

# Prosody by Phase: Evidence from Focus Intonation–*Wh*-scope Correspondence in Japanese\*

Shinichiro Ishihara

University of Potsdam

Japanese *wh*-questions always exhibit focus intonation (FI). Furthermore, the domain of FI exhibits a correspondence to the *wh*-scope. I propose that this phonology-semantics correspondence is a result of the cyclic computation of FI, which is explained under the notion of *Multiple Spell-Out* in the recent Minimalist framework. The proposed analysis makes two predictions: (1) embedding of an FI into another is possible; (2) (overt) movement of a *wh*-phrase to a phase edge position causes a mismatch between FI and *wh*-scope. Both predictions are tested experimentally, and shown to be borne out.

*Keywords:* Japanese, *wh*-question, prosody, focus intonation, *wh*-scope, cyclicity, phase, *Multiple Spell-Out*

## 1 Introduction

Recently, much attention has been paid to the prosodic properties of *wh*-questions in Japanese and their interaction with syntax and processing (Deguchi and Kitagawa, 2002; Ishihara, 2002; Kitagawa and Tomioka, 2003; Kitagawa and Fodor, 2003; Hirotani, 2003; Ishihara, 2003, among others). It has been claimed that there is a correspondence between the domain of *focus intonation* (henceforth, *FI*)<sup>1</sup> observed in *wh*-questions and the scope of *wh*-questions. It has

\* I would like to thank the following people for their comments, help, suggestions and discussions: Noam Chomsky, Cornelia Endriss, Gisbert Fanselow, Caroline Féry, Danny Fox, Stefan Hinterwimmer, Michael Kenstowicz, David Pesetsky, Donca Steriade, and Hubert Truckenbrodt. Some parts of this paper were presented at GLOW2004 (Thessaloniki, Greece) and at the Workshop on Prosody, Syntax and Information Structure: A Japanese Perspective (WPSI) (Bloomington, IN, USA). I would like to thank the participants of these meetings for their comments and discussion. All the remaining errors are of my own.

<sup>1</sup> Deguchi and Kitagawa (2002) calls it *Emphatic Prosody (EPD)*

also been claimed that this prosody-scope correspondence influences (apparent) syntactic judgments and sentence processing.

In this paper, I will focus on how this prosody-scope correspondence is created. I will claim that prosody is computed cyclically during the course of derivation. Adopting the recent Minimalist framework (Chomsky, 2000, 2001a,b), I propose that cyclic (and hence multiple) application of the so-called *Spell-Out* derives the phonology-semantics correspondence. That is, prosody, the domain of FI in particular, is computed ‘phase-by-phase’.

The proposed model makes two predictions. First, it predicts that the cyclic computation of prosody would allow an embedding of an FI into another. Such a pitch contour has not been reported in the literature of Japanese intonation. In fact, standard analyses of Japanese FI (Pierrehumbert and Beckman, 1988; Nagahara, 1994) would not expect such a contour. Second, when a *wh*-phrase is scrambled out of its *wh*-scope, the Multiple Spell-Out analysis predicts that the prosody-scope correspondence will collapse, and result in a mismatch between the FI domain and *wh*-scope. This prediction contradicts the claims made earlier (Ishihara, 2002; Kitagawa and Fodor, 2003), which take the prosody-scope correlation as a principle that Japanese *wh*-questions always comply to. The Multiple Spell-Out analysis proposed here, on the contrary, *derives* the correspondence as a result of the cyclic computation. Under this analysis, the prosody-scope mismatch is a natural consequence of the overt movement of the *wh*-phrase out of its scope. These two predictions are tested experimentally. As we will see, the results of the experiments further support the proposed model.

This paper is organized as follows. In §2, the Focus Intonation–*Wh*-scope Correspondence will be illustrated with actual examples. Then I will propose the Multiple Spell-Out model of FI creation in §3. §4 introduces the two predictions that the proposed model makes. These two predictions are discussed in §5 and §6, respectively, based on the results of the experiments.

## 2 Focus Intonation–*Wh*-Scope Correspondence (FI=WH)

Japanese *wh*-questions are always accompanied by a focus intonation.<sup>2</sup> Interestingly, the domain of FI exhibits a correspondence to the scope of the *wh*-question, as we will see below. In this section, we will look at some examples showing this phonology-semantics correspondence.

### 2.1 Focus intonation (FI) in Japanese *wh*-question

Maekawa (1991a,b) showed that Japanese (Tokyo dialect) *wh*-questions exhibit FIs. FIs in Japanese can be characterized by two phonetic phenomena:  $F_0$ -boosting on the focalized phrase and the  $F_0$ -lowering of the material following the focalized phrase. We will call these phenomena the *P(rosodic)-focalization* and the *post-FOCUS reduction (PFR)*, respectively.

#### (1) *Focus Intonation (FI) in Japanese*

##### a. **P(rosodic)-focalization**

The  $F_0$  peak of a narrowly focused phrase is raised.

##### b. **Post-FOCUS reduction (PFR)**

The  $F_0$  peaks of the material after the P-focalized phrase is lowered.

A simple illustration of the FI in a *wh*-question is given in (2)<sup>3,4</sup>:

<sup>2</sup> There is one more *wh*-construction in Japanese that exhibits FI, namely, the so-called *Mo*-construction (a.k.a. indeterminate construction) (cf. Kuroda, 1965; Nishigauchi, 1990; Shimoyama, 2001; Hiraiwa, 2002). See fn. 20. See also Ishihara (2003) and Kuroda (2004) for discussion on the prosody of *Mo*-construction.

<sup>3</sup> For expository purpose, I will only use lexically accented words in the examples throughout the paper. The location of lexical pitch accent is marked with ‘’.

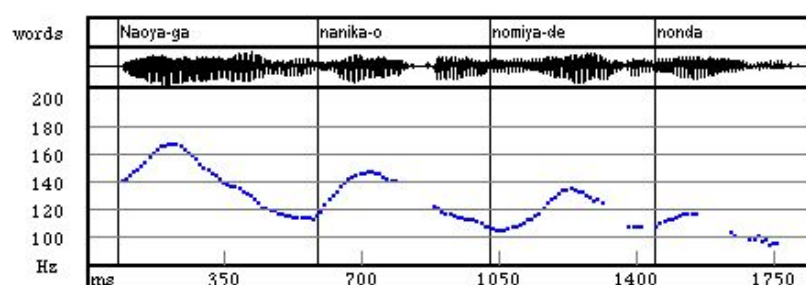
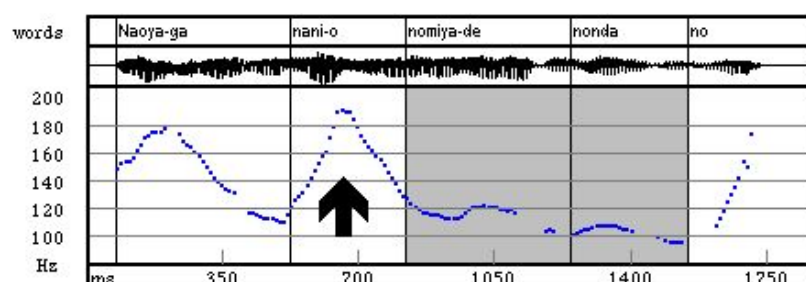
<sup>4</sup> The pitch contours in this examples are recordings of my own voice. All the other pitch contours presented in this paper are obtained from the experiment.

(2) a. *Non-interrogative sentence*

Náoya-ga **nánika**-o nomíya-de nón-da  
 Naoya-NOM something-ACC bar-LOC drank  
 ‘Naoya drank something at the bar.’

b. *Wh-question*

Náoya-ga **náni**-o nomíya-de nón-da **no**?  
 Naoya-NOM what-ACC bar-LOC drank Q  
 ‘What<sub>i</sub> did Naoya drink  $t_i$ ?’

(2') a. *Non-interrogative sentence*b. *Wh-question*

(2a) is a declarative sentence without any narrow/contrastive focus. In this case, the  $F_0$  peaks of the phrases (SUB, OBJ, PP) are all clearly observed.<sup>5,6</sup> On

<sup>5</sup> There appears some downstep-like lowering effect on DO and PP in this pitch contour, since they are clearly lower than their preceding phrases and this lowering effect is too large to attribute to time-dependent declination. This lowering effect, however, is not relevant for our discussion, as long as we can observe the contrasts between the declarative sentence and the *wh*-question.

<sup>6</sup> Generally speaking, the  $F_0$ -peak of the verb is realized much smaller than XPs (DPs/PPs). I will assume that this is due to *downstep* (a.k.a. *Catathesis*), following Selkirk and Tateishi

the other hand, (2b) is a *wh*-question. The *wh*-phrase DO *nani-o* ‘what-ACC’ is clearly realized at a higher pitch than the non-*wh*-counterpart in (2a), since the P-focalization on the *wh*-phrase boosts its  $F_0$  peak. In addition to that, the  $F_0$ -peaks of the post-*wh*-material, i.e., PP *nomiya-de* ‘bar-LOC’ and V *nonda* ‘drank’, are significantly lowered, due to the post-FOCUS reduction.<sup>7</sup>

For the purpose of clarity, I will make one assumption regarding the phonetic nature of P-focalization and PRF, although our main discussion does not hinge on it. Standard analyses of Japanese FI (Pierrehumbert and Beckman, 1988; Nagahara, 1994; Truckenbrodt, 1995, among others) assume that FI is obtained by modifying phonological phrasing, more specifically, by modifying *Major Phrase (MaP)* (a.k.a. *intermediate phrase*) boundaries. A new MaP boundary is created at the focalized phrase while all the MaP boundaries are deleted thereafter. As a result of the restructuring of MaP phrasing, downstep takes place within the newly created large MaP containing the focalized phrase and all the post-FOCUS material. In other words, P-focalization and PFR are captured by the obligatory insertion of a MaP boundary and by downstep, respectively. In this paper, however, I will assume that P-focalization and PFR are pitch-boosting/compression phenomena that are independent of MaP phrasing. This means that I assume that downstep and PFR are different phenomena.<sup>8</sup>

---

(1991). Downstep is a  $F_0$ -lowering phenomenon triggered by H\*L pitch accent within a *Major Phrase*. In principle, verbs always receive downstep effect unless they receive a narrow focus. For a detailed discussion on downstep in Japanese, see, among others, Pierrehumbert and Beckman (1988); Selkirk and Tateishi (1991) and Kubozono (1993).

<sup>7</sup> Since the  $F_0$  peaks on verbs are already reduced by downstep (see fn. 6), the effect of PFR may be very small on verbs. Therefore it may often be the case that the expected contrast due to PFR cannot be clearly observed on the verb (e.g., (3') below). For this reason we will mainly examine the  $F_0$  peaks of non-verbal post-*wh*-phrases.

<sup>8</sup> There are several reasons to take this stance instead of the standard one. Sugahara (2003) shows, for example, that there are cases where MaP boundaries are maintained in the post-focus domain. Even in such cases, however,  $F_0$ -lowering is observed, which suggests that PFR is independent of MaP phrasing. See Ishihara (2003) for a more detailed discussion.

## 2.2 FI–Wh-scope Correspondence (FI=WH)

In addition to this prosodic property of *wh*-questions, Deguchi and Kitagawa (2002) and Ishihara (2002) further showed the following property: When a *wh*-question takes matrix scope, its PFR continues until the end of the matrix clause. When a *wh*-question takes embedded scope, its PFR continues until the end of the embedded clause.<sup>9</sup>

**Matrix *wh*-question** In the case of a matrix *wh*-question like (3), P-focalization boosts the F<sub>0</sub>-peak of the *wh*-phrase, and the PFR compresses the F<sub>0</sub> until the end of the matrix clause, where the question particle *no* appears.

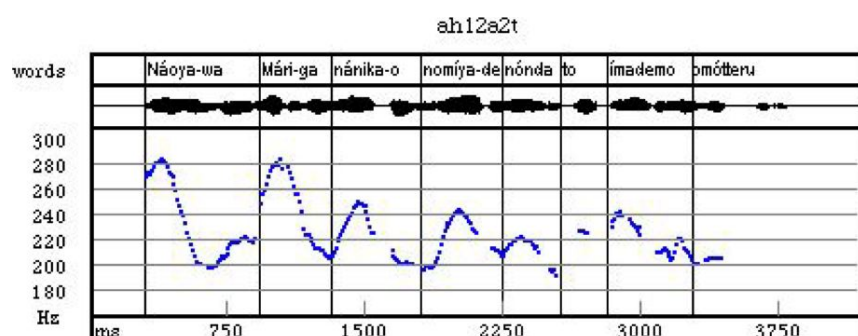
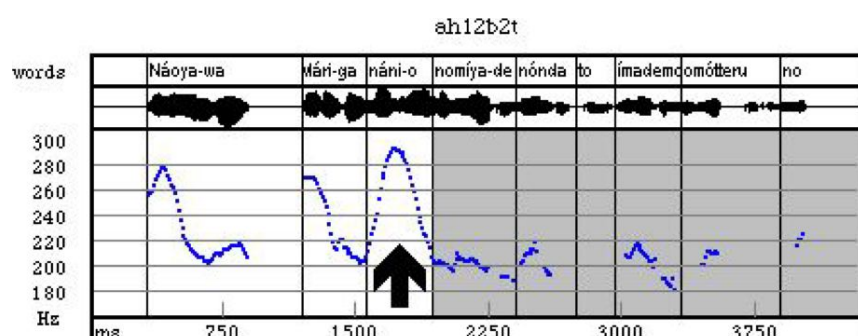
(3) a. *Non-interrogative sentence*

Náoya-wa [ Mári-ga **nánika**-o nomíya-de nónda to ]  
 Naoya-TOP Mari-NOM something-ACC bar-LOC drank that  
 ímademo omótteru  
 even.now think  
 ‘Naoya still thinks that Mari drank something at the bar.’

b. *Wh-question*

Náoya-wa [ Mári-ga **náni**-o nomíya-de nónda to ]  
 Naoya-TOP Mari-NOM what-ACC bar-LOC drank that  
 ímademo omótteru **no**?  
 even.now think Q  
 ‘What<sub>i</sub> did Naoya still think that Mari drank *t<sub>i</sub>* at the bar?’

<sup>9</sup> This property is already reported earlier by Tomioka (1997). Thanks to Masa Deguchi for pointing this out to me.

(3') a. *Non-interrogative sentence*b. *Wh-question*

**Indirect *wh*-question** In the case of the indirect *wh*-question in (4), an FI is again observed, but crucially, in a different manner. P-focalization is observed on the *wh*-phrase, as expected. The PFR, however, does not continue until the end of the matrix clause, but stops at the end of the embedded clause, where the embedded Q-particle *ka* appears. In these cases,  $F_0$  exhibits a *pitch reset* phenomenon after the embedded clause: The post-embedded clause material (e.g. *ímademo* in (4b)) shows clear  $F_0$  peaks.

(4) a. *Indirect Yes/No-question*

Náoya-wa [ Mári-ga **nánika-o** nomíya-de nónda ka ]  
 Naoya-TOP Mari-NOM something-ACC bar-LOC drank Q  
 ímademo obóeteru  
 even.now remember

‘Naoya still remembers whether Mari drank something at the bar.’

b. *Indirect wh-question*

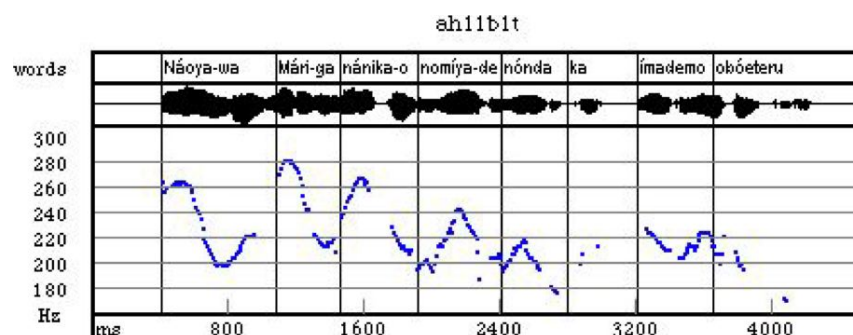
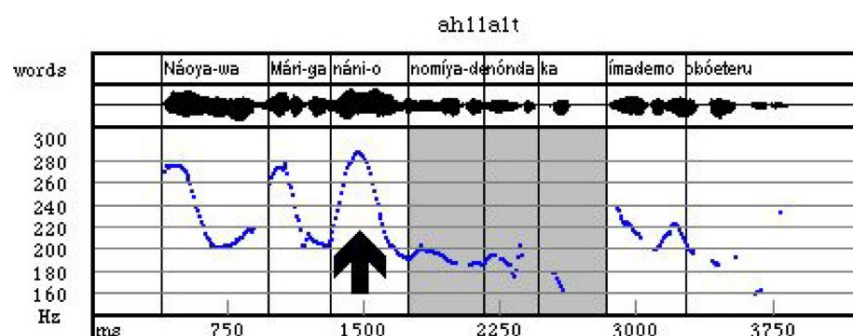
Náoya-wa [ Mári-ga **náni-o** nomíya-de nónda **ka** ]

Naoya-TOP Mari-NOM what-ACC bar-LOC drank Q

ímademo obóeteru

even.now remember

‘Naoya still remembers what<sub>i</sub> Mari drank  $t_i$  at the bar.’

(4') a. *Indirect Yes/No-question*b. *Indirect wh-question*

The facts lead us to the following generalization:

(5) *Focus Intonation–Wh-scope Correspondence (FI=WH)<sup>10,11</sup>*

The domain of FI corresponds to the scope of a *wh*-question.

<sup>10</sup> See Hirotani (2003) for a critical discussion about this generalization.

<sup>11</sup> See also Truckenbrodt (1995, Ch. 4) for a relevant discussion. He claimed that the scope of FOCUS (in the sense of Rooth, 1992) corresponds to the phonological domain at which a focus prominence is assigned.



It should be noted that (5) is just a generalization of the facts we have seen so far. In §6, I will present experimental evidence for the case of *FI–Wh-scope Mismatch*, where the *FI–Wh-scope Correspondence* is no longer observed.<sup>12</sup>

The main goal of this paper is to propose a production model that derives this prosody-semantics correspondence, and to present empirical evidence for this model. Although there are many interesting issues regarding the possible effects of prosody on perception or grammatical judgments,<sup>13</sup> I will concentrate on the issues of production in this paper. In the next section (§3), I will present an analysis that accounts for FI=WH, which is based on the recent Minimalist framework (Chomsky, 2000, 2001a,b).

### 3 A Multiple Spell-Out Account

I propose that FI=WH is a result of the cyclic computation of prosody, which is triggered by the cyclic computation of syntax.<sup>14</sup> This cyclicity in FI creation will be explained in terms of the recent Minimalist framework (Chomsky, 2000, 2001a,b) with the notion of *Multiple Spell-Out*. The syntactic operation *Spell-Out* takes place cyclically at each *phase* in the course of syntactic derivation. My proposal is that prosody, in particular, the domain of FI, is also computed ‘phase-by-phase’. In this section, I will present the mechanism of the model I propose.

#### 3.1 Multiple Spell-Out

*Multiple Spell-Out* is a notion in the recent Minimalist framework proposed by Chomsky (2000, 2001a,b). In this framework, it is proposed that syntactic

<sup>12</sup> As we will discuss later (§4.2 and §6), such a case contradicts the empirical claims made earlier by myself (Ishihara, 2002) and by Kitagawa and Fodor (2003).

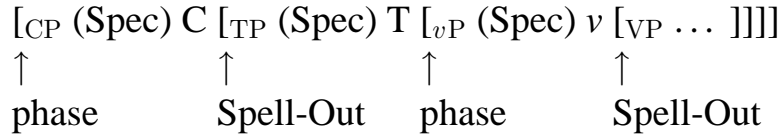
<sup>13</sup> See Deguchi and Kitagawa (2002); Ishihara (2002); Kitagawa and Fodor (2003); Ishihara (2004) for discussion related to perception issues.

<sup>14</sup> The idea of cyclic phonological computation dates back to Bresnan (1972).

computation is done in a cyclic manner. The unit of this cyclic computation is called the *phase*. At each phase, a certain part of the derivation is transferred (via operation *Transfer*) from the narrow syntax *NS* to two interface levels,  $\Phi$  and  $\Sigma$ . The phonological part of *Transfer*, i.e., the operation that transfers the syntactic derivation to the phonological component ( $NS \rightarrow \Phi$ ) is called *Spell-Out*. Since there is more than one phase in a single syntactic derivation, *Spell-Out* takes place more than once in a cyclic manner during the course of derivation, hence ‘Multiple’ *Spell-Out*. The relevant assumptions are listed below.

(6) *Multiple Spell-Out* (Chomsky, 2000, 2001a,b)

- a. CPs and *v*P are *phases*.<sup>15</sup>
- b. When a syntactic derivation reaches a phase (*v*P/CP) in the narrow syntax, the complement of the phase head (i.e., VP/TP) is transferred to the interface levels ( $\Phi/\Sigma$ ). The phonological part of the *Transfer* ( $NS \rightarrow \Phi$ ) is called *Spell-Out*.



### 3.2 Proposal

As mentioned at the beginning of the section, the basic claim of the paper is that “FI is created phase-by-phase.” In this subsection, I present three relevant assumptions of the cyclic FI prosody model I propose.

**FOCUS feature assignment by C** First, we assume that the creation of FI is induced by a FOCUS feature interpreted at the phonological component  $\Phi$ .<sup>16</sup> I

<sup>15</sup> Strictly speaking, only the *v*P of the transitive verb, labeled as *v*\*P, functions as a phase.

<sup>16</sup> I assume that the FOCUS feature is also interpreted at the semantic component  $\Sigma$ . At  $\Sigma$ , it introduces an alternative set for the focus semantic value (Rooth, 1992). We will not discuss the semantics any further in this paper.

propose that this feature is assigned to *wh*-phrases at the syntactic component by the relevant Complementizers, i.e., Q-particles. Therefore, at that point in a syntactic derivation where a *wh*-phrase is merged to the structure, the *wh*-phrase does not carry a FOCUS feature. It will be assigned to a *wh*-phrase when the relevant Q-particle is merged to the derivation.

(7) *FOCUS feature assignment by C*



**Timing of FI creation** The FOCUS feature assigned to a *wh*-phrase is interpreted at  $\Phi$  as soon as it enters into  $\Phi$  via Spell-Out operation. The FOCUS feature induces P-focalization on the FOCUS phrase and PFR thereafter. Since the Complementizer assigns the FOCUS feature to *wh*-phrases, it is not until C is introduced to the syntactic derivation and a CP phase is formed that the FI creation is induced at  $\Phi$ .

For example, let us look at the matrix *wh*-question sentence (8), which contains the *wh*-phrase *nani-o* as its object.

- (8) [CP [TP Táro-wa [<sub>v</sub>P [VP **náni-o** nónda ] *v* ] T ] **no** ]  
           Taro-TOP           what-ACC drank           Q  
           ‘What did Taro drink?’

When the *v*P phase is created, its Spell-Out domain (VP) contains the *wh*-phrase, but the *wh*-phrase is not yet assigned a FOCUS feature. Thus the FI is not yet created at the Spell-Out of this phase, as in (9a). At the CP phase, the Q-particle *no* is merged to the derivation and assigns a FOCUS feature to the *wh*-phrase. The Spell-Out domain (TP) now contains a FOCUS feature, as in (9b). Hence the FI is created at this Spell-Out cycle.

- (9) a. *vP phase: No FI created*

[<sub>vP</sub> [<sub>VP</sub> **náni**-o nónda ] *v* ]

↑

No FOCUS feature assigned

- b. *CP phase: FI created*

[<sub>CP</sub> [<sub>TP</sub> Táro-wa [<sub>vP</sub> [<sub>VP</sub> **náni**<sub>FOC</sub>-o nónda ] *v* ] T ] **no** ]

↑

FOCUS feature assigned by C

**FOCUS feature deletion** Lastly, we assume that the FOCUS feature is deleted after the FI is created. This means, once the FOCUS feature is used to create an FI at some Spell-Out cycle, it will not affect prosody created at any later Spell-Out cycle. Let us see how the model works with some examples.

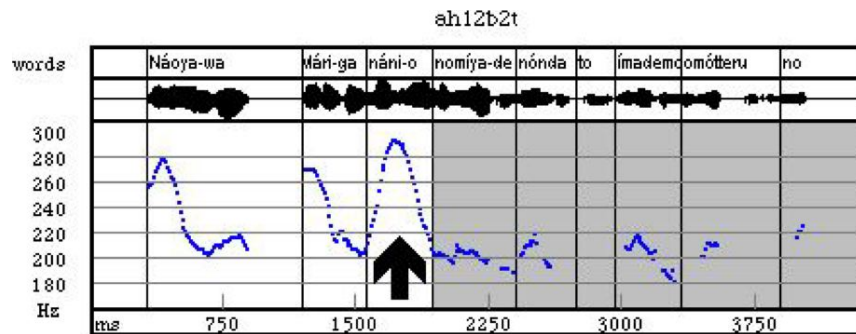
### 3.3 Examples

The proposed analysis nicely explains the difference in FI realization between the matrix *wh*-question (3b) and the indirect *wh*-question (4b), repeated below. It predicts that the FIs of these two sentences are created at different Spell-Out domains: In the former case, the FI is created at the Spell-Out domain of the matrix CP phase, while in the latter, it is created at the Spell-Out domain of the embedded CP phase. Let us take a closer look at how their FIs are derived.

- (3b) *Matrix wh-question: FI created at the matrix CP phase*

Náoya-wa [ Mári-ga **náni**-o nomíya-de nónda to ] ímademo  
 Naoya-TOP Mari-NOM what-ACC bar-LOC drank that even.now  
 omóteru **no**?  
 think Q

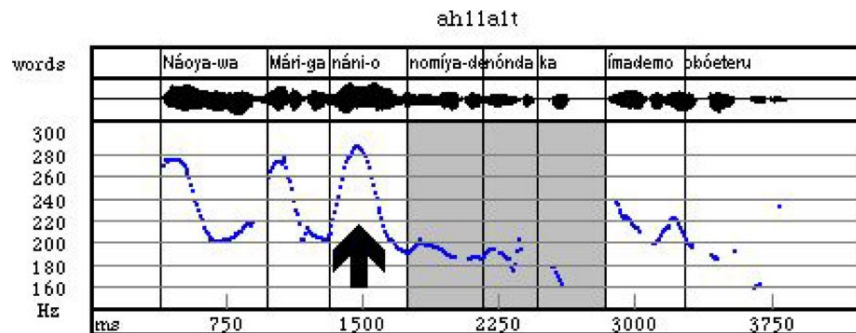
‘What<sub>i</sub> did Naoya still think that Mari drank *t<sub>i</sub>* at the bar?’



(4b) *Indirect wh-question: FI created at the embedded CP phase*

Náoya-wa [ Mári-ga **náni-o** nomíya-de nón<sup>da</sup> **ka** ] ímademo  
 Naoya-TOP Mari-NOM what-ACC bar-LOC drank Q even.now  
 obóeteru  
 remember

‘Naoya still remembers what<sub>i</sub> Mari drank *t<sub>i</sub>* at the bar.’



**Matrix *wh*-question** In the case of matrix *wh*-questions, the *wh*-phrase is P-focalized, and the PFR after the *wh*-phrase continues until the end of the sentence. (In the examples hereafter, P-focalization is indicated by box, and PFR by underline.)

(10) *Matrix wh-question*

[<sub>CP</sub> ...  $\alpha$  ... [<sub>CP</sub> ...  $\beta$  ... WH ...  $\gamma$  ... ] ...  $\delta$  ... Q]

At the embedded CP phase (11a)<sup>17</sup>, the *wh*-phrase is not yet assigned a FOCUS feature, since the Q-particle is not yet merged to the derivation. Therefore its Spell-Out domain, the embedded TP, does not contain any FOCUS feature. Since there is no FOCUS feature, no FI is created at  $\Phi$  at this point of the derivation, as in (11b).

(11) a. *Embedded CP phase*

[<sub>CP</sub> [<sub>TP</sub> ...  $\beta$  ... WH ...  $\gamma$  ... ] ] (No FOCUS assignment)

b. *Output at  $\Phi$*

[<sub>TP</sub> ...  $\beta$  ... WH ...  $\gamma$  ... ] (No FI)

The derivation continues to the matrix CP phase. A Q-particle is merged as the matrix C, and assigns a FOCUS feature to the *wh*-phrase, as in (12a). As a result, the Spell-Out domain, the matrix TP, now contains a FOCUS feature. Accordingly, an FI is created at  $\Phi$ : The *wh*-phrase is P-focalized, and the PFR applies to the all the post-*wh*-phrases, as in (12b). Since the FI is created at the matrix CP phase, its Spell-Out domain, i.e., the matrix TP, serves as the domain of FI. This means that the PFR domain contains the post-*wh*-phrase material in the embedded CP ( $\gamma$ ) as well as the one in the matrix CP ( $\delta$ ).

(12) a. *Matrix CP phase*

[<sub>CP</sub> [<sub>TP</sub> ...  $\alpha$  ... [<sub>CP</sub> ...  $\beta$  ... WH<sub>FOC</sub> ...  $\gamma$  ... ] ...  $\delta$  ... ] Q ]

↑  
FOCUS assignment

b. *Output at  $\Phi$*

[<sub>TP</sub> ...  $\alpha$  ... [<sub>CP</sub> ...  $\beta$  ... WH ...  $\gamma$  ... ] ...  $\delta$  ... ]

FI creation

At this point, there are a few more elements that have not been transferred to  $\Phi$ , namely, phrases in the Spec,CP (if any), and the phase head, i.e., Q-particle.

<sup>17</sup> Although I will omit *vP* phases for brevity, the explanation presented here for the CP phase not headed by a Q-particle holds for *vP* phases as well. See also fn. 20 for discussion of the *Mo*-construction, in which *vP* phase seems relevant.

I will assume that there is another Spell-Out operation that applies to the root of the derivation, which I will call the *root Spell-Out*. The Spec,CP and the phase head C are transferred to  $\Phi$  at the root Spell-Out.

Since Q-particles are phase heads and appear outside the Spell-Out domain (i.e., TP), the proposed analysis would predict that they are not to be inside the domain of PFR. In reality, however, these particles seem to be within the PFR domain.<sup>18</sup> I suggest that this is because they do not behave as Prosodic Words by themselves and have no ability to create a new prosodic boundary at any level (Minor Phrase, Major Phrase, or Intonation Phrase). Hence, they are always integrated into the prosodic phrase of the preceding phrase (i.e., verbal complex). Their  $F_0$  is therefore always dependent on that of the preceding phrase.

**Indirect *wh*-question** In the case of indirect *wh*-questions like (4b), FI is only observed within the embedded CP. After the embedded CP, a pitch reset is observed. The matrix material after the embedded CP ( $\delta$ ) is outside the FI domain.

(13) *Indirect wh-question*

[<sub>CP</sub> ...  $\alpha$  ... [<sub>CP</sub> ...  $\beta$  ... WH ...  $\gamma$  ... Q ] ...  $\delta$  ... ]

↑  
Pitch reset

At the embedded CP phase (14a), the Q-particle assigns FOCUS to the *wh*-phrase. When Spell-Out applies to the derivation, the sister of the Q-particle, i.e., TP, is transferred to  $\Phi$ . Since this Spell-Out domain contains a FOCUS

<sup>18</sup> In the case of the matrix Q-particles like the one in (12a), a question-final rising intonation is normally observed on the Q-particle. Therefore it looks as if they were outside the PFR domain. This rising intonation, however, is not a property of the Q-particle itself, but rather a utterance-final boundary tone that is realized on the final mora of the utterance. If a non-monomoraic Q-particle *ndai* (cf. Yoshida, 1998) is used, for example, the rising intonation is realized on the last mora of this particle, instead of the beginning of this particle. Even if the Q-particle is omitted (cf. Yoshida and Yoshida, 1996), the rising intonation is still observed on the last mora of the verbal complex. See also fn. 19 about the Q-particle in the embedded clause.

feature, an FI is created: P-focalization on the *wh*-phrase followed by the PFR of the post-FOCUS material ( $\gamma$ ), as in (14b).<sup>19</sup>

(14) a. *Embedded CP phase*

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

↑  
FOCUS assignment

b. *Output at  $\Phi$*

$$[_{TP} \dots \beta \dots \boxed{\text{WH}} \dots \gamma \dots ]$$

FI creation

Note that the FOCUS feature is deleted after the FI is created. At the matrix CP phase, therefore, no more FI is created, as in (15). Since the FI is created at the earlier Spell-Out cycle, it does not affect the material introduced at the matrix cycle ( $\alpha$ ,  $\delta$ ). Accordingly, a pitch reset is observed after the embedded CP.

(15) a. *Matrix CP phase*

$$[_{CP} [_{TP} \dots \alpha \dots [_{CP} \dots \beta \dots \boxed{\text{WH}} \dots \gamma \dots Q ] \dots \delta \dots ] ]$$

(No more FOCUS assignment)

b. *Output at  $\Phi$*

$$[_{TP} \dots \alpha \dots [_{CP} \dots \beta \dots \boxed{\text{WH}} \dots \gamma \dots Q ] \dots \delta \dots ]$$

↑  
Pitch reset

In sum, the FI for a *wh*-phrase is created at the phase whose head is the Q-particle that binds the *wh*-phrase. When the Q-particle is the matrix C (i.e.,

<sup>19</sup> In this case again, the Q-particle, which is outside of the Spell-Out domain of the embedded CP phase, appears to be contained in the PFR domain. In my experimental data, there were cases where a sharp  $F_0$  rise is observed on Q-particles, which could potentially be analyzed as a beginning of a new phonological phrase. My impression was, however, that the occurrence of this rise were inconsistent enough to conclude that Q-particles always start a new phonological phrase. Therefore I will assume here that this rise is some sort of boundary tone at the end of the PFR domain. I will leave the investigation of this rise for future research.



when the sentence is a matrix *wh*-question), the FI is created at Spell-Out of the matrix CP phase (i.e., the matrix TP). When the Q-particle is the embedded C (i.e., when the sentence is an embedded *wh*-question), the FI is created at the Spell-Out of the embedded CP (i.e., the embedded TP). Accordingly, the domain of FI corresponds to the scope of the *wh*-question. FI=WH is a result of the cyclic computation of FI.

It should be noted that there is no direct interaction between the phonological and the semantic component during this process. FI=WH is not a result of the direct interaction between phonology and semantics. It is rather the result of the cyclic syntactic computation. One advantage of this model is that the phonological process is as simple as possible. The phonological component only looks for a FOCUS feature each time a new syntactic material is transferred via Spell-Out. When it finds one, it immediately creates an FI. The phonological component is completely indifferent to the semantic scope. Note that the phonetic rules to create an FI are also simple: boosting the  $F_0$  peak of the phrase bearing a FOCUS feature, and lowering everything thereafter. It does not involve specifying where PFR ends. The end point of PFR is automatically derived, since PFR only applies to a relevant Spell-Out domain, not to the whole sentence.<sup>20</sup>

<sup>20</sup> One might wonder if there is a case in which an FI is created at a *vP* phase. In the so-called *Mo*-construction (Shimoyama, 2001) (the *indeterminate construction* of Kuroda, 1965), FIs can be found between *wh*-phrase and the particle *mo*, which may appear after C, Verb, or Case-markers. (i) is an example where *mo* attaches to *vP*.

(i) *Mo-construction*

Máři-wa    [<sub>vP</sub> **náni**-o      nomíya-de    nómi ]-**mo**    si-**nákat**-ta  
 Mari-TOP    what-ACC    bar-LOC      drink -MO    do-NEG-PST  
 ‘For no *x*, Mari drink *x* at the bar.’

This suggests that *vP* and *DP* are also phases and an FI can be created at their Spell-Out.

## 4 Two Predictions

The Multiple Spell-Out account proposed in the previous section derives FI=WH as a consequence of cyclic computation in syntax, namely, Multiple Spell-Out. Because of this cyclic property, the proposed analysis makes two interesting predictions. These two predictions are not expected in other possible analyses for FI and for FI=WH phenomena in Japanese. These two predictions are experimentally tested. As we will see, the results strongly support the Multiple Spell-Out analysis.

### 4.1 Prediction 1: FI embedding

FIs are created cyclically under the Multiple Spell-Out model, it would be possible for a single derivation to create two FIs at different Spell-Out domains. We can therefore make the following prediction:

(16) *FI embedding*

When there are two independent WH-Q dependencies with different scopes, an FI is embedded into another.

$$[ \text{WH1} \dots [ \dots \text{WH2} \dots \alpha \dots \text{Q}_{emb} ] \dots \beta \dots \text{Q}_{mat} ]$$

The diagram shows a nested structure of brackets. The outermost bracket starts at 'WH1' and ends at 'Q\_mat'. Inside this, there is a bracket starting at 'WH2' and ending at 'Q\_emb'. A third bracket starts at 'α' and ends at 'Q\_emb'. A long horizontal line with a vertical drop at the end is positioned below the 'Q\_emb' bracket, extending from the level of 'WH1' to the level of 'Q\_mat'.

The resulted contour would realize an FI at the matrix CP (between WH1 and  $\text{Q}_{mat}$ ) which contains ‘residues’ of another FI that are created at the embedded CP (between WH2 and  $\text{Q}_{emb}$ ). WH2 would be first P-focalized at the embedded CP phase, and then reduced by PFR at the matrix cycle induced by WH1. Also the post-WH material  $\alpha$  would exhibit the PFR effects of both FIs, while the post-embedded CP material  $\beta$  would only show the PFR effect of the matrix FI.

Such a pitch contour has never been reported for Japanese, at least to my knowledge. If such a contour is in fact observed, standard analyses of Japanese

FI (Nagahara, 1994; Truckenbrodt, 1995; Selkirk, 2000; Sugahara, 2003, among others) would require some modifications. As mentioned earlier (§2.1), they assume that FI is obtained by restructuring MaP phrasing. The FI embedding would then be analyzed as an embedding of a MaP into another. Such a prosodic phrasing structure would violate the Non-recursivity of the Strict Layer Hypothesis (Selkirk, 1984; Nespor and Vogel, 1986).<sup>21</sup> Also, Selkirk's (2003) claim that a (contrastive) focus is always associated with prominence at the Intonation Phrase (IP) level would not hold in the FI embedding case, because the realization of the matrix focus (WH1 in (16)) and that of the embedded focus (WH2) are expected to be different: The embedded focus would have a more compressed realization than the matrix focus. In §5, I present and discuss the result of the experiment conducted to test this prediction. In the next subsection, we consider the second prediction.

#### 4.2 Prediction 2: FI–WH Mismatch (FI≠WH) due to movement

The second prediction of the Multiple Spell-Out analysis is related to syntactic movement. So far, we have only seen examples where the *wh*-phrases stay in-situ. In all these examples, we observed FI=WH. Once the *wh*-phrase overtly moves outside the Spell-Out domain via so-called 'edge' position of phases (i.e., the specifier of the phase head), however, the Multiple Spell-Out model expects a different FI than what we have seen so far.

If a *wh*-phrase moves out of the *wh*-scope phase, by moving to the 'edge' positions in a successive cyclic manner, it will be excluded from the Spell-Out domain of each phase. As a result, the creation of an FI will be postponed to

<sup>21</sup> There have been, however, cases reported in the literature that violate Non-recursivity (Selkirk, 1993; Truckenbrodt, 1995). Therefore if FI embedding is in fact the case, it could serve as evidence for MaP embedding. See Kubozono (2004) for the recursive structure of MaP in Japanese downstep. See also Féry and Truckenbrodt (2003); Truckenbrodt and Féry (2003) for a recursive model of downstep for German.

a later Spell-Out cycle than the Spell-Out of the phase where the *wh*-scope is fixed. As a result, the FI domain becomes larger than the actual *wh*-scope.

(17) *FI–Wh-scope Mismatch (FI ≠ WH)*

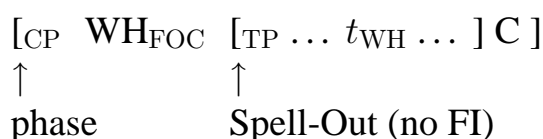
Once the *wh*-phrase bearing a FOCUS feature is moved out of its *wh*-scope via phase ‘edge’ positions, the FI will be created at the later Spell-Out cycle. As a result, *FI–Wh-scope Mismatch (FI ≠ WH)* will arise.

This prediction is drawn from the following theoretical assumptions:

- (18) a. The landing site of  $\bar{A}$ -scrambling (including all instances of long-distance scrambling) is Spec,CP (Mahajan, 1994).  
 b. Spec,CP is the phase ‘edge’ position, which is *outside* the Spell-Out domain of this CP phase.

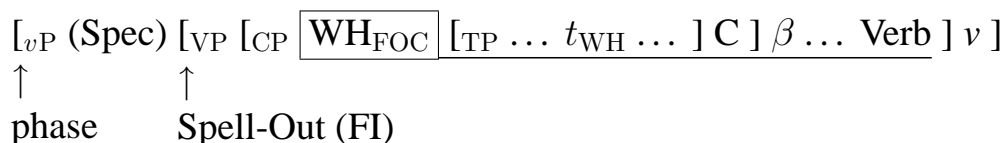
This means that any *wh*-phrase scrambled to a Spec,CP will be excluded from the Spell-Out domain of this CP phase, as in (19).

(19) *Embedded CP phase*



The FOCUS feature of the scrambled *wh*-phrase, then, will be carried to the next phase, i.e., the *vP* phase. As a result, the FI will be created at the Spell-Out of the *vP* phase, namely, VP, which includes not only the embedded clause but also post-embedded-CP phrases (i.e.,  $\beta$  in (20)) and the verb.

(20) *Matrix vP phase*



If the *wh*-phrase further moves to a higher position (e.g., Spec,*v*P in (20)) via successive cyclic movement, the FI creation will be delayed further. Semantically, however, the *wh*-phrase is interpreted in-situ, due to the radical reconstruction effect of long-distance scrambling (Saito, 1989). As a result, the domain of FI and the scope of the *wh*-question no longer exhibit a correspondence.

A relevant case can be found in the literature. The example we will examine here is from Saito (1989), in which he showed the radical reconstruction property of long-distance scrambling. In (21), the *wh*-phrase has an embedded scope, regardless of whether the *wh*-phrase is in situ as in (21a), or it is long-distance scrambled to the beginning of the matrix clause as in (21b).

(21) *Saito's (1989) example: Long-distance-scrambled wh-phrase*

- a. [ Mary-ga [ John-ga **dono** hon-o tosyokan-kara  
 Mary-NOM John-NOM which book-ACC library-from  
 karidasita **ka** ] siritagatteiru ] koto  
 checked.out Q want.to.know fact  
 'The fact that Mary wants to know [which book]<sub>i</sub> John checked  
 out *t*<sub>i</sub> from the library.'
- b. ? [ **dono** hon<sub>i</sub>-o [ Mary-ga [ John-ga *t*<sub>i</sub> tosyokan-kara  
 which book-ACC Mary-NOM John-NOM library-from  
 karidasita **ka** ] siritagatteiru ] ] koto  
 checked.out Q want.to.know fact

(Saito, 1989, p. 191–192, ex. (34))

We already saw in (4) that the embedded *wh*-question like (21a) exhibits an FI in the embedded clause, between the in-situ *wh*-phrase and the embedded Q-particle. Now the question is how sentences like (21b) would be pronounced.

If one assumes a non-cyclic model to explain FI=WH, one could generalize FI=WH by stipulating that an FI starts from the *wh*-phrase and ends at the Q-particle that binds the *wh*-phrase. (This was in fact the generalization I made in Ishihara, 2002. See also Kitagawa and Fodor, 2003 for the same claim.) Under such an observation, the expected contour for (21b) would show an FI from

the scrambled *wh*-phrase until the embedded Q-particle *ka*, and a pitch reset thereafter, as illustrated in (22a).

On the other hand, the Multiple Spell-Out model proposed here would predict that the FI is created at the root Spell-Out instead of the embedded CP phase, even though the scope of the *wh*-question is still the embedded CP, due to the radical reconstruction, as in (22b). As a result, we would no longer expect FI=WH. We would rather expect a mismatch between the phonological domain of FI and the *wh*-scope.

(22) a. *F<sub>0</sub> contour predicted by the generalization in Ishihara (2002)*

[<sub>CP</sub> WH [<sub>TP</sub> α ... [<sub>CP</sub> [<sub>TP</sub> ... *t<sub>WH</sub>* ... ] **ka** ] β ... ]

↑  
Pitch reset

b. *F<sub>0</sub> contour predicted by the multiple Spell-Out model*

[<sub>CP</sub> WH [<sub>TP</sub> α ... [<sub>CP</sub> [<sub>TP</sub> ... *t<sub>WH</sub>* ... ] **ka** ] β ... ]

↑  
No pitch reset

If this prediction is borne out, it would pose a challenge to any model assuming direct phonology-semantics interaction to account for FI=WH, because FI=WH no longer holds once the *wh*-phrase is scrambled out of the scope of the Q-particle binding it. If a direct phonology-semantics interaction is assumed to account for FI=WH, such a mismatch would not be expected.

In this section, we discussed the two prediction made by the Multiple Spell-Out model. These two predictions were experimentally tested. In the next two sections (§5, §6), the results of the experiments will be presented.

## 5 Experiment 1: FI Embedding

Let us examine the first prediction, i.e., FI embedding. In this section, I present the result of an experiment, and claim that FI embedding is in fact attested.

## 5.1 Method

The experiment was conducted using five subjects (four females, AH, CS, CK, NM, and a male, YY), who are all non-linguists brought up in Tokyo or surrounding areas. Stimuli consisting of 32 target sentences (see below for detail) mixed with 104 filler sentences are 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. 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. Each subject makes 3 recordings of the entire set of stimuli. Each recording uses a different pseudo-randomized order of the stimuli sentences.

## 5.2 Stimuli

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

(23) *4 sentence types to be examined*

A. **non-WH/WH:** *Indirect wh-question*

Náoya-wa [ Mári-ga **náni**-o nomíya-de nónda **ka** ]  
 Naoya-TOP Mari-NOM what-ACC bar-LOC drank Q  
 ímademo obóeteru  
 even.now remember

‘Naoya still remembers what<sub>i</sub> Mari drank *t<sub>i</sub>* at the bar.’

B. **non-WH/non-WH:** *Indirect Yes/No-question*

Náoya-wa [ Mári-ga **nánika**-o nomíya-de nónda **ka** ]  
 Naoya-TOP Mari-NOM something-ACC bar-LOC drank Q  
 ímademo obóeteru  
 even.now remember

‘Naoya still remembers whether Mari drank something at the bar.’

C. **WH/WH:** *Wh-question with an indirect wh-question*

**dáre**-ga [ **Má**ri-ga **náni**-o nomíya-de nónda **ka** ]  
 who-NOM Mari-NOM what-ACC bar-LOC drank Q  
 ímademo obóeteru **no**?  
 even.now remember Q

‘Who still remembers what<sub>i</sub> Mari drank *t<sub>i</sub>* at the bar?’

D. **WH/non-WH:** *Wh-question with an indirect Yes/No-question*

**dáre**-ga [ **Má**ri-ga **nánika**-o nomíya-de nónda **ka** ]  
 who-NOM Mari-NOM something-ACC bar-LOC drank Q  
 ímademo obóeteru **no**?  
 even.now remember Q

‘Who still remembers whether Mari drank something at the bar?’

(23C) is the FI embedding sentence, which contains one *wh*-phrase in the matrix clause (taking the matrix scope), and another *wh*-phrase in the embedded clause (taking the embedded scope). This sentence is compared with (23D), where the embedded *wh*-phrase is replaced by a non-*wh*-phrase. (23D) would only show an FI at the matrix clause. If FI embedding is possible at all, (23C) would show FI effects at the embedded clause, even though the entire embedded clause is compressed by the PFR of the matrix FI. (23A) and (23B), in which the matrix *wh*-phrase is replaced by a non-*wh*-phrases, are compared with (23C) and (23D), respectively, to make sure that the matrix FI effects are observed in (23C) and (23D).

Among the  $F_0$  peaks in the sentences, those of the following five phrases are measured to examine the FI effects. They are labeled as P1, P2, ... P5, respectively.

(24) *Labels of the relevant  $F_0$  peaks*

[ (Non-)WH	[ ... (Non-)WH ...	$\alpha$ ...	Verb $C_{[+Q]}$ ]	$\beta$ ...	$C_{[\pm Q]}$ ]
P1	P2	P3	P4	P5	



**P1: Matrix *wh*/non-*wh*-phrase** P1 indicates the P-focalization effect at the matrix CP cycle. (If P1 is a *wh*-phrase, it is P-focalized.)

**P2: Embedded *wh*/non-*wh*-phrase** P2 indicates the P-focalization effect at the embedded CP cycle. (If P2 is a *wh*-phrase, it is P-focalized.) It also indicates the PFR effect at the matrix CP cycle. (If P1 is a *wh*-phrase, P2 is lowered by PFR.)

**P3: Phrase immediately following P2** P3 shows the PFR effects of both the embedded and the matrix CP cycle. (If P1 and/or P2 are *wh*-phrases, P3 is lowered by PFR.)

**P4: Embedded clause verb** P4 is not directly relevant to the test.<sup>22</sup> However, since it is the last and the lowest  $F_0$  peak in the embedded clause, it helps us see more clearly the effect of pitch reset expected on P5.

**P5: Material immediately following the embedded clause** P5 indicates the PFR effect at the matrix CP cycle, but not the PFR effect at the embedded CP cycle. In other words, P5 indicates the amount of pitch reset after the embedded clause. (If P1 is a *wh*-phrase, P5 is lowered by PFR. If P2 is a *wh*-phrase, P3 and P4 are lowered by PFR, but P5 is not.)

### 5.3 Predictions

The stimulus set is schematically illustrated in (25). Also, all the expected contrasts are depicted in a graph in (26).

<sup>22</sup> Because all the effects expected on this peak are exactly the same as those of P3.

(25) *Stimulus set (with predicted P-focalization and PFR)*A. *non-WH/WH*

[ P1<sub>[-WH]</sub> [ ... P2<sub>[+WH]</sub> ... P3 ... P4 C<sub>[+Q]</sub> ] P5 ... C<sub>[-Q]</sub> ]

B. *non-WH/non-WH*

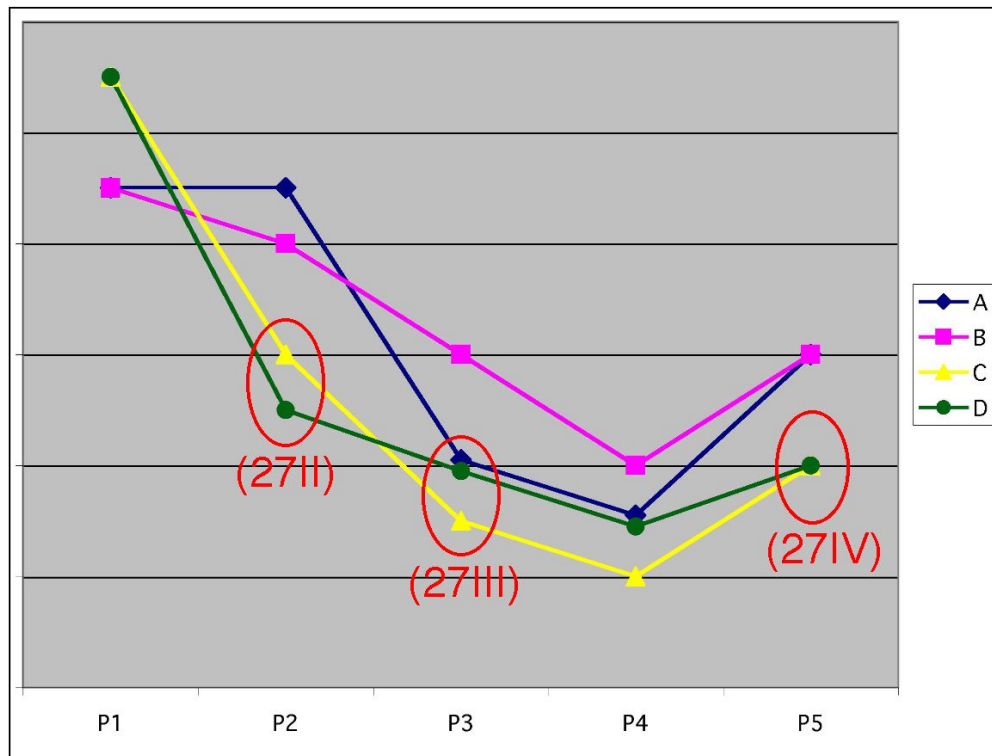
[ P1<sub>[-WH]</sub> [ ... P2<sub>[-WH]</sub> ... P3 ... P4 C<sub>[+Q]</sub> ] P5 ... C<sub>[-Q]</sub> ]

C. *WH/WH*

[ P1<sub>[+WH]</sub> [ ... P2<sub>[+WH]</sub> ... P3 ... P4 C<sub>[+Q]</sub> ] P5 ... C<sub>[+Q]</sub> ]

D. *WH/non-WH*

[ P1<sub>[+WH]</sub> [ ... P2<sub>[-WH]</sub> ... P3 ... P4 C<sub>[+Q]</sub> ] P5 ... C<sub>[+Q]</sub> ]

(26) *Prediction (NB: not an actual result)*<sup>23</sup>

As mentioned above, the crucial contrasts to be examined is those between C and D, especially, regarding P2, P3, and P5. First of all, in C and D, all these

<sup>23</sup> This graph simply illustrates the expected contrasts in terms of relative height at each peak among the sentence types. No quantitative predictions are made.

peaks are expected to be lowered by the PFR after the matrix *wh*-phrase (P1). It is therefore expected that P2, P3, and P5 are lower in C and D than in B (Prediction (27I)).

Inside this PFR domain, we would expect the difference on P2 and P3 between C and D. P2 is expected to be higher in C due to the P-focalization of the *wh*-phrase (Prediction (27II)); and P3 is expected to be lower in C due to the PFR induced by this *wh*-phrase (Prediction (27III)).

P5, however, is expected to show no difference between the two sentence types, since the embedded FI would not affect this peak (Prediction (27IV)).

(27) *Crucial predictions*

I. **P2–5: B > C, D**

In both C and D, P1 is P-focalized, and P2–P5 are lowered by PFR. Therefore, P2–P5 in C and D are expected to be lower than those in B, where no PFR takes place.

II. **P2: C > D**

P2 in C is P-focalized at the embedded CP cycle, while it is not in D. Accordingly, C is expected to be higher than D.

III. **P3: C < D**

P3 in C is lowered by PFR at the embedded CP cycle, while it is not in D. Accordingly, C is expected to be lower than D.

IV. **P5: C = D**

P5 in C and D are expected to reach the same height, due to the pitch reset after the embedded CP cycle in D.

## 5.4 Results and discussion

The results are first analyzed for each subject. Then the data from four of the five subjects (excluding MN's data<sup>24</sup>) are normalized to see if the embedded FI can be observed as a general property among these speakers.<sup>25</sup>

### (28) *Data normalization*

- a. Each subject's data is normalized according to the following formula:

$$y = \frac{x - R_2}{R_1 - R_2}$$

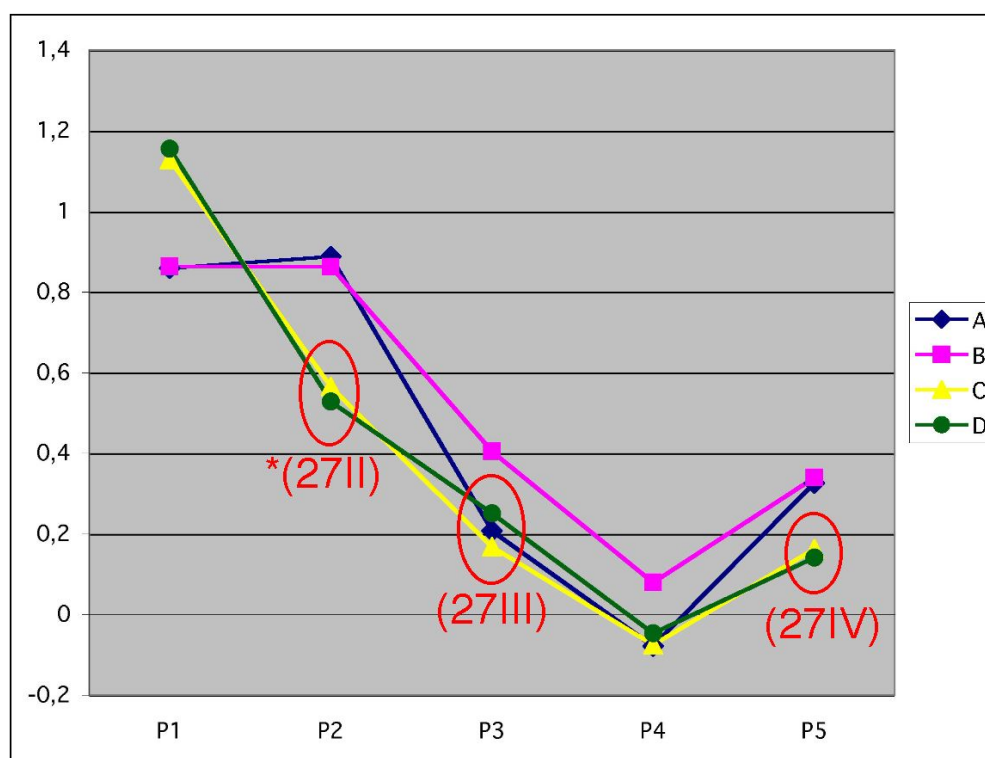
where  $R_1$  and  $R_2$  are the reference points calculated independently for each subject.

- b. The following two values are chosen as the reference points ( $R_1, R_2$ ) for the normalization:
  - $R_1$  = Mean value of P1 ( $F_0$ -peak on the 1st (non)-*wh*-phrase)
  - $R_2$  = Mean value of P4 ( $F_0$ -peak of the embedded verb)

The normalized results are shown in the graph below.

<sup>24</sup> 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 syntax/semantics of the sentences, and read them mere as sequences of words. Such data would not tell us anything important for our purpose.

<sup>25</sup> In this paper, I will only present the normalized data due to space limitations. For the results of the individual subjects and detailed analyses of them, see Ishihara (2003).

(29) *Normalized Result*

First of all, it is clear from (29) that P2–P5 of C and D are much lower than that of B (i.e., Prediction (27I)). In fact, the contrasts are all statistically significant ( $p < 0.00001$  at all relevant peaks).

Given that P2–P5 are all lowered, we can now compare between C and D to verify the rest of the predictions in (27). The t-test results are shown below:

(30) *Mean differences between C and D*

Peak	p	Statistically ...	Relevant prediction
P2	= 0.306	Not significant	*(27II) C > D
P3	< 0.0001	Significant	(27III) C < D
P5	= 0.231	Not significant	(27IV) C = D

As shown above, (27III) and (27IV) are supported by the data. There is a statistically significant contrast on P3, showing that P3 is lower in C than in D

(i.e., (27III)). This means that even though P3 is lowered both in C and D by the matrix PFR effect, there is also an embedded PFR effect only in C. This embedded PFR effect is further proved by the fact that P5 in C and D reaches the same point, indicating that there was a pitch reset after the embedded PFR in C. Since the embedded PFR effect in C is limited to the embedded CP, P5, which belongs to the matrix clause, is not affected by this effect. As a result, P5 in C is only affected by the matrix PFR effect, just like in D.

The contrast on P2, however, is not statistically significant. This fact by itself may appear to indicate that there is no embedded P-focalization effect. This lack of expected contrast on P2, however, seems due to the experimental design. As the non-*wh*-counterparts for this position, indefinite pronouns such as *nanika* ‘something’ and *dareka* ‘someone’ were used, because they are phonologically minimally different from *wh*-phrases, *nani* ‘what’, *dare* ‘who’, etc. I speculate, however, that this similarity made it difficult for the subjects to notice the difference between *wh*-phrase and non-*wh*-counterpart. To my ear, some subjects consistently P-focalized the indefinites as well. As a result, the expected contrast became much smaller than expected. Note that the contrast on P2 is also very small between A and B, as is clear from (29) ( $p > 0.333$ ). Such a lack of contrast is unexpected, given that the P-focalization effect is clearly attested on P1 (note the difference on P1 between A/B and C/D), where no indefinites were used for the non-*wh*-counterparts. Also note that P2 in B, the F<sub>0</sub> peak of the indefinites, is almost as high as P1. This mean value for P2 is slightly higher than we would expect, given that the time-dependent declination effect would make P2 lower than P1. The speculation about the unexpected P-focalization of indefinites would naturally explain these apparently unexpected facts. Since we do not observe a contrast between A and B, we cannot expect a contrast between C and D either. Given these considerations, the fact that the prediction (27II) is not borne out does not necessarily falsify the analysis.

On the contrary, the other two predictions, (27III) and (27IV), are supported by the result. Given that these contrasts are found within the matrix PFR domain, as (27I) ensures, they clearly indicate the existence of FI embedding.

In this section, we tested the first of the two predictions made by the Multiple Spell-Out model, namely, FI embedding, and discussed the result of the experiment. Although the embedded P-focalization effect was not confirmed by the result, the embedded PFR effect, along with the pitch reset after it, was attested. This result strongly supports the Multiple Spell-Out model proposed in §3. In the next section, we will test the other prediction, namely, FI–WH mismatch.

## 6 Experiment 2: FI–WH Mismatch (FI $\neq$ WH)

In the previous section, we saw that the FI embedding is in fact attested, confirming the first prediction made by the Multiple Spell-Out model. In this section, we will examine the second prediction, namely, the FI–*Wh*-scope Mismatch (FI $\neq$ WH). In this experiment, we will examine the pitch contour of Saito’s (1989) radical reconstruction sentences like (21b).

### 6.1 Method

The procedure of the experiment is exactly the same as the one in the FI embedding experiment (see §5.1), except that the number of target sentences is 28 instead of 32, and the number of filler sentences is 108 instead of 104.

### 6.2 Stimuli

Stimuli are made of 7 sets of four sentence types (28 sentences in total), one of which is given below:

(31) *Stimulus set example*A. *No scrambling, Non-wh-sentence*

Náoya-wa [ Mári-ga **rámu**-o nomíya-de nónda to ]  
 Naoya-TOP Mari-NOM rum-ACC bar-LOC drank that  
 ímademo omóteru  
 even.now 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** ]  
 Naoya-TOP Mari-NOM what-ACC bar-LOC drank Q  
 ímademo obóeteru  
 even.now remember

‘Naoya still remembers what<sub>i</sub> Mari drank  $t_i$  at the bar.’

C. *Scrambling, Non-wh-sentence*

**rámu**<sub>i</sub>-o Náoya-wa [ Mári-ga  $t_i$  nomíya-de nónda to ]  
 rum-ACC Naoya-TOP Mari-NOM bar-LOC drank that  
 ímademo omóteru  
 even.now think

‘Naoya still thinks that Mari drank rum at the bar.’

D. *Scrambling, Indirect wh-question*

**náni**<sub>i</sub>-o Náoya-wa [ Mári-ga  $t_i$  nomíya-de nónda **ka** ]  
 what-ACC Naoya-TOP Mari-NOM bar-LOC drank Q  
 ímademo obóeteru  
 even.now remember

‘Naoya still remembers what<sub>i</sub> Mari drank  $t_i$  at the bar.’

(31A) and (31B) are sentences with a canonical word order (i.e., no scrambling). (31B) is an embedded *wh*-question, containing a *wh*-phrase and a Q-particle in the embedded clause.



(31C) and (31D) are the scrambled versions of (31A) and (31B), respectively. (31D) is Saito's (1989) example, where the embedded *wh*-phrase is scrambled to the beginning of the matrix clause.

### 6.3 Predictions

In this experiment, we are interested in the FI domain of sentences like (31D). What we need to verify is to see whether the FI domain continues after the embedded clause (as the Multiple Spell-Out model predicts) or not (as claimed earlier by Ishihara, 2002; Kitagawa and Fodor, 2003). To test this, we focus on the  $F_0$ -peak of the embedded Verb (P1) and that of the phrase after the embedded clause (P2). (In (31): P1 = *nón*da; P2 = *í*mademo)

- (32) *Labels of the relevant  $F_0$  peaks*
- |  |
|--|
| [ <sub>CP</sub> ((Non-)WH) ... [ <sub>CP</sub> ... ((Non-)WH) ... Verb Q ] α ... ] |
| P1                  P2   |

**P1: Embedded clause verb** P1 is inside the embedded CP. Hence it will be lowered if an FI is created either at the embedded CP cycle or at the matrix CP cycle.

**P2: Material immediately following the embedded clause** P2 is outside the embedded CP. Hence it will be lowered only if an FI is created at the matrix CP cycle. It will be insensitive to the FI within the embedded CP.

Under the Multiple Spell-Out model, we will have the following predictions for the non-scrambling sentences (A and B) and for the scrambling sentences (C and D), Respectively:

## (33) Non-scrambling sentences: A vs. B

A.  $[_{CP} [_{TP} \dots [_{CP} [_{TP} \dots \text{Non-WH} \dots P1] C] P2 \dots]]$   
 $\uparrow \quad \uparrow$   
 No PFR No PFR

B.  $[_{CP} [_{TP} \dots [_{CP} [_{TP} \dots \boxed{\text{WH}} \dots P1] \mathbf{ka}] P2 \dots]]$   
 $\uparrow \quad \uparrow$   
 PFR No PFR

- In B, an FI is created at the embedded CP cycle.
- P1 of B is lower than in A due to the PFR.
- P2 of A and B are of the same height, because P2 should not be affected by the PFR in the embedded CP cycle. Hence, a pitch reset takes place.

## (34) Scrambling sentences: C vs. D

C.  $[_{CP} \text{Non-WH}_i [_{TP} \dots [_{CP} [_{TP} \dots t_i \dots P1] C] P2 \dots]]$   
 $\uparrow \quad \uparrow$   
 No PFR No PFR

D.  $[_{CP} \boxed{\text{WH}_i} [_{TP} \dots [_{CP} [_{TP} \dots t_i \dots P1] \mathbf{ka}] P2 \dots]]$   
 $\uparrow \quad \uparrow$   
 PFR PFR

- In D, an FI is created at the matrix CP cycle.
- P1 of D is lower than that of C.
- P2 of D is also lower than that of C.

## 6.4 Result and discussion

Data of the four subjects (AH, CS, CK and YY) are normalized.<sup>26</sup>

<sup>26</sup> Again, the data of one subject (NM) is excluded from the analysis. See fn. 24.

(35) *Data normalization*a. *Formula for normalization:*

$$y = \frac{x - R_2}{R_1 - R_2}$$

where  $R_1$  and  $R_2$  are the reference points calculated independently for each subject.

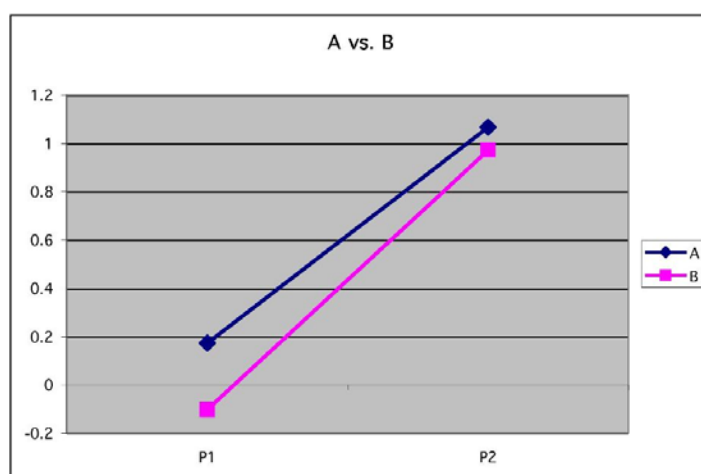
b. *Reference points ( $R_1, R_2$ ):*

- $R_1$  = Mean value of P2 ( $F_0$ -peak on the phrase immediately following the embedded clause)
- $R_2$  = Mean value of P1 ( $F_0$ -peak on the embedded verb)

The normalized data show the expected results. In the non-scrambled sentences A and B, P1 (the embedded verb) is lowered in B due to the PFR after the *wh*-phrase. The difference between A and B is statistically significant. On P2 (the post-embedded-CP phrase), although there still is a difference between A and B, it is much smaller than the one on P1. It is in fact statistically not significant. This means that in B a pitch reset takes place and the pitch register of the P2 is set back to the non-reduced value. Hence there is no more significant difference on P2.

## (36) A vs. B

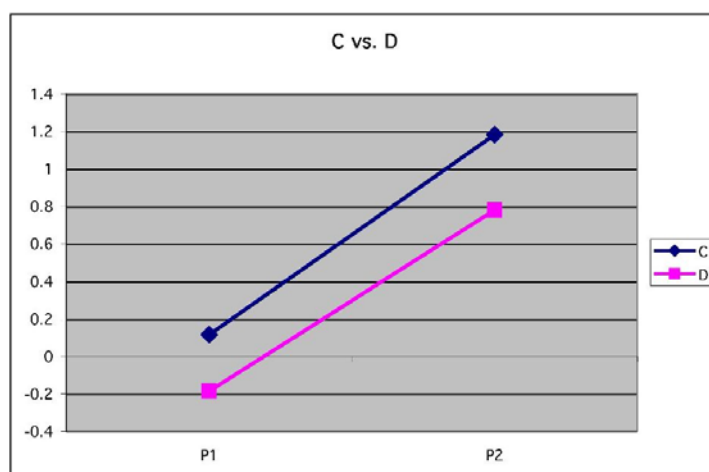
	A	B	diff.	p
Mean(P1)	0.174	-0.103	0.276	< .001
Mean(P2)	1.066	0.971	0.095	= .257



In the scrambled sentences C and D, P1 shows the same result as in the non-scrambled version, as expected. P1 is lower in D than in C due to the PFR. On P2, the differences between C and D are not reduced at all, and in fact, they are still statistically significant. This means that the PFR continues to the matrix material, unlike the non-scrambled version.

(37) C vs. D

	C	D	diff.	p
Mean(P1)	0.115	-0.185	0.301	< .001
Mean(P2)	1.182	0.780	0.402	< .0001



The result of the experiment reinforces the Multiple Spell-Out analysis proposed in this paper, as the prediction is in fact empirically supported. This result also denies the earlier empirical claim in Ishihara (2002) and Kitagawa and Fodor (2003) that the FI is always observed between the *wh*-phrase and the Q-particle.<sup>27</sup>

(38) a. \* *Observation in Ishihara (2002); Kitagawa and Fodor (2003)*

[<sub>CP</sub> WH [<sub>TP</sub> α ... [<sub>CP</sub> [<sub>TP</sub> ... *t*<sub>WH</sub> ... ] **ka** ] β ... ]  
↑  
No PFR

b. *Actually attested pitch contour*

[<sub>CP</sub> WH [<sub>TP</sub> α ... [<sub>CP</sub> [<sub>TP</sub> ... *t*<sub>WH</sub> ... ] **ka** ] β ... ]  
↑  
PFR

This result is particularly important because it suggests that FI=WH is not a result of the direct phonology-semantics interaction. If it were the case, we would expect an FI only inside the embedded clause in Saito's (1989) example like (31D). FI=WH is rather a result of the cyclic computation, which usually computes the domain of FI and the *wh*-scope at the same phase, unless the syntactic movement creates a mismatch between the phonological domain of FI and the semantic *wh*-scope.

In this section, we examined the second prediction of the Multiple Spell-Out analysis, namely, the FI–*Wh*-scope Mismatch. The result of the experiment

<sup>27</sup> An question remains as to why both Kitagawa and Fodor (2003) and I (Ishihara, 2002) acknowledged that (38a) is the correct pitch contour. In fact, I still feel that (38a) is not entirely impossible. 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 (which is in fact the main point of discussion in Ishihara, 2002 and Kitagawa and Fodor, 2003), and maybe also because I may be too sensitive to the FI–*Wh*-scope correspondence to give a naive judgement. If, however, this intuition turns out to be real, there must be some additional mechanism that allows a contour like (38a), because the Multiple Spell-Out model would never allow such a contour. I will leave this question for future research. In this paper, I will take the result of the experiment as the real and correct description of the fact.

presented in this section in fact supports this prediction. When a *wh*-phrase is scrambled out of its *wh*-scope, the FI creation is postponed to a later Spell-Out cycle, and the domain is extended. As a result, the FI domain and the *wh*-scope no longer shows a correspondence. Together with the FI embedding discussed in §5, this experimental result strongly supports the proposed analysis.

## 7 Conclusion

In this paper, we discussed the Focus Intonation–*Wh*-scope Correspondence (FI=WH) in Japanese *wh*-questions. I proposed that FI=WH is derived by the cyclic syntactic computation and the Spell-Out mechanism.

The Multiple Spell-Out model proposed here is further supported by the results of the two experiments. The first experiment showed that FIs may be embedded when there are two WH-Q dependencies that take different scopes. FI embedding is naturally explained under the proposed model. The second experiment showed that FI=WH breaks down once the *wh*-phrase is scrambled out of its *wh*-scope. The *wh*-scope remains the same if scrambling takes place, thanks to the radical reconstruction effect. The FI prosody, in contrast, is created later in the derivation, namely, at the Spell-Out domain at which the scrambled *wh*-phrase is transferred to the phonological component. As a result, FI≠WH takes place.

This analysis not only explains FI=WH and FI≠WH in Japanese *wh*-questions, but also has further theoretical implications. First, under this analysis, the phonological component computes prosodic information in a cyclic fashion. This means that not only segmental phonological material, but also suprasegmental information such as intonation is computed cyclically phase by phase, and superimposed each time. The FI embedding experiment (§5) suggests that this is in fact the case. If so, it raises further interesting questions such as how the phonological component implements such cyclic suprasegmental in-

formation, how it is realized phonetically, how the cyclic Spell-Out is related to phonological phrasing, etc.

Also, this analysis gives support for the phase ‘edge’ position. In the current Minimalist framework, phase ‘edge’ positions are needed at the syntactic component to allow successive cyclic movement. The material (dis)located to this position escapes from Spell-Out at this phase, remaining accessible to the next phase. The FI–WH Mismatch experiment (§6) provides support for this claim. Material moved to this position is in fact spelt-out at a later cycle.<sup>28</sup>

As interesting discussion has already been made recently (see §1 for references), prosody and its impact on syntactic ‘judgment’ has to be studied more in detail. What is interesting about the prosody of Japanese *wh*-questions is that FIs appear obligatorily in the sentence. The situation is clearly different from non-*wh*-sentences. Since the appearance of focus heavily depends on the discourse and information structure of the sentence, an FI may or may not appear in a non-*wh*-sentence, depending on the context. This does not necessarily mean, however, that *wh*-questions may not have any additional FIs optionally. Some *wh*-question sentences may contain both obligatory and optional FIs. It is therefore important for future research to specify how these ‘obligatory’ and ‘optional’ FIs may interact with each other. Such studies would help us understand better how prosody influences syntactic judgments.

## Bibliography

Bresnan, Joan W. 1972. On sentence stress and syntactic transformation. In *Contributions to generative phonology*, ed. Michael Brame, 73–107. Austin:

<sup>28</sup> The real test case would be where the scrambled *wh*-phrase stops at the specifier of the embedded CP. The Multiple Spell-Out model predicts that the FI will be created at the matrix *v*P phase. This property could potentially be used to distinguish the two types of scrambling. The scrambling to Spec,CP is considered to be A'-scrambling, while the scrambling to Spec,TP is A-scrambling (Miyagawa, 2001). If so, A'-scrambling would create an FI at the matrix *v*P phase, while A-scrambling creates an FI at the embedded CP phase. I leave these further questions for future research.

University of Texas Press.

Chomsky, Noam. 2000. Minimalist inquiries: The framework. In *Step by step: In honor of Howard Lasnik*, ed. Roger Martin, David Michaels, and Juan Uriagereka, 85–155. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001a. Derivation by phase. In Kenstowicz (2001), 1–52.

Chomsky, Noam. 2001b. *Beyond explanatory adequacy*, volume 20 of *MIT Occasional Papers in Linguistics*. Cambridge, MA: MITWPL.

Deguchi, Masanori, and Yoshihisa Kitagawa. 2002. Prosody and wh-questions. In *NELS 32: Proceedings of the North East Linguistics Society*, ed. Masako Hirotani, volume 1, 73–92. Amherst, MA: GLSA, University of Massachusetts, Amherst.

Féry, Caroline, and Hubert Truckenbrodt. 2003. Sisterhood and tonal scaling. Unpublished ms.

Hiraiwa, Ken. 2002. Indeterminate agreement and varieties of raising in Japanese. Paper accepted at the 21st West Coast Conference on Formal Linguistics (WCCFL XXI).

Hirotani, Masako. 2003. Prosodic effect on the comprehension of Japanese wh-question. Poster presentation at the Sixteenth Annual CUNY Conference on Human Sentence Processing, May 27–29 2003.

Ishihara, Shinichiro. 2002. Invisible but audible wh-scope marking: Wh-constructions and deaccenting in Japanese. In *Proceedings of the 21st West Coast Conference on Formal Linguistics (WCCFL 21)*, ed. Line Mikkelsen and Chris Potts, 180–193. Somerville, MA: Cascadilla Press.

Ishihara, Shinichiro. 2003. Intonation and interface conditions. Doctoral Dissertation, Massachusetts Institute of Technology.

Ishihara, Shinichiro. 2004. Where does ‘?’ come from?: Syntax-phonology cyclicity mismatch. Talk presented at the Workshop on Prosody, Syntax and Information Structure (WPSI), May 27–29 2004.



- Kenstowicz, Michael, ed. 2001. *Ken Hale: A life in language*. Cambridge, MA: MIT Press.
- Kitagawa, Yoshihisa, and Janet Dean Fodor. 2003. Default prosody explains neglected syntactic analysis in Japanese. In *Japanese/Korean Linguistics*, ed. William McClure, volume 12, 267–279. Stanford: CSLI Publications.
- Kitagawa, Yoshihisa, and Satoshi Tomioka. 2003. Masked island effects in Japanese. Talk presented at Workshop on Altaic Formal Linguistics (WAFL), May 16–18 2003.
- Kubozono, Haruo. 1993. *The organization of Japanese prosody*. Tokyo: Kurosio Publishers.
- Kubozono, Haruo. 2004. Downstep and phonological phrasing in Japanese. Talk presented at the Workshop on Prosody, Syntax and Information Structure (WPSI), May 27–29 2004.
- Kuroda, S.-Y. 1965. Generative grammatical studies in the Japanese language. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Kuroda, S.-Y. 2004. Prosody and the syntax of indeterminates. Talk presented at the Workshop on Prosody, Syntax and Information Structure (WPSI), May 27–29 2004.
- Maekawa, Kikuo. 1991a. Perception of intonation characteristics of WH and non-WH questions in Tokyo Japanese. In *Proceedings of the XXII<sup>nd</sup> International Congress of Phonetic Science*, volume 4, 202–205. Aix-en-Provence, France: Université de Provence.
- Maekawa, Kikuo. 1991b. Tokyo hoogen gimon-bun no intoneesyon [The intonation of question sentences in Tokyo dialect]. In *Heisei 3-nendo Nihon Onsei Gakkai Zenkoku Taikai Kenkyuu Happyoo Ronshyuu [Proceedings of the 2001 PSJ Meeting]*, 42–47. Nihon Onsei Gakkai.
- Mahajan, Anoop. 1994. Toward a unified theory of scrambling. In *Studies on scrambling: Movement and non-movement approach to free word-order*

- phenomena*, ed. Norbert Corver and Henk van Riemsdijk, 301–330. Berlin and New York: Mouton de Gruyter.
- 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.
- Nespor, Marina, and Irene Vogel. 1986. *Prosodic phonology*. Dordrecht: Foris Publications.
- Nishigauchi, Taisuke. 1990. *Quantification in the theory of grammar*. Dordrecht: Kluwer Academic Publishers.
- Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese tone structure*. Cambridge, MA: MIT Press.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1:75–116.
- Saito, Mamoru. 1989. Scrambling as semantically vacuous A'-movement. In *Alternative conceptions of phrase structure*, ed. M. Baltin and A. Krock, 192–200. University of Chicago Press.
- Selkirk, Elisabeth. 1993. The prosodic structure of function words. In *Signal to syntax: Bootstrapping from speech to grammar in early acquisition*, ed. James L. Morgan and Katherine Demuth, 187–213. Mahwah, NJ: Lawrence Erlbaum Associates.
- Selkirk, Elisabeth. 2000. The interaction of constraints on prosodic phrasing. In *Prosody: Theory and experiment*, ed. Merle Horne, 231–261. Dordrecht and Boston and London: Kluwer Academic Publishers.
- Selkirk, Elisabeth. 2003. Bengali intonation revisited: An optimality theoretic analysis in which FOCUS stress prominence drives FOCUS phrasing. In *Papers in optimality theory II*, ed. A. Carpenter, A. Coetzee, and P. de Lacy, volume 26 of *University of Massachusetts Occasional Papers (UMOP)*. Amherst, MA: GSLA, University of Massachusetts, Amherst.

- Selkirk, Elisabeth, and Koichi Tateishi. 1991. Syntax and downstep in Japanese. In *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, ed. Carol Georgopoulos and Roberta Ishihara, 519–543. Dordrecht: Kluwer Academic Publishers.
- Selkirk, Elisabeth O. 1984. *Phonology and syntax: The relation between sound and structure*. Cambridge, MA: MIT Press.
- Shimoyama, Junko. 2001. Wh-constructions in Japanese. Doctoral Dissertation, University of Massachusetts, Amherst.
- Sugahara, Mariko. 2003. Downtrends and post-FOCUS intonation in Japanese. Doctoral Dissertation, University of Massachusetts, Amherst, MA.
- Tomioka, Satoshi. 1997. Wh-insitu, subjacency, and LF syntax. Handout from LSA Summer Institute, Cornell University, Ithaca, NY.
- Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge, MA.
- Truckenbrodt, Hubert, and Caroline Féry. 2003. More on hierarchical organization and tonal scaling. Unpublished ms.
- Yoshida, Keiko, and Tomoyuki Yoshida. 1996. Question marker drop in Japanese. *ICU Language Research Bulletin* 11:37–57.
- Yoshida, Tomoyuki. 1998. Wh-operator vs. Yes/No-operator. *ICU Language Research Bulletin* 13:159–172.

*Shinichiro Ishihara*  
*Universität Potsdam*  
*SFB632 „Informationsstruktur“*  
*Postfach 601553*  
*14415 Potsdam*  
*Deutschland*  
*ishihara@rz.uni-potsdam.de*  
*[http://alum.mit.edu/www/s\\_i/](http://alum.mit.edu/www/s_i/)*