Multiple Spell-Out analysis proposed in Ishihara (2003, 2004) accounts for the FI-Wh-scope Correspondence as well as the Mismatch. The FI-Wh-scope Correspondence/Mismatch is captured by the phase-by-phase computation of wh-scope and FI-prosody with respect to the FOC-feature. In the usual case of matching, the wh-scope and the FI-prosody are established at the same phase. When the wh-phrase is scrambled out of the wh-scope, the creation of FI-prosody is delayed to a later Spell-Out cycle. This delay induces the mismatch.

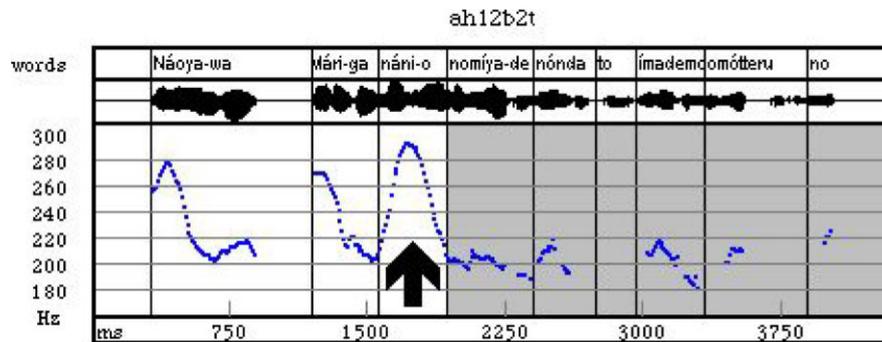
The biggest advantage of this cyclic, derivational analysis is that we can account for both the cases discussed above without any additional rules or assumptions. If we assume a non-cyclic, representational analysis of FI such as the standard analyses by Nagahara (1994), Truckenbrodt (1995), Sugahara (2003), among others, we will run into a problem one way or the other.

First of all, none of the standard analyses has taken the FI-Wh-scope Correspondence into account. In these analyses, everything that follows a focused phrase will be included in a single MaP phrase. That is to say, MaP boundaries after the focused phrase will be deleted until the end of the sentence. In the case of non-scrambled embedded wh-question like (1b), repeated below, the standard analysis would wrongly predict that the FI would be formed between the embedded wh-phrase náni-o until the end of the sentence, just like in the case of matrix wh-question like (1a).

(1) a. Matrix wh-question

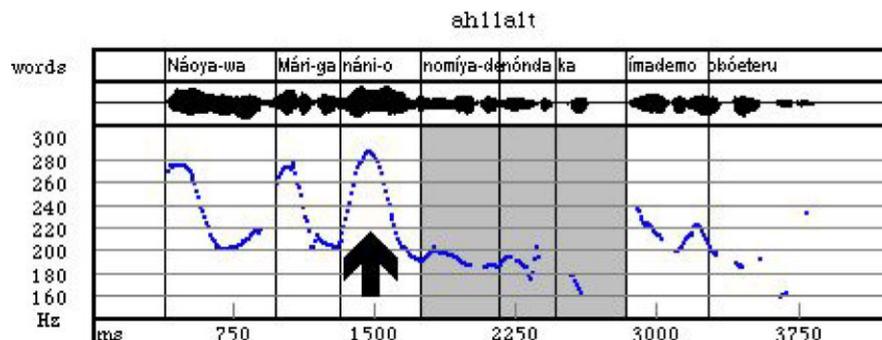
Náoya-wa [Mári-ga **náni**-o nomíya-de nónda to] ímademo
 Naoya-Top Mari-Nom what-Acc bar-Loc drank that even.now
 omótteru **no**?
 think Q

‘What_i did Naoya still think that Mari drank t_i at the bar?’



b. Indirect wh-question

Náoya-wa [Mári-ga **náni-o** nomíya-de nónda **ka**] ímademo
 Naoya-Top Mari-Nom what-Acc bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’



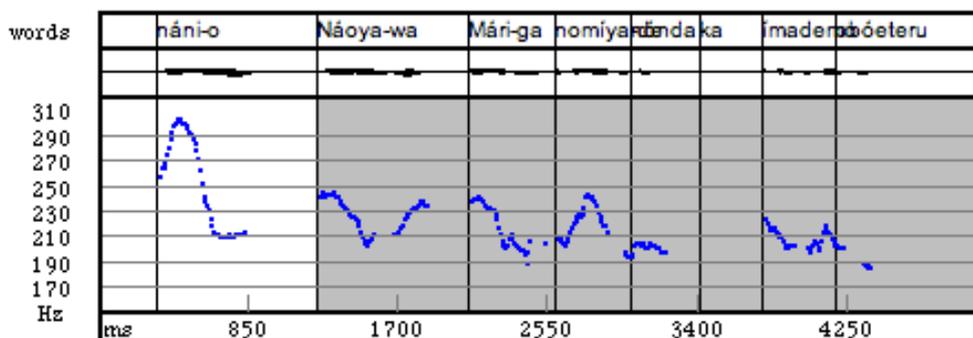
Since we already know that this prediction is not correct, a modification would be needed to account for the FI-Wh-scope Correspondence. The question is how we explain the prosody-scope correspondence in a representational analysis. One possibility is to postulate a rule or constraint that determines the end point of FI (or the right edge of MaP that contains a focus). Such a rule or constraint would state that the post-focus MaP boundary deletion (which corresponds to the post-focus reduction in the analysis proposed in this paper) continue within the scope of the focus (i.e. the wh-scope). Alternatively, the rule/constraint would state that the MaP boundary deletion continue until the Q-

particle ka. In this way, it is possible to restrict the domain of MaP within the scope of focus or wh-question.

Once we postulate such a rule/constraint, however, we run into trouble in the Mismatch case like (2b), repeated below. Recall that in the Mismatch case, the FI continues until the end of the matrix clause, i.e. outside the wh-scope, unlike the non-scrambling case like (1b).

(2b) Indirect wh-question with wh-scrambling

náni_i-o Náoya-wa [Mári-ga t_i nomíya-de nónda **ka**] ímademo
 what-Acc Naoya-Top Mari-Nom bar-Loc drank Q even.now
 obóeteru
 remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’



In a representational analysis, it is necessary to postulate a rule/constraint that determines the end point of FI, because the end point of FI is not derived automatically from the system itself. If one postulates a phonology-syntax interface rule/constraint to accommodate the prosody-scope correspondence, it causes a problem in the Mismatch case. If we give up such a rule/constraint, the FI-Wh-scope Correspondence will be left unexplained.

In the cyclic, derivational analysis proposed here, the end point of FI will be derived automatically. The phonological system does not need to specify the end point or the right edge of the FI domain. That is, phonological rules (FI rules in (22)) remain as simple as possible. Furthermore, the FI-Wh-scope Correspondence will be derived as a

result of the computation. There is no need to stipulate such a phonology-syntax interface condition. In such an account, the Mismatch case can be also accounted for without any additional complication in the grammar.

5 Remaining Questions

5.1 Initial Observation by Ishihara (2002) and Kitagawa and Fodor (2003)

There are a few remaining questions to be answered, which I will leave for future research. First, given that the experiment shows results that contradicts the previous observation by Ishihara (2002) and Kitagawa and Fodor (2003), it becomes a mystery why both acknowledged that (17a), repeated below, is the correct pitch contour. In fact, I, as a native speaker, still feel that (17a) is not entirely impossible.

- (17) a. *Observation in Ishihara (2002), Ishihara (2003) (=3)
- $$[\text{CP } \boxed{\text{WH}} [\text{TP } \dots [\text{CP } [\text{TP } \dots \underline{\text{t}_{\text{WH}} \dots}] \mathbf{ka}] \alpha \dots]]$$
- ↑
Pitch reset
- b. Actually attested pitch contour
- $$[\text{CP } \boxed{\text{WH}} [\text{TP } \dots [\text{CP } [\text{TP } \dots \underline{\text{t}_{\text{WH}} \dots}] \mathbf{ka}] \alpha \dots]]$$
- ↑
F₀-reduction

It is, however, hard to decide whether this intuition is real and has to be accounted for, because this sentence always involves unnaturalness in judgement,¹⁶ and also because I may be too sensitive to the FI-Wh-scope Correspondence to give a naïve judgement.

One possible explanation, pointed out by an anonymous reviewer, is that an impressionistic (or syntagmatic) examination of the pitch contour led to an incorrect conclusion. As we have seen in the experiment, there is always an F₀-rise from the embedded clause verb (P1 in the experiment) to the post-embedded-CP phrase (P2), regardless of the existence/absence of post-focus reduction effect (see (15)). This F₀-rise is due to the pitch

reset of downstep. If one only looks at the wh-scrambling sentence (D in the experiment), however, s/he may interpret this F_0 -rise as an indication of the end point (i.e. pitch reset) of post-focus reduction.

Another possible explanation is that a wh-scrambling sentence causes a conflict between the production mechanism and some perception/processing principle, due to the FI-Wh-scope Mismatch. Suppose there is a perception principle that a hearer interprets wh-scope according to the FI found in the sentence. If this principle is at work, the wh-scrambling sentence with the intonation in (17a) would be wrongly interpreted as a matrix question, which does not correspond to its syntactic structure.¹⁷ Given that, when one produces the wh-scrambling sentence with a hearer in mind, s/he would try to maximize the correspondence (i.e. minimize the mismatch) between the FI domain and wh-scope so that the hearer would not have such a misinterpretation. As a result, he/she would try to produce a contour like (17a) instead of (17b).

In the production experiment in this paper, where no specific hearer is expected, however, subjects might have only followed the production mechanism, which would create a contour like (17b) (assuming that the Multiple Spell-Out model is correct).

If (17a) turns out to be a possible output form for the wh-scrambling sentence, there needs to be some additional operation that creates a contour like (17a), since the Multiple Spell-Out model per se does not induce a pitch reset after the embedded clause.

The interaction between the production mechanism and perception/processing mechanism appears to be an important key in answering this question. I do not discuss perception/processing issues any further, and leave them for future research. For the moment, I take the results of the experiment as the real and correct description of the fact, and propose that the Multiple Spell-Out model is the correct production model that accounts for them.

5.2 Predictions about Clause-Initial Scrambling

The proposed Multiple Spell-Out model makes another prediction which remains to be tested. Suppose that the wh-phrase of an indirect wh-question like (34a) moved to the beginning of the embedded clause, but not any further, as in (34b).

- (34) a. Náoya-wa [_{CP1} Mári-ga **náni**-o nónda **ka**] obóeteru
 Naoya-Top Mari-Nom what-Acc drank Q remember
 ‘Naoya remembers what_i Mari drank t_i.’
- b. Náoya-wa [_{CP1} **náni**_i-o Mári-ga t_i nónda **ka**] obóeteru
 Naoya-Top what-Acc Mari-Nom drank Q remember

The expected pitch contour for (34b) varies depending on the theoretical assumptions about the landing site of the wh-scrambling.

In contrast to the long-distance scrambling like in (2b), which always exhibits A'-properties, clause-initial scrambling has been claimed to exhibit both A- and A'-properties (Mahajan (1990), Tada (1993), among others). Mahajan (1994) and Miyagawa (1997, 2001) claim that there are two different types of scrambling, A-scrambling and A'-scrambling, and that they have different landing sites, namely, Spec,TP for A-scrambling, Spec,CP for A'-scrambling.¹⁸

If the scrambling in (34b) is an A-scrambling, the landing site is Spec,TP. Therefore the wh-phrase is transferred to Φ at the Spell-Out of the embedded CP phase. In this case, the FI will be observed only inside the embedded clause, and a pitch reset takes place after the embedded clause.

If the scrambling in (34b) is an A'-scrambling, the landing site is Spec,CP. In this case the wh-phrase escapes from the Spell-Out of the embedded CP phase. It is instead transferred to Φ at the next Spell-Out cycle, i.e. at the matrix *v*P phase, which includes the matrix verb. In this case, the FI will include the whole embedded clause and the matrix verb. This essentially means that FI-Wh-scope Mismatch occurs even though the wh-phrase is not moved out of the wh-scope.

One obvious question is whether this prediction is in fact correct. This question can be answered only by testing it. There seem, however, to be a few technical difficulties in testing this prediction. How to control the type of scrambling in the experiment is not an obvious issue, for example. It would also be important to check the end point of the FI in the Mismatch case. In (34b), the FI is expected to end at the end of matrix verb obóeteru ‘remember’ (assuming A'-scrambling). This means that if we embed (34b) into another clause, we expect a pitch reset thereafter. Sentences with such deep embedding, however, might be too complicated for phonetic experimentation. Given these complications, these issues remain to be considered further.

6 Conclusion

In this paper, I discussed the Focus Intonation–Wh-scope Correspondence found in Tokyo Japanese wh-questions. In particular, I showed, based on the experimental data, that when the wh-phrase is scrambled out of its scope, an FI–Wh-scope Mismatch occurs. Then I illustrated how the Multiple Spell-Out account proposed in Ishihara (2003, 2004) correctly accounts for both the matching and mismatch cases.

Furthermore, the discussion in this paper has an important theoretical implication. The FI–Wh-scope Match and Mismatch phenomena, along with the Multiple Spell-Out account, provide empirical evidence for the effect of phase in phonology. The phase theory was originally proposed in the Minimalist framework in order to account for various syntactic mechanisms and computational efficiency. It is, however, natural to expect that the theory has impacts on phonology, since the interaction between narrow syntax and the PF interface level becomes more frequent than the traditional Y-model assumed in the GB framework (Chomsky (1981)) or in an earlier stage of Minimalist framework (Chomsky (1995)). The discussion in this paper shows that cyclic phase-by-phase computation does indeed have an effect on phonology, in particular, on the focus intonation in Tokyo

Japanese.

Bibliography

- Bresnan, Joan W. (1972) “On Sentence Stress and Syntactic Transformation,” Contributions to Generative Phonology, ed. by Michael Brame, 73–107, University of Texas Press, Austin.
- Chomsky, Noam (1981) Lectures on Government and Binding: The Pisa Lectures, Foris, Dordrecht.
- Chomsky, Noam (1995) Minimalist Program, MIT Press, Cambridge, MA.
- Chomsky, Noam (2000) “Minimalist Inquiries: The Framework,” Step by Step: In Honor of Howard Lasnik, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 85–155, MIT Press, Cambridge, MA.
- Chomsky, Noam (2001a) “Derivation by Phase,” In Kenstowicz (2001), 1–52.
- Chomsky, Noam (2001b) Beyond Explanatory Adequacy, MIT Occasional Papers in Linguistics 20, MITWPL, Cambridge, MA.
- Deguchi, Masanori, and Yoshihisa Kitagawa (2002) “Prosody and Wh-questions,” NELS 32, Volume 1, 73–92.
- Hirotsu, Masako (2004) Prosody and LF: Processing Japanese Wh-Questions, Doctoral dissertation, University of Massachusetts, Amherst.
- Inkelas, Sharon (1989) Prosodic Constituency in the Lexicon, Doctoral dissertation, Stanford University.
- Ishihara, Shinichiro (2002) “Invisible but Audible Wh-Scope Marking: Wh-Constructions and Deaccenting in Japanese,” WCCFL 21, 180–193.

- Ishihara, Shinichiro (2003) Intonation and Interface Conditions, Doctoral dissertation, MIT.
- Ishihara, Shinichiro (2004) “Prosody by Phase: Evidence from Focus Intonation–Wh-Scope Correspondence in Japanese,” Interdisciplinary Studies on Information Structure 1: Working Papers of the SFB 632, ed. by Shinichiro Ishihara, Michaela Schmitz, and Anne Schwarz, 77–119, Universität Potsdam, Potsdam.
- Ishihara, Shinichiro (2005) “Intonation of Wh-Scrambling: Experiment Overview,” ms., University of Potsdam, [Available at <http://alum.mit.edu/www/s.i/>].
- Kenstowicz, Michael, ed. (2001) Ken Hale: A Life in Language, MIT Press, Cambridge, MA.
- Kiparsky, Paul (1982) “Lexical Phonology and Morphology,” Linguistics in the Morning Calm, ed. by I. S. Yang, 3–91, Hanshin, Seoul.
- Kitagawa, Yoshihisa, and Janet Dean Fodor (2003) “Default Prosody Explains Neglected Syntactic Analysis in Japanese,” Japanese/Korean Linguistics 12, ed. by William McClure, 267–279, CSLI, Stanford.
- Kubo, Tomoyuki (1989) “Fukuoka-shi Hoogen no Dare, Nani Too no Gimon-Shi o Fukumu Bun no Pitti Pataan [The Pitch Pattern of the Sentences Containing Interrogative Markers Such As Dare, Nani, etc.],” Kokugo-gaku 156, 71–82.
- Kubozono, Haruo (1989) “Syntactic and Rhythmic Effects on Downstep in Japanese,” Phonology 6, 39–67.
- Maekawa, Kikuo (1991) “Tokyo Hoogen Gimon-Bun no Intoneesyon [The Intonation of Question Sentences in Tokyo Dialect],” Heisei 3-nendo Nihon Onsei Gakkai Zenkoku Taikai Kenkyuu Happyoo Ronshyuu [Proceedings of the 2001 PSJ Meeting], 42–47, Nihon Onsei Gakkai.

- Maekawa, Kikuo (1997) “Nihongo Gimonsi Gimon-Bun no Intoneesyōn [The Intonation of the Japanese Wh-Question],” Bunpoo to Onsei [Grammar and Sound], ed. by Onsei Bunpoo Kenkyū-kai, 45–53, Kurosio, Tokyo.
- Mahajan, Anoop (1990) The A/A-Bar Distinction and Movement Theory, Doctoral dissertation, MIT.
- Mahajan, Anoop (1994) “Toward a Unified Theory of Scrambling,” Studies on Scrambling: Movement and Non-movement Approach to Free Word-Order Phenomena, ed. by Norbert Corver and Henk van Riemsdijk, 301–330, Mouton de Gruyter, Berlin/New York.
- Miyagawa, Shigeru (1997) “Against Optional Scrambling,” Linguistic Inquiry 28, 1–26.
- Miyagawa, Shigeru (2001) “The EPP, Scrambling, and Wh-in-Situ,” In Kenstowicz (2001), 293–338.
- Nagahara, Hiroyuki (1994) Phonological Phrasing in Japanese, Doctoral dissertation, University of California, Los Angeles.
- Pesetsky, David (1977) “Russian Morphology and Lexical Theory,” ms., MIT.
- Pierrehumbert, Janet, and Mary Beckman (1988) Japanese Tone Structure, MIT Press, Cambridge, MA.
- Saito, Mamoru (1989) “Scrambling as Semantically Vacuous A'-Movement,” Alternative Conceptions of Phrase Structure, ed. by M. Baltin and A. Klock, 192–200, University of Chicago Press.
- Saito, Mamoru (2001) “Toward Reunification of Japanese Scrambling,” MIT Working Papers in Linguistics 41: Proceedings of the 3rd Formal Approaches to Japanese Linguistics (FAJL3), ed. by María Cristina Cuervo, Daniel Harbour, Ken Hiraiwa, and Shinichiro Ishihara, MITWPL, Cambridge, MA.

- Selkirk, Elisabeth (1993) “The Prosodic Structure of Function Words,” Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition, ed. by James L. Morgan and Katherine Demuth, 187–213, Lawrence Erlbaum Associates, Mahwah, NJ.
- Selkirk, Elisabeth (2003) “Bengali Intonation Revisited: An Optimality Theoretic Analysis in Which FOCUS Stress Prominence Drives FOCUS Phrasing,” University of Massachusetts Occasional Papers (UMOP) 26: Papers in Optimality Theory II, ed. by A. Carpenter, A. Coetzee, and P. de Lacy, GSLA, University of Massachusetts, Amherst, Amherst, MA.
- Selkirk, Elisabeth, and Koichi Tateishi (1991) “Syntax and Downstep in Japanese,” In Interdisciplinary Approaches to Language: Essays in Honor of S.-Y. Kuroda, ed. by Carol Georgopoulos and Roberta Ishihara, 519–543, Kluwer, Dordrecht.
- Selkirk, Elisabeth O. (1984) Phonology and Syntax: The Relation Between Sound and Structure, MIT Press, Cambridge, MA.
- Sugahara, Mariko (2003) Downtrends and Post-FOCUS Intonation in Japanese, Doctoral dissertation, University of Massachusetts, Amherst.
- Tada, Hiroaki (1993) A / A-Bar Partition in Derivation, Doctoral dissertation, MIT.
- Truckenbrodt, Hubert (1995) Phonological Phrases: Their Relation to Syntax, Focus, and Prominence, Doctoral dissertation, MIT.
- Zec, Draga, and Sharon Inkelas (1990) “Prosodically Constrained Syntax,” In The Phonology-Syntax Connection, ed. by Sharon Inkelas and Draga Zec, 365–378, CSLI, Stanford.

Universität Potsdam

SFB 632 „Informationsstruktur“

Postfach 601553

14415 Potsdam

Germany

E-mail: s_i@alum.mit.edu

URL: http://alum.mit.edu/www/s_i/

Notes

¹In this paper, we discuss intonation of wh-question in Tokyo dialect. In Fukuoka dialect, wh-question is marked by intonation as in Tokyo dialect, but the realization is different from that of Tokyo dialect (Kubo (1989)).

²Hirotsu (2004) discusses processing of wh-questions in relation to prosody. Based on the results of her production and perception experiments, she claims that the processing principle, which she calls Scope Prosody Correspondence (SPC), does not call for a strict one-to-one mapping between prosody and wh-scope like the one discussed here. See Hirotsu (2004) for detail.

³For expository purpose, I will only use lexically accented words in the examples throughout the paper. The location of lexical pitch accent is marked with ‘’’. See Pierrehumbert and Beckman (1988, Ch. 4) for the difference in focus realization between accented and unaccented phrases in Japanese.

⁴In Selkirk’s (2003) analysis, it is Intonation Phrase boundary that is inserted, although the basic idea remains the same.

⁵How the prosodic phrasing and FI (under the assumption adopted here) interact with each other remains to be studied. I will leave this question for future research.

⁶Such a claim has been already made for study of downstep (Kubozono (1989)).

⁷For the complete set of stimuli, see Ishihara (2005).

⁸Cyclic phonological computation is not a novel idea in itself, e.g. Bresnan (1972) or lexical phonology (Pesetsky (1977), Kiparsky (1982)).

⁹I assume some kind of invisible wh-movement in the case of long-distance Agree relation, to avoid Phase Impenetrability Condition (Chomsky (2000, 2001a,b)) violation.

Our discussion here does not hinge on what kind of invisible wh-movement (e.g. invisible operator movement, feature movement, copy theory) one should assume.

¹⁰I assume that (22a) applies to the phonological content of the phrase with FOC. That is, if the wh-phrase undergoes an invisible movement (see fn. 9), for example, F₀-boosting takes place in its in-situ position.

¹¹An anonymous reviewer pointed out that it would be technically problematic to delete a syntactic FOC-feature after a phonological operation (i.e. application of FI Rules). S/he suggested instead that some kind of conservation principle is involved that prohibits the redefinition of FI that has already been defined at an earlier derivation. I will leave this problem for future research.

¹²An anonymous reviewer pointed out that under the copy theory the basic computation determining which copy to pronounce is also done in a domain bigger than a Spell-Out domain. As long as syntactic computation is done cyclically, the global aspect discussed here should not be a theoretical problem.

¹³The behavioral difference between function words and content words in terms of prosodic phrasing has been often reported in the literature (e.g. Selkirk (1984, 1993), Inkelas (1989), Zec and Inkelas (1990)). Thanks to an anonymous reviewer for pointing this out to me.

¹⁴In the example below, a PP nomíya-de ‘bar-Loc’ and the matrix adverb ímademo ‘even.now’ is deleted to save space. The omission of these words does not affect the current discussion.

¹⁵There is an alternative analysis for Japanese scrambling, in which all scrambled phrases adjoin to TP (cf. Saito (1989, 2001)). Note that this analysis is incompatible with the Multiple Spell-Out analysis proposed here. If the scrambled phrase is in TP, it would be transferred to Φ at Spell-Out, and hence would not be available for further

syntactic movement at the next phase. This means that successive cyclic (i.e. long-distance) scrambling is unavailable universally, unless an additional assumption is made (either that TP-adjoined position is a phase edge, or that there is an additional movement from TP-adjoined position to Spec,CP). Once this additional assumption is made, the prediction becomes the same.

¹⁶This unnaturalness was first acknowledged by Saito (1989). This unnaturalness in judgment is in fact the main point of discussion in Ishihara (2002) and Kitagawa and Fodor (2003).

¹⁷Hirotsu (2004) proposed a processing principle named Scope-Prosody Correspondence (SPC), which is similar to what I suggest here. SPC, however, makes a different prediction from the one entertained here. It in fact predicts that the sentence is correctly interpreted with the intonation in (17b). No misinterpretation would be expected. See Hirotsu (2004) for detail.

¹⁸The alternative account, in which the landing site of scrambling is uniformly TP-adjoined position (cf. Saito (1989, 2001)) is already out of the picture here for the reason explained in fn. 15.