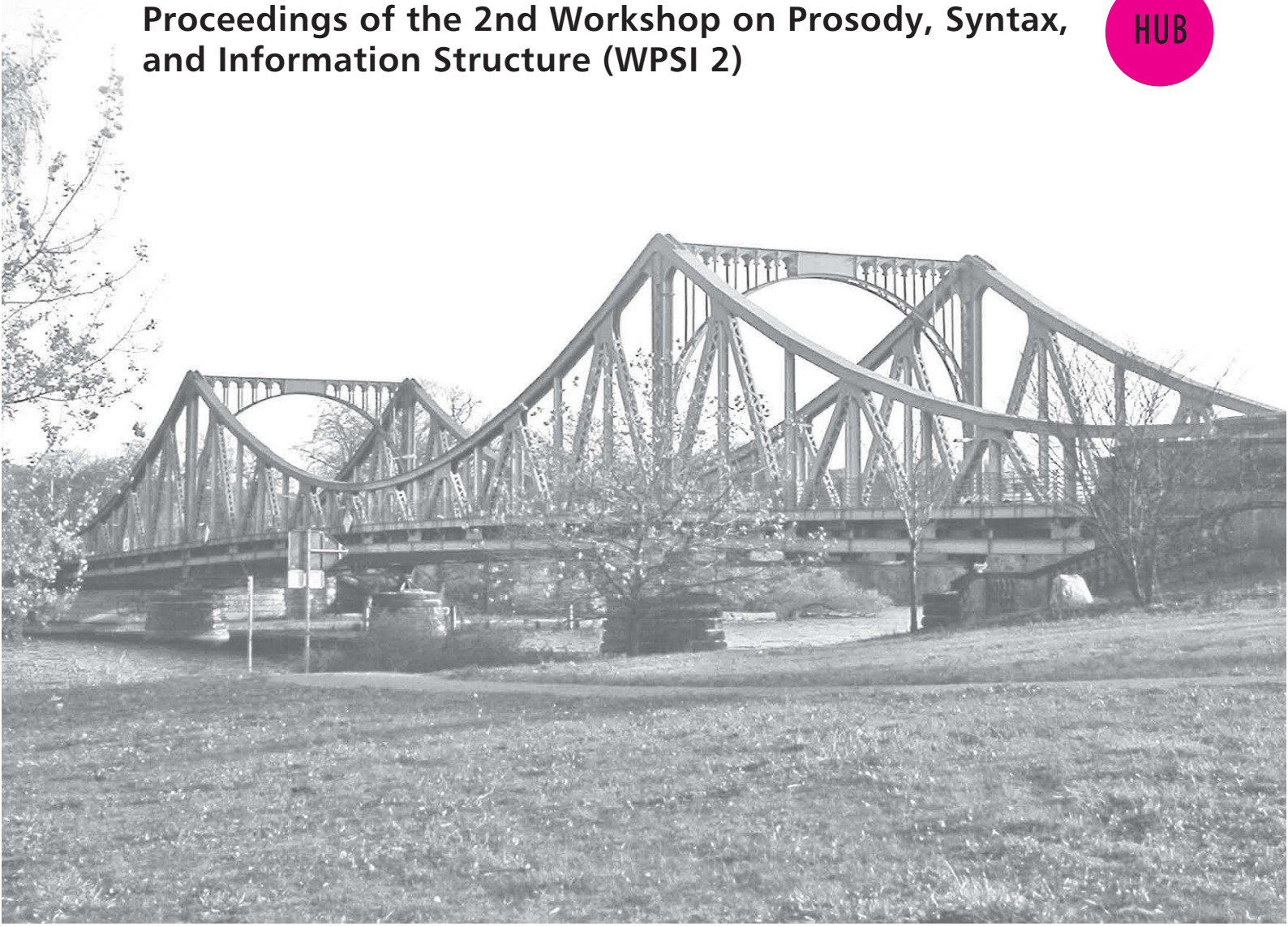


Interdisciplinary Studies on Information Structure Vol. 9

UP

Proceedings of the 2nd Workshop on Prosody, Syntax,
and Information Structure (WPSI 2)

HUB



Interdisciplinary Studies on Information Structure

**Proceedings of the 2nd Workshop on Prosody,
Syntax, and Information Structure (WPSI 2)**

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Universitätsverlag Potsdam 2007

<http://info.ub.uni-potsdam.de/verlag.htm>

Universitätsverlag Potsdam, Am Neuen Palais 10, 14469 Potsdam

Tel.: +49 (0)331 977 4623 / Fax: 4625

E-Mail: ubpub@uni-potsdam.de

Interdisciplinary Studies on Information Structure (ISIS)

9 (2007)

Series Editors:

isis.editors@googlemail.com

<http://www.sfb632.uni-potsdam.de/isis.html>

Shinichiro Ishihara

Universität Potsdam, SFB 632

Am Neuen Palais 10, D-14469 Potsdam

Svetlana Petrova | Anne Schwarz

Humboldt-Universität zu Berlin, SFB 632

Sitz: Mohrenstr. 40 - 41

Unter den Linden 6, D-10099 Berlin

ISIS issues do not appear according to strict schedule.

Copyrights of articles remain with the authors.

ISSN 1866-4725

Online published at the Institutional Repository of the Potsdam University

URL <http://pub.ub.uni-potsdam.de/volltexte/2008/2223/>

URN <urn:nbn:de:kobv:517-opus-22234>

[<http://nbn-resolving.de/urn:nbn:de:kobv:517-opus-22234>]

Preface

The 2nd Workshop on Prosody, Syntax, and Information Structure (WPSI2) was held at University of Potsdam on March 18, 2005. This workshop was held as a continuation of the one held in the previous year at Indiana University.

WPSI 2 was aimed to discuss issues on the interaction of prosody, syntax, and information structure with interdisciplinary perspectives. All the speakers of this workshop have been recently working on relevant issues, especially looking at the phenomena related to the intonation of focus and (*wh*-)questions in Japanese.

This volume includes papers from four of the five speakers of this workshop: Haruo Kubozono, Shinichiro Ishihara, Yoshihisa Kitagawa, and Satoshi Tomioka. (Unfortunately, Masako Hirotsu could not contribute to our volume.)

Haruo Kubozono's "Focus and Intonation in Japanese: Does Focus Trigger Pitch Reset?" presents experimental data that cast doubt on the traditional analysis that focus blocks downstep. The results of his experiment show that a *wh*-phrase does not block downstep, although it receives a higher pitch than its preceding element. He concludes that focus does not trigger a pitch reset.

Yoshihisa Kitagawa's "When We Fail to Question in Japanese" discusses several interesting cases of *wh*-questions in which certain interpretations are unavailable with certain prosodic representations. He proposes an elaborated version of the "E-agreement" analysis, and claims that this model, together with some extra-grammatical factors, accounts for those phenomena that are otherwise puzzling.

Shinichiro Ishihara's "Intonation of Sentences with an NPI" reports the results of a phonetic production experiment that examines the intonation of sentences containing a negative polarity item (NPI). The results show that an NPI such as *-sika* in Tokyo Japanese triggers a focus intonation: a focal F_0 -rise on the NPI followed by a focal F_0 -downtrend that continues until the negation that binds the NPI. This contour is parallel to that of *wh*-questions, which also exhibit a focus intonation between a *wh*-phrase and the question particle binding it.

Satoshi Tomioka's "Intervention Effects in Focus: From a Japanese Point of View" discusses the so-called LF-intervention effects. After introducing analyses based on the focus semantics by Beck (2006) and Kim (2002, 2005), he points out some pragmatic constraints observed in the intervention phenomena

in Japanese, which the focus-based analyses cannot handle by themselves. Based on his pragmatic account, he suggests some modifications to the focus-based analyses, and concludes that focus-based accounts need be augmented by some auxiliary constraint(s) to account for the complexity and subtlety of the intervention phenomena.

As for the WPSI series, the latest meeting, WPSI 3, was held on September 14–15, 2007 at Indiana University. Also this research program has been further studied under the NSF Grant Research “Wh-interrogatives at the Prosody-Syntax-Pragmatics Crossroad.” See the web sites for detail. I hope that this volume, together with the WPSI meetings and researches at the NSF project mentioned above, interests many colleagues working on interface issues between prosody, syntax, and information structure, and stimulates further discussion in the field.

Shin Ishihara

Related web sites:

WPSI 1: <http://www.iub.edu/~ykring/wpsi.html>

WPSI 2: http://alum.mit.edu/www/s_i/misc/wpsi2/

WPSI 3: http://www.indiana.edu/%7Egpsi/wpsi_index.html#

NSF Grant Research

“Wh-interrogatives at the Prosody-Syntax-Pragmatics Crossroad”:

<http://www.ling.udel.edu/focus/>

Contents

Focus and Intonation in Japanese: Does Focus Trigger Pitch Reset? <i>Haruo Kubozono</i>	1
When We Fail to Question in Japanese <i>Yoshihisa Kitagawa</i>	29
Intonation of Sentences with an NPI <i>Shinichiro Ishihara</i>	65
Intervention Effects in Focus: From a Japanese Point of View <i>Satoshi Tomioka</i>	97

Focus and Intonation in Japanese: Does Focus Trigger Pitch Reset?*

Haruo Kubozono

Kobe University

This paper discusses how focus changes prosodic structure in Tokyo Japanese. It is generally believed that focus blocks the intonational process of downstep and causes a pitch reset. This paper presents experimental evidence against this traditional view by looking at the prosodic behavior of Wh words, which receive focus lexically in Japanese as in other languages. It is demonstrated, specifically, that the focused Wh element does not block downstep although it receives a much higher pitch than its preceding element. This suggests that presence of lexical focus does not trigger pitch reset in Japanese.

Keywords: Focus, Intonation, Japanese, Pitch Reset, Downstep

1 Introduction

The relationship between prosody and information structure is attracting serious attention in the literature (e.g. Zubizarreta 1998). This topic has been discussed in the literature of Japanese prosody, too (see, for example, Truckenbrodt 1995 and Ishihara 2003), but there is not much work specifically dealing with the interface between intonation and focus. As far as I know, Poser (1984) presented the first experimental work that looked at this issue in terms of the prosodic

* An earlier version of this paper was read at the Second Workshop on Prosody, Syntax and Information Structure (WPSI 2) held at Potsdam University in March 2005. I would like to thank the participants in this workshop for various valuable comments. I am also grateful to Donna Erickson for checking and commenting on this manuscript, to Shingo Sugiyama and Shinji Ogawa for their help in the experiment and to all those who participated in the experiment as subjects. All remaining errors are of course my own. The work reported in this paper has been supported by Grant-in-Aid for Scientific Research (A) (Japan Society for the Promotion of Science, Grant no. 17202010).

organization of utterances in (Tokyo) Japanese.¹ By carrying out an experiment in which a certain element bears contrastive focus (or emphasis, in his terminology), Poser reported that focus does not block the intonational process of catathesis (or downstep) and, hence, that focus does not introduce a major intonational phrase boundary. On the other hand, Pierrehumbert and Beckman (1988) (henceforth P&B) claimed that focus blocks catathesis/downstep and, hence, starts a new major intonational phrase which they called ‘intermediate phrase’.² Since this intonational phrase is defined as the domain of pitch reset, as we will see shortly below, P&B argued that focus resets the pitch range. Their view regarding the interaction between focus and intonation structure seems popularly accepted in the studies of Japanese prosody in general.

As in the general literature dealing with focus, both Poser and P&B mean contrastive or corrective focus by the term ‘focus’ and they are not different in this respect. However, they reached entirely different conclusions as regards the effect of focus on intonation structure. One thing that can be said with some certainty is that it is generally difficult to obtain sufficiently natural speech when we examine contrastive focus in a controlled experiment. When we use corrective or contrastive focus in the material, subjects tend to exaggerate the contrast and hence to put extra boost on the focused element. There are two ways to avoid this potential problem: one is to analyze spontaneous speech in a non-controlled experiment, while the other is to look at the kind of focus in a controlled experiment that does not involve contrast. We chose the second option in this study and decided to analyze sentences with a Wh element which

¹ In this paper, ‘Japanese’ refers to the standard Tokyo Japanese unless otherwise stated.

² “Focus blocked propagation of catathesis...focus caused the introduction of a prosodic boundary, at which the pitch range was reset. We will call this level of phrasing the intermediate phrase” (P&B 1988: 19–20).

is known to receive focus lexically, independent of the pragmatics of the sentence.

Before going into the details of the experiment, let us define some basic notions and shared ideas about the intonation structure of Japanese (section 2). In section 3, we will review past works on the interface between focus and intonation by specifically comparing Poser's (1984) experimental results with P&B's (1988). In section 4, we will sketch our experiment and present its major results. This will be followed by a discussion of the data in section 5, where we consider the implications of the data for the modeling of Japanese intonation, especially as regards the hierarchical organization of prosody.

2 Background

We need to define four basic notions here: lexical accent, downstep, minor phrase and major phrase.

2.1 Lexical accent

'Lexical accent' is ambiguous in Japanese phonetics and phonology. First of all, lexical items in Tokyo Japanese fall into two types: 'accented' and 'unaccented'. 'Accented' words involve a sudden pitch fall at the phonetic output, whereas 'unaccented' words do not show any such fall even when they are followed by a particle (like the nominative particle *ga*). This distinction is generally preserved in phrases and sentences, where accented and unaccented words give rise to phrases with or without a sudden pitch fall, respectively. Some examples of the two lexical types are given in (1): accent is denoted by an apostrophe placed on the vowel immediately before the pitch fall, whereas unaccented words are transcribed with no such mark.

- (1) a. Accented words
 kyóoto ‘Kyoto’, kóobe ‘Kobe’, nagásaki ‘Nagasaki’, náoko ‘Naoko’,
 dóitu ‘Germany’, pótudamu ‘Potsdam’, umái ‘tasty’, haréru ‘to clear’
- b. Unaccented words
 tookyoo ‘Tokyo’, oosaka ‘Osaka’, hirosima ‘Hiroshima’, naomi ‘Naomi’,
 berurin ‘Berlin’, amerika ‘America’, amai ‘sweet’, hareru ‘to swell up’

The term ‘lexical accent’ sometimes refers to accent patterns specified at the word-level, including the unaccented one in (1b). This convention of referring to both accented and unaccented words as having lexical accent is popular in the traditional literature of phonetics and phonology in Japan (Akinaga 1985). Quite often, however, the term ‘lexical accent’ is used to refer to the sudden pitch fall observed in accented words; words in (1a) bear a lexical accent on the vowel marked by the accent mark, whereas those in (1b) lack such a phonological feature. In this paper, we will follow this second convention, thus characterizing (1b) as lexical items with no ‘lexical accent’.

The distinction between accented and unaccented words is supposed to be lexical in the sense that all morphologically simplex words are specified in the lexicon with respect to their accentedness. While this is largely true in the native vocabulary, it is not always the case in the foreign and Sino-Japanese vocabulary. Recent studies on Japanese accent have revealed that accentedness is predictable to a considerable extent in these two types of words, especially on the basis of their syllable and mora structure (Kubozono 1996, Kubozono & Fukui 1996, Ogawa 2004).

2.2 Downstep

The lexical distinction between accented and unaccented words exerts profound effects on Japanese intonation. In addition to the presence or absence of an abrupt pitch fall, accented and unaccented words/phrases also differ in pitch

height (see Figure 1 below). Namely, the former exhibits a higher pitch peak than the latter (Poser 1984); very often, it also starts with a higher pitch than the latter (Kubozono 1988/1993). The extra pitch boost in accented words, or ‘accentual boost’ (Kubozono 1988), can be interpreted as a concomitant feature of the sudden pitch fall that follows. In other words, accented words receive a higher pitch in order to display the following pitch fall more clearly.

Lexical accent exerts a considerable effect on the following material, too, by lowering the pitch level of the following phrase. This can be seen very clearly by comparing two sentences or expressions, one with an accented first phrase and the other with an unaccented one. In Figure 1, for example, the second phrase *nomímono* ‘drink’ is realized at a considerably lower level when it follows an accented phrase *umái* ‘tasty’ than when it follows an unaccented one *amai* ‘sweet’.

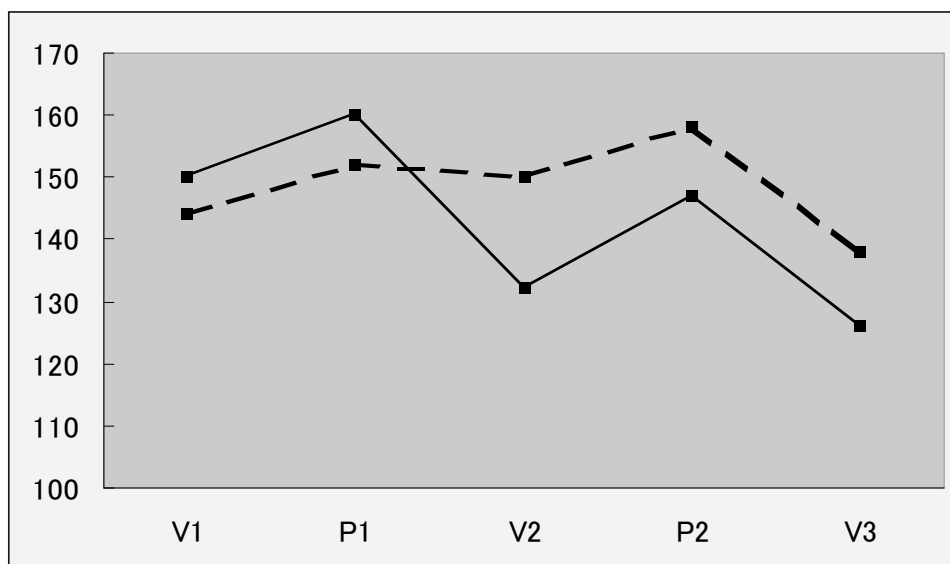


Figure 1 Schematic comparison of two contours: a sequence of two accented phrases (solid line) vs. a sequence of an unaccented phrase plus an accented one (dashed line).

Poser (1984) and P&B (1988) called this lowering process ‘catathesis’, for which Kubozono (1988), Selkirk and Tateishi (1991) and others used the term ‘downstep’. We will use the latter term throughout this paper.

Figure 1 illustrates the two effects of lexical accent on Japanese intonation. Accented words are realized at a higher pitch level than their unaccented counterparts on the one hand (*accentual boost*), but they realize the following material in a lower region on the other (*downstep*). Because of these effects, the two contours in Figure 1 show a clear separation when the pitch peak in the second phrase is plotted as against the pitch peak in the first phrase: the solid line has a greater pitch value than the dashed line for the first phrase but a smaller pitch value for the second phrase (P&B 1988).

In addition to these, it was also demonstrated that downstep occurs iteratively (Poser 1984, Kubozono 1988); the pitch level goes further down if a third phrase follows in Figure 1. In this sense, the downstep effect is cumulative.

Whether downstep has occurred between two adjacent phrases can be tested in two independent ways. In a syntagmatic dimension, it can be tested by comparing the pitch level of two adjacent phrases within a single utterance. In Figure 1, for example, the second phrase is downstepped in the solid line since it is lower in pitch than its preceding phrase. This high-low relation of two adjacent phrases within a single utterance is the syntagmatic clue to downstep. In a paradigmatic dimension, on the other hand, downstep can be defined by comparing the pitch contours of two utterances, like the two contours in Figure 1, one beginning with an accented phrase and the other beginning with an unaccented one. Downstep can be identified if the second phrase is realized at different pitch levels depending on the accentedness of the first phrase. Unlike the syntagmatic definition, this definition of downstep does not require us to

compare two peaks within a single utterance, but rather the peaks of corresponding elements on two different utterances.

Of the two definitions of downstep, it is the paradigmatic definition that was originally proposed (Poser 1984), and was generally adopted in 1980s (e.g., P&B 1988, Kubozono 1988). In contrast, the syntagmatic definition was adopted popularly in 1990s and later years (e.g., Selkirk & Tateishi 1991, Nagahara 1994). This approach to downstep, although more simplistic than the paradigmatic one, poses some serious problems in actual prosodic analysis. The most serious problem is that of ambiguity concerning the identification of downstep. Namely, it is not clear how much lower a given phrase has to be before it is recognized as being ‘downstepped’ as relative to its preceding phrase. It often happens that a phrase is only slightly lower in pitch than its preceding accented phrase. In such a case, it is difficult to tell in any objective sense whether or not downstep has occurred between the two phrases. Equally serious is the existence of cases where the two definitions of downstep do not agree with each other. These cases will be discussed in detail in sections 3 through 5 below.

The process of downstep thus defined has both phonological and phonetic aspects. It is a phonological process in the sense that it is triggered by a phonological factor, i.e., lexical accent. It is a phonetic process, on the other hand, if one looks at its effect. The second phrase in Figure 1 shows a bigger pitch rise in the utterance with downstep (solid line) than in the utterance without downstep (dashed line) to compensate for the greater pitch fall triggered by the lexical accent. This compensatory pitch rise is obviously insufficient, with the result that the downstepped phrase nevertheless attains a lower peak (P2 in Figure 1) than the non-downstepped phrase. Thus, the phrase following an accented phrase is realized at a lower pitch level than the one following an

unaccented phrase. This whole process is phonetic in nature since the magnitude of the lowering process is not binary but variable depending on several factors such as the phonological length of the phrases, the location of accent, etc. (Kubozono 1988). That the effect of downstep is basically phonetic rather than phonological can also be supported by perceptual evidence. Shinya (2005), for example, showed that generally speaking, native speakers are not conscious of the downstep effect in either a syntagmatic or paradigmatic sense. In particular, they are not conscious of the pitch difference between the first and second phrases in the utterance with downstep (solid line in Figure 1) or between the downstepped second phrase (solid line) and the non-downstepped one (dashed line).

2.3 Major and Minor Phrases

Most theoretical studies of Japanese intonation today assume two prosodic levels under the topmost ‘Utterance’, which is only vaguely defined in the literature. The level just under the utterance is the level of ‘intermediate phrase’ (P&B 1988) or ‘major phrase’ (Poser 1984, Kubozono 1988, Selkirk & Tateishi 1991). Thus, each utterance is supposed to consist of one or more major phrases. The prosodic level that is further lower is that of ‘accentual phrase’ or ‘minor phrase’: each major phrase consists of one or more minor phrases.

The minor phrase is generally defined as the domain of initial pitch rise as well as the domain within which at most one lexical accent can be phonetically realized. A sequence of accented words/phrases usually display a staircase like the solid line in Figure 1, where each syntactic phrase (often referred to in Japanese as *bunsetsu*) constitutes one minor phrase with an independent initial pitch rise and accentual fall (if it contains a lexically accented word). On the other hand, a sequence of two syntactic phrases is often amalgamated into one

minor phrase if the first phrase does not have a lexical accent (Poser 1984, Kubozono 1988). In this case, the whole sequence shows only one initial pitch rise and at most one accentual fall.

The Major Phrase is generally defined as the domain of downstep, the pitch lowering process described above. If downstep is identified between two adjacent minor phrases as in the solid contour in Figure 1, those phrases belong to one and the same Major Phrase. If, on the other hand, it is blocked between two minor phrases, there is a Major Phrase boundary between them.

In P&B's work, the Major Phrase (or 'intermediate phrase' in their terminology) is also defined as the domain of pitch reset. So if downstep is blocked between two minor phrases, there is a Major Phrase boundary between them, with the pitch being 'reset' at the beginning of the second phrase. We follow P&B to assume that the Major Phrase is the domain of pitch reset as well as downstep.

Selkirk & Tateishi (1991) do adopt the general definition of the Major Phrase as the domain of downstep, but they also take a syntax-driven approach in identifying Major Phrase boundaries. According to their top-down approach, a Major Phrase boundary is inserted at the left edge of every XP. We will see evidence against this approach later (see footnote 3 and section 4.2).

3 Review of past work on focus prosody

With a view to examining the effect of focus on intonation structure, Poser (1984) used the set of four phrases in (2) and compared the peaks of the second adjective *aóí* 'blue'. In (2a) and (2b), no word is emphasized, whereas the adjective *aóí* is emphasized in (2c) and (2d). Emphasized words are capitalized in (2) and the rest of this paper.

- (2)
- | | | |
|----|--|---------------------|
| a. | <i>amai</i> <i>aói</i> <i>kudámono</i> | ‘sweet, blue fruit’ |
| b. | <i>umái</i> <i>aói</i> <i>kudámono</i> | ‘tasty, blue fruit’ |
| c. | <i>amai</i> AÓI <i>kudámono</i> | ‘sweet, BLUE fruit’ |
| d. | <i>umái</i> AÓI <i>kudámono</i> | ‘tasty, BLUE fruit’ |

Poser obtained the following average peak values (Hz) from a single speaker for the three phrases/words comprising the sentences in (2): the first adjective (*umái* or *amai*), the second adjective (*aói*) and the noun (*kudámono*).

(3)

	<i>amai</i> or <i>umái</i>	<i>aói</i>	<i>kudámono</i>
a.	171.3	168.9	159.8
b.	176.7	163.2	135.7
c.	168.4	187.7	151.8
d.	175.9	183.4	133.6

A comparison of these values indicates that emphasized elements are more boosted in pitch than non-emphasized ones. For example, the adjective *aói* is considerably higher in (3c,d) than in (3a,b), respectively. In terms of downstep, a comparison of (3a) and (3b) reveals that the adjective *aói* was significantly lower in (2b) than in (2a) ($T=3.48$, $p<0.005$). This indicates the presence or absence of a lexical accent in the immediately preceding adjective (*umái* vs. *amai*) has exerted an effect on the height of the adjective *aói*; namely, *aói* is downstepped in (2b) due to the presence of a lexical accent in the preceding adjective. Poser then compared the pitch peaks of the same adjective in (2c) and (2d), in both of which the adjective itself is emphasized. He reports that *aói* is lower in (2d) than in (2c) with a difference that is ‘only marginally significant’ ($T=1.98$, $p=0.03$) (Poser 1984: 301).

What Poser’s data suggest is that emphasis does not block downstep although it may weaken the lowering effect as compared with the non-emphasis context. This result is particularly interesting because the downstepped element,

aóí, has a higher peak than its preceding element, *umái*, in (2/3d): 183.4 Hz vs. 175.9 Hz. We will see similar paradoxical cases in the following sections.

While Poser's data suggest that focus on a particular element does not block downstep and, hence, it does not introduce a Major Phrase boundary, P&B (1988) presented quite different data and drew an entirely different conclusion. As mentioned in section 1, they claimed that focus blocks the intonational process of downstep and, hence, starts a new Major Phrase, or what they called 'intermediate phrase'. The sentences they used to make this claim are the following:

- (4) a. *amerika-níwa amai KÉEKI-wa arimásu-ga amai AME-wa arimasén.*
 'In America there are sweet CAKES, but there aren't sweet CANDIES.'
- b. *amerika-níwa umái NINZIN-wa arimásu-ga umái MAMÉ-wa arimasén.*
 'In America there are tasty CARROTS, but there aren't tasty BEANS.'
- c. *amerika-níwa umái KÉEKI-wa arimásu-ga umái AME-wa arimasén.*
 'In America there are tasty CAKES, but there aren't tasty CANDIES.'
- d. *amerika-níwa amai NINZIN-wa arimásu-ga amai MAMÉ-wa arimasén.*
 'In America there are sweet CARROTS, but there aren't sweet BEANS.'

In order to see an effect of focus on downstep, P&B compared (4a) and (4b) with respect to the pitch peaks of the adjective-noun sequences, where *wa* is a particle denoting contrast:

- (5) a. (=4a) *amai AME-wa* 'sweet CANDIES'
 b. (=4b) *umái MAMÉ-wa* 'tasty BEANS'

They demonstrated that these two phrases do not exhibit a typical downstep pattern, a pattern whereby they are clearly separated when the peak of their first word is plotted against the peak of the second word. On the basis of

this experimental result, P&B claimed that the focused element is realized at the same pitch level irrespective of any difference in phonological structure of the preceding material; thus, the pitch contour is ‘reset’ by focus.

Given P&B’s data thus described, one may quite naturally wonder why they did not compare (4a) and (4c) or (4b) and (4d), respectively. These suggested comparisons are given in (6) and (7).

- (6) a. (=4a) amai AMÉ-wa ‘sweet CANDIES’
 b. (=4c) umái AMÉ-wa ‘tasty CANDIES’
- (7) a. (=4b) umái MAMÉ-wa ‘tasty BEANS’
 b. (=4d) amai MAMÉ-wa ‘sweet BEANS’

It is not clear why P&B did not compare the two phrases in these pairs, but this appears to be a drawback in their analysis. The two phrases in (5) differ not only in the accentedness of the first word but in that of the second word as well. A pair of phrases like this cannot be used to examine an effect of downstep since the second word in (5b) differs from the second word in (5a) in more than one way. In terms of accentual boost, the second word in (5b) should bear a higher pitch than its counterpart in (5a) due to the lexical accent it contains, or ‘accentual boost’. If downstep takes place, on the other hand, the second word in (5b) should bear a lower pitch than its counterpart in (5a) because of the lexical accent in its preceding word, *umái*. Given these antagonistic forces that may operate on the second word in (5b), this particular word may well be realized at much the same pitch level as its counterpart in (5a). In other words, the effect of accentual boost may well diminish the effect of downstep in the second word in (5b).

This interpretation is supported by Kubozono’s (1988) experimental data, which contain the following pairs of phrases.

-
- (8) a. umái méron ‘tasty melon’
 b. amai oimo ‘sweet potato’
- (9) a. umái nomímono ‘tasty drink’
 b. amai yamaimo ‘sweet yam (potato)’

When pronounced in a non-contrastive context, the second words in each pair did not exhibit a noticeable pitch difference with respect to their height. However, the phrases in (8a) and (9a) did show a clear effect of downstep when they were compared with the phrases in (10) and (11), respectively: *méron* is realized at a significantly lower pitch level in (8a) than in (10), and *nomímono* is significantly lower in (9a) than in (11).

- (10) amai méron ‘sweet melon’
- (11) amai nomímono ‘sweet drink’

Thus, the second words in (8a) and (9a) are downstepped due to the presence of an accent in the preceding material, but this downstep effect cannot be seen—i.e., it is masked—when they are compared with the second words in (8b) and (9b), respectively. After all, in order to see whether or not there is a downstep effect, one must compare two phrases that contrast minimally with each other. Comparing the two phrases in (5), as P&B did, is equivalent to comparing the two phrases in (8) or those in (9), and will not answer the question of whether downstep has taken place between the two relevant phrases.

4 Experiment

4.1 Method

We used the following frame sentence in our experiment: TOP, ACC and Q stand for topic, accusative and question markers, respectively.³

- (12) *anáta-wa X-de Y-to náni-o mimásita-ka?*
 You- TOP X-in Y-with what-ACC see-PAST-Q
 ‘What did you see with Y in X?’

We put either *aómori* ‘Aomori’ or *oomori* ‘Oomori’ in the X slot and *náoko* ‘Naoko’ or *naomi* ‘Naomi’ in the Y slot. This created the four combinations in (13), which differ in the accentedness of the nouns in the two slots. [AA-Wh] stands for a sequence of two accented phrases plus the Wh word (which is always lexically accented in Tokyo Japanese), whereas [UU-Wh] refers to a sequence of unaccented noun phrases plus the Wh word.

- (13) a. [AA-Wh] ...*aómori-de náoko-to náni-o...*
 ‘What did you see with Naoko in Aomori?’
- b. [AU-Wh] ...*aómori-de naomi-to náni-o...*
 ‘What did you see with Naomi in Aomori?’
- c. [UA-Wh] ...*oomori-de náoko-to náni-o...*
 ‘What did you see with Naoko in Oomori?’
- d. [UU-Wh] ...*oomori-de naomi-to náni-o...*
 ‘What did you see with Naomi in Oomori?’

³ Selkirk & Tateishi’s (1991) top-down approach to intonational structure would predict downstep is blocked between *X-de* and *Y-to* as well as between *Y-to* and *náni-o* in (12) since these positions correspond to the left edge of XPs and, hence, introduce a Major Phrase boundary. This prediction cannot be borne out in our experiment, as we will see in section 4.2.

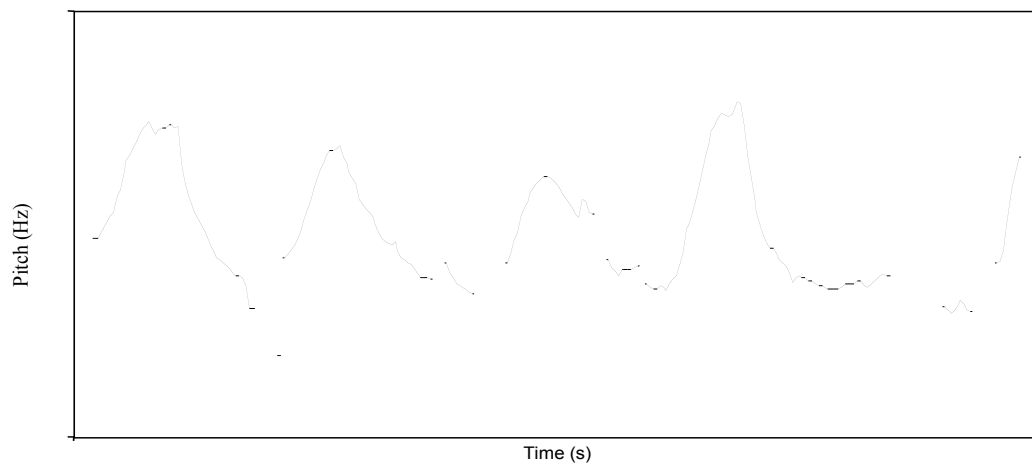
These four test sentences were mixed with dummy sentences in a random way. They were read by seven native speakers of Tokyo Japanese (five male and two female speakers), who pronounced the sentences eleven times at a normal tempo—one repetition per set for a total of eleven sets. This recorded eleven tokens of each test sentence, the first of which was discarded for analysis. A total of 280 tokens were subsequently analyzed (10 tokens x 4 sentences x 7 speakers).

In the analysis, we measured the pitch peak and valley of the three phrases comprising the sentences in (13), with particular focus on the peak of the Wh word. If the Wh element turns out to have the same peak value irrespective of the accentedness of its preceding phrases, then it will mean that downstep has been blocked by the Wh element with the pitch being reset in this position. This will be the result in accordance with the result reported by P&B. If, on the other hand, the pitch is not reset by the Wh element, this element will exhibit different pitch heights among the four test sentences, reflecting the differences in the phonological structure of its preceding materials. In particular, the Wh element will be realized at a considerably lower pitch level when following a sequence of accented phrases as in (13a) than when following a sequence of unaccented phrases as in (13d). If this were the case, it would be a clear case of downstep in the traditional (paradigmatic) sense of the term, indicating that the Wh element does not block this lowering process. This will be the result compatible with Poser's (1984). As we will see in the next section, our experimental data support this second scenario.

4.2 Results

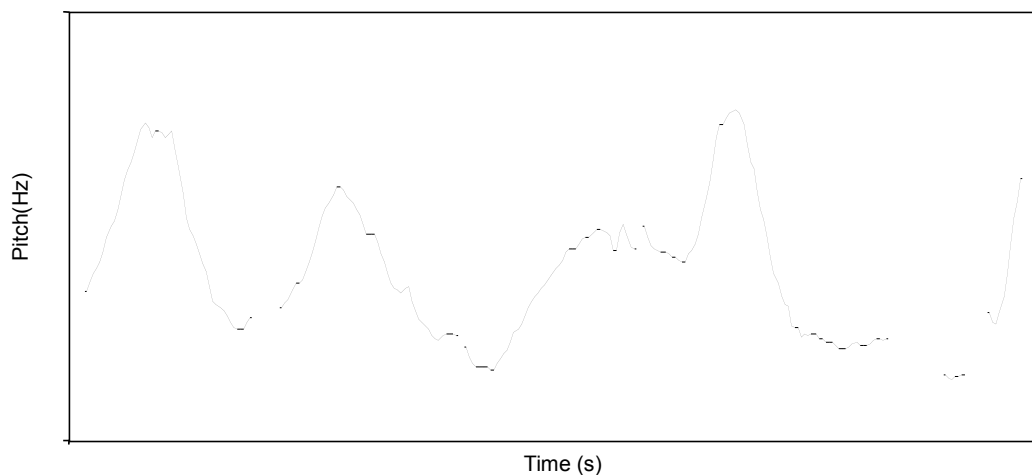
First of all, the Wh element attains a very high pitch level in all the four sentences in (13). In fact, *náni-o* 'what' showed the highest pitch peak in every

one of the 280 tokens that were analyzed, higher than the sentence-initial phrase and much higher than the phrase immediately preceding it. This is illustrated in Figures 2 and 3, which show the typical pitch contours of the sentences in (13a) and (13b), respectively. These contours indicate that the Wh phrase is the very focused element in the test sentences.



anáta-wa aómori-de náoko-to náni-o mimásita-ka?

Figure 2 A typical pitch contour of the sentence in (13a)



anáta-wa aómori-de naomi-to náni-o mimásita-ka?

Figure 3 A typical pitch contour of the sentence in (13b)

More significantly, the Wh element attained different peak levels among the four test sentences in (13). The biggest difference was found between (13a) and (13d), with the former showing a considerably lower peak than the latter. This tendency was shown by all seven speakers: for five of these speakers the difference was statistically significant (by a two-tailed t test), while the other two speakers exhibited a similar tendency although the difference did not reach a level of statistical significance (Table 1).

Table 1 Statistics of the peak F0 values (in Hz) in (13a) and (13d)

Speaker	Peak of Wh in (13a)	Peak of Wh in (13d)	T value	P value
TS (male)	146.2	153.7	2.630	p<0.05
TY (male)	139.7	156.3	2.297	p<0.05
JI (male)	205.9	222.6	3.393	p<0.01
AO (male)	184.3	196.2	4.005	p<0.01
AK (female)	277.8	289.4	2.202	p<0.05
NI (male)	179.4	182.5	0.415	p=0.683
MM (female)	271.9	281.9	1.445	p=0.165

Figure 4 schematizes the overall differences that Speaker AO showed between (13a) and (13d); the average peak and valley values of the three phrases comprising (13a) and (13d) are plotted, respectively.

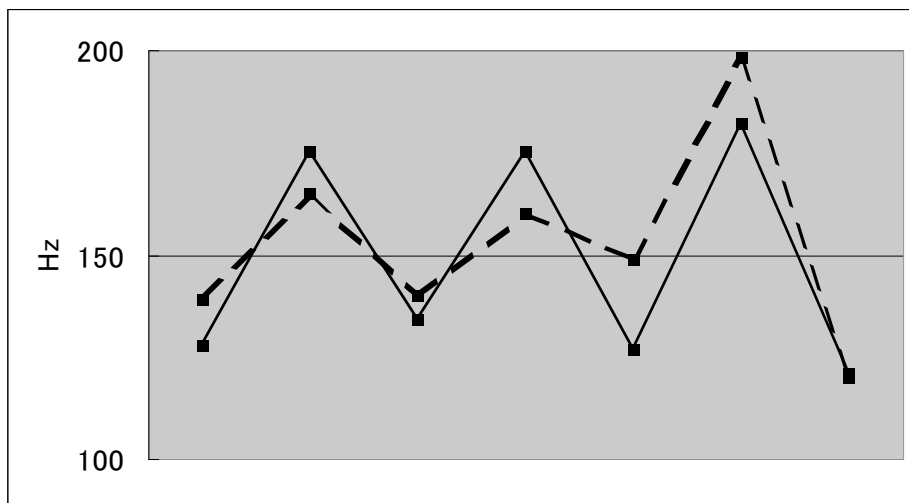


Figure 4 Schematic comparison of the pitch contours [AA-Wh] (solid line) vs. [UU-Wh] (dashed line) [Speaker AO]

The pitch contours in Figure 4 differ from each other in two crucial respects. First, the two phrases before the Wh phrase exhibit a higher peak in [AA-Wh] sequence than in [UU-Wh] sequence: *aómori-de* and *náoko-to* in [AA-Wh] have higher peaks than *oomori-de* and *naomi-to* in [UU-Wh], respectively. This is due largely to the presence or absence of lexical accent, or the effect of accentual boost described in section 2.2 above. A second and more important difference between the two contours in Figure 4 is that the Wh element, *náni-o*, has a considerably lower peak in [AA-Wh] than in [UU-Wh]. This difference was statistically significant for most speakers, as summarized in Table 1.

The overall difference between the two test sentences is clear. [AA-Wh] has a higher peak than [UU-Wh] in the pre-Wh context, but it has a lower pitch in the Wh element itself. This is the same situation that we saw in Figure 1 above, where the two definitions of downstep were explained. In other words, the Wh element in [AA-Wh] is lowered in pitch by the accent(s) of its preceding elements as compared to its counterpart in [UU-Wh]. This clearly shows that

downstep has occurred between the Wh element and its preceding phrase in [AA-Wh] according to the original, paradigmatic definition of the term.

In fact, the only substantial difference between the solid line in Figure 1 and the [AA-Wh] contour in Figure 4 is that in the latter, the downstepped element (*náni-o*) is higher in pitch than its preceding phrase (*náoko-to*) despite that the former phrase has been lowered by the accent of the latter phrase in a paradigmatic sense. In this sense, the [AA-Wh] contour in Figure 4 presents a paradoxical case, a case where the downstepped element (Wh) is realized at a higher pitch level than the element whose accent has triggered the lowering process.

5 Discussion

5.1 Downstep or no downstep?

We are now faced with a puzzling case where the two definitions of downstep—syntagmatic and paradigmatic—do not agree with each other. In the syntagmatic dimension, it appears that the Wh phrase in the [AA-Wh] sentence in Figure 4 is not downstepped as relative to its preceding phrase since it is higher in pitch than the latter. On the other hand, the same Wh element is realized at a lower pitch region in the [AA-Wh] sentence than in its [UU-Wh] counterpart, reflecting the difference in the accentedness of the pre-Wh material. This is a clear case of downstep in the paradigmatic sense of the term.

The question is how we can interpret this puzzling case. One thing that is worth serious attention here is that the paradoxical case in question is not an isolated phenomenon in Japanese prosody. There are at least two independent cases reported in the literature. One case has already been mentioned in section 3, where we discussed Poser's (1984) experimental data. A comparison of the

pitch patterns exhibited by (2c) and (2d) showed that the focused adjective *AÓI* is significantly lower in pitch when preceded by a lexically accented word, *umái*, than when preceded by a lexically unaccented one, *amai*. Namely, the focused adjective is downstepped as relative to its preceding word in (2d). However, this downstepped adjective is realized at a higher pitch level than its preceding word as the values in (3d) clearly show.

A paradoxical case of the same nature has also been reported in my previous experiments (Kubozono 1988, 1989 and 1992). One case concerns the two sentences in (14).

- (14) a. [AAAA]
 [[[náoko-no] [áni-no]] [[aói] [erímaki]]]
 ‘(I saw) Naoko’s brother’s blue muffler’
- b. [AUAA]
 [[[náoko-no] [ane-no]] [[aói] [erímaki]]]
 ‘(I saw) Naoko’s sister’s blue muffler’

These two sentences have an identical syntactic construction, with four phrases constituting a binary branching structure. In phonological terms, they only differ in the accentedness of the second phrases, *áni-no* ‘brother’s’ vs. *ane-no* ‘sister’s’. These two sentences exhibit pitch contours as schematized in Figure 5.

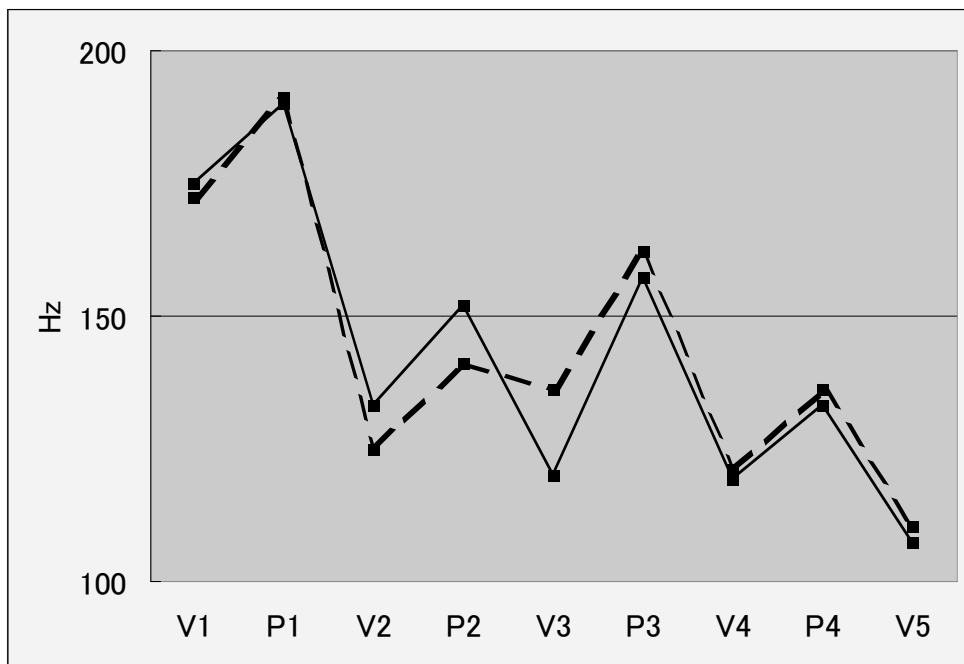


Figure 5 Comparison of the pitch contours shown by the two sentences in (14): (14a) (solid line) vs. (14b) (dashed line).

In every utterance of the two sentences in (14), the third phrase shows a higher peak than the second contour, suggesting that there is some prosodic boundary between the second and third phrases. On the other hand, the third phrase is realized at a lower pitch level when following the accented phrase, *áni-no* ‘brother’s’, than when following the unaccented one, *ane-no* ‘sister’s’. This indicates that the accent of the second phrase has exerted a lowering effect on the third phrase in (14a) although the lowered phrase is realized at a higher pitch level than the phrase whose accent has triggered the lowering process. This is a situation that is identical to the one we saw in Figure 4 above.

What do the paradoxical cases in Figures 4 and 5 tell us? For one thing, they clearly show that the two definitions of downstep—syntagmatic and paradigmatic—do not always agree with each other. More specifically, they suggest that the lowering effect exerted by the accent of a phrase cannot be

identified by comparing the pitch values of two consecutive phrases of a single contour. This speaks against the syntagmatic definition of downstep. In both cases illustrated in Figures 4 and 5, an accented phrase has exerted a lowering effect on the following material in a clear and objective way. This is a clear case of downstep according to the original and more objective definition of the term. There may be some prosodic boundary between the two relevant phrases, i.e., between the Wh phrase and its preceding phrase in Figure 4, and between the second and third phrases in Figure 5. However, this boundary cannot be a Major Phrase boundary. On the contrary, given that downstep is not blocked between the two relevant phrases, it follows that the two phrases do belong to one and the same Major Phrase as long as we adopt the general definition of this prosodic category described in section 2.3 above. Consequently, pitch reset has not taken place between the two phrases.

5.2 Implications

The data presented and analyzed in sections 4 and 5.1 have some important implications for the modeling of Japanese intonation while raising interesting questions for future research. Here we will focus on two issues, one concerning the interface between focus and intonation in general, and the other regarding the hierarchical structure of Japanese intonation.

5.2.1 Two types of focus?

We have seen in section 4 that Wh words do not block downstep and, hence, do not trigger pitch reset. How can we compromise this finding with the existing data about contrastive focus? As mentioned in section 3, Poser (1984) presented data suggesting that contrastive focus fails to block downstep and, hence, to trigger pitch reset. His analysis was supported by Shinya (1999), who also looked at the effect of contrastive focus on downstep. On the other hand, P&B

(1988) claimed that contrastive focus does block downstep, introducing a Major Phrase boundary immediately before the focused element. However, P&B's analysis requires reconsideration, as pointed out in section 3. It seems to follow from these considerations that contrastive focus does not block downstep or trigger pitch reset. This analysis is compatible with our experimental data dealing with non-contrastive focus. This suggests that focus generally fails to block downstep in Japanese, whether it is contrastive or otherwise.

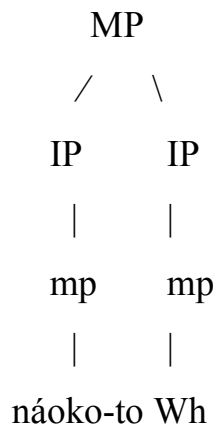
5.2.2 Intonation structure

The data presented in the preceding sections have a significant implication for the prosodic organization of Japanese utterances, too. We have confirmed that Wh elements do not block downstep in the sense that their pitch height is influenced by the accentedness of the material immediately preceding them. On the other hand, Wh elements are realized in a higher pitch region than their preceding phrases. This latter fact suggests that there may be some prosodic boundary immediately before the Wh elements.

Assuming that this interpretation is correct, one may naturally ask what this prosodic boundary is. This cannot be a Major Phrase boundary since, as mentioned above, the Major Phrase is defined as the domain of downstep. As long as downstep is observed between the Wh element and its preceding phrase, there cannot be a Major Phrase boundary between the two elements. The boundary in question cannot be a minor phrase boundary, either, since the Wh element and its preceding phrase clearly form two independent minor phrases, as can be seen from the pitch contours in Figures 2–4. This will raise a challenging question for the intonation model of Japanese since previous studies of Japanese intonation did not posit any intermediate phrase between the Major Phrase and the minor phrase.

One solution to this will be to posit an independent prosodic level/phrase between the two existing prosodic phrases. Let us tentatively call it an ‘intermediate phrase’ (IP), noting that it is different from what P&B called by the same name (by which they meant what we call the Major Phrase here). Restricting ourselves to the string of the Wh phrase and its preceding phrase in Figure 2, this analysis will assume a prosodic hierarchy as illustrated in (15), where the two phrases belong to one and the same Major Phrase, with each phrase constituting an intermediate phrase and a minor phrase on its own.

(15)



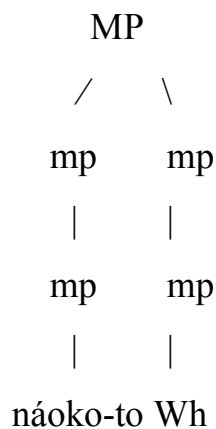
Under this analysis, the Wh phrase undergoes downstep because its preceding phrase is accented and belongs to the same Major Phrase. The Wh phrase can have a higher pitch than its preceding phrase since it belongs to a different intermediate phrase from the latter.

While this appears to be an appealing solution, it falls into a problem regarding the definition of the new phrase. As mentioned in section 2, both the Major Phrase and the minor phrase have been defined in prosodic terms, as the domains of certain prosodic processes. However, the IP, which we have tentatively proposed in (15), does not have such an objective definition since it cannot be defined as the domain of any independent prosodic process. This

seems to be a critical problem if one wants to propose a reasonably constrained model of intonation.

A solution to avoid using this unmotivated prosodic phrase is to appeal to the notion of ‘recursive’ category proposed by Ladd (1996) and supported by Kubozono (1988, 1989, 1992). This analysis allows a certain prosodic phrase/level to occur recursively in the prosodic hierarchy, as illustrated in (16).

(16)



Since this analysis introduces no new prosodic phrase/level, it is free from the kind of problem that the analysis in (15) poses.⁴ Other merits as well as demerits of this analysis need to be explored in detail.

Another solution to the puzzling case in question might be to claim that focus intonation is independent of prosodic phrasing (cf. Ishihara 2005). This approach is certainly incompatible with the assumption generally adopted in the literature, i.e., that focus effects on intonation can be captured in terms of prosodic phrasing. However, this might allow us to solve our puzzle and,

⁴ A reviewer suggests another recursive model of intonation, where the Major Phrase rather than the minor phrase occurs recursively. In this analysis, *náoko-to* and the Wh element in (16) belong to different Major Phrases. This analysis would have to abandon the traditional definition of Major Phrase as the domain of downstep and, hence, require a new definition of this intonational phrase.

moreover, to understand the seemingly complicated interaction between focus and intonation in a reasonably principled way. We would like to leave this issue as a topic for future research.

References

- Akinaga, Kazue. 1985. *Kyootuugo no akusento* (Accent of Standard Japanese). In the appendix to NHK (ed.), *NHK Pronunciation and Accent Dictionary*. Nihon Hoso Shuppan Kyokai, Tokyo.
- Ishihara, Shinichiro. 2003. *Intonation and Interface Conditions*. Ph.D. dissertation, MIT.
- Ishihara, Shinichiro. 2005. Prosody-scope match and mismatch in Tokyo Japanese *Wh*-Questions. *English Linguistics* 22–2: 347–379
- Kubozono, Haruo. 1988. *The Organization of Japanese Prosody*. Ph.D. dissertation, Edinburgh University. [Kurosio Publishers, 1993].
- Kubozono, Haruo. 1989. Syntactic and rhythmic effects on downstep in Japanese. *Phonology* 6–1: 39–67.
- Kubozono, Haruo. 1992. Modeling syntactic effects on downstep in Japanese. In G. J. Docherty and D. R. Ladd (eds.) *Papers in Laboratory Phonology II: Gesture, Segment, Prosody*. 368–387. Cambridge: Cambridge University Press.
- Kubozono, Haruo. 1996. Syllable and accent in Japanese: evidence from loanword accentuation. *The Bulletin* (Phonetic Society of Japan) 211, 71–82.
- Kubozono, Haruo and Misa Fukui. 1996. Phonological structure and unaccented nouns in Tokyo and Osaka Japanese. In Timothy J. Vance and Kimberly Jones (eds.) *Japanese/Korean Linguistics 14*, 39–50. Stanford: CSLI Publications & SLI.
- Ladd, D. Robert. 1986. Intonational phrasing: the case for recursive prosodic structure. *Phonology Yearbook* 3, 311–340.
- Nagahara, Hiroyuki. 1994. *Phonological Phrasing in Japanese*. Ph.D. dissertation, UCLA.

-
- Ogawa, Shinji. 2004. Sino-Japanese word accent and syllable structure. *Phonological Studies* 7, 41–48.
- Pierrehumbert, Janet and Mary Beckman. 1988. *Japanese Tone Structure*. Cambridge, Mass.: MIT Press.
- Poser, William. 1984. The Phonetics and Phonology of Tone and Intonation in Japanese. Ph.D. dissertation, MIT.
- Selkirk, Elizabeth and Koichi Tateishi. 1991. Syntax and downstep in Japanese. In Carol Georgopoulos and Roberta Ishihara (eds.) *Interdisciplinary Approaches to Language*, 519–543. Dordrecht: Kluwer Academic Publishers.
- Shinya, Takahito. 1999. Eigo to nihongo ni okeru fookasu ni yoru daunsuteppu no sosi to tyooon undoo no tyoogoo [The blocking of downstep by focus and articulatory overlap in English and Japanese]. In *Proceedings of Sophia Linguistics Society*, 14, 35–51.
- Shinya, Takahito. 2005. Lexical accent and the perception of intonational peaks in Japanese. Talk presented at the monthly meeting of the Phonological Association in Kansai (PAIK), June 25, 2005.
- Truckenbrodt, Hubert. 1995. Phonological Phrases: Their Relation to Syntax, Focus and Prominence. Ph.D. dissertation, MIT.
- Zubizarreta, Maria. 1998. Prosody, *Focus and Word Order*. Cambridge, Massachusetts, The MIT Press.

Haruo Kubozono
Department of Linguistics
Faculty of Letters, Kobe University
Rokkodai-cho, Nada-ku
Kobe 657-8501
Japan
kubozno@lit.kobe-u.ac.jp

When We Fail to Question in Japanese*

Yoshihisa Kitagawa

Indiana University

When we pay close attention to the prosody of Wh-questions in Japanese, we discover many novel and interesting empirical puzzles that would require us to devise a much finer syntactic component of grammar. This paper addresses the issues that pose some problems to such an elaborated grammar, and offers solutions, making an appeal to the information structure and sentence processing involved in the interpretation of interrogative and focus constructions.

Keywords: focus, (implicit) prosody, information structure, processing, Wh-question

1 Background — Some Recent Development in Formal Syntax

In this paper, I will take up some Wh-constructions in Japanese which do not seem to pose any grammatical problem but disallow us to obtain certain type of expected interrogative interpretations. In Section 1, I will summarize some recent development of a research method incorporating prosodic and other extra-syntactic/extra-grammatical analyses into the formal study of syntax. In Section 2, I will sketch out the elaborated version of "*LF E-agreement*" proposed and argued for by Kitagawa (2006). In Section 3, I will investigate

* I am grateful to the participants of WAFL 2, my colloquium presentation in Hiroshima Gengogaku Danwakai and syntax seminar at Indiana University, especially to Leslie Gabriel, Shin Ishihara and Satoshi Tomioka, for their comments. The research presented in this work has been partially supported by COAS Grants-in-Aid and Faculty Research Incentive Fund at Indiana University. This material is also based upon work supported by the National Science Foundation under Grant No. BCS-0650415.

into the nature of the puzzling phenomenon mentioned above, which would permit us to account for what is left unattended in the *LF E-agreement* approach.

1.1 Pivotal Observations

Deguchi and Kitagawa (2002) and Ishihara (2002), among others, converged on the view that we must examine prosody in order to understand the semantic and formal properties of Wh-questions in Japanese more precisely. They pointed out that Wh-questions in Japanese must be generally accompanied by "Emphatic Prosody (EPD)" (or "Focus Intonation (FI)" in Ishihara's terminology) as in (1a) below. EPD consists of, first, an *emphatic accent* on the Wh-focus, which consists of sharp rise of F_0 (indicated by **BOLD CAPITALS**) followed by its fall, and second, *post-focal reduction*, which virtually (though not entirely) suppresses all lexical accents up to the end of some clause by compressing their pitch and amplitude range (indicated by **shade**). Independently of EPD, interrogative rise intonation (indicated by \uparrow) is added at the end of an utterance in the matrix Wh-question, which terminates post-focal reduction.

- (1) a. **DA**re-ga itumo ohiru-ni piza-o taberu-**no** \uparrow ?
 who-NOM always lunch-for pizza-ACC eat-COMP_{Wh}
 'Who always eats pizza for lunch?'
- b. # **da**re-ga **i**tumo o **hi**ru-ni **pi**za-o taberu-no \uparrow ?
 who-NOM
- c. **Jo**hn-wa **i**tumo o **hi**ru-ni **pi**za-o taberu.
 John-TOP
 'John always eats pizza for lunch.'

Need for EPD in (1a) can be demonstrated when we observe that the same Wh-question sounds unnatural when it is pronounced without EPD as in (1b), with the lexical accent of the head of each phrase retained (as indicated by a

Ⓒircle) and downstepped. (# on this and other examples indicates that the sentence is unacceptable with the indicated prosody (and interpretation).) This non-emphatic prosody, on the contrary, is perfectly natural in a declarative sentence as in (1c), which further indicates the close association between EPD and Wh-questions.¹

They then pointed out that the domain of EPD coincides with the scope domain of Wh — the [+Wh] CP at which EPD ends corresponds to the scope domain of a Wh-phrase. Therefore, when a Wh-question is accompanied by *Local EPD*, which ends at the subordinate COMP as in (2) below, subordinate Wh-scope is obtained and the sentence is interpreted as containing an indirect Wh-question.²

- (2) Hokenzyo-wa [syokutyuudoku-kanzya-zen'in-ga
 health.department-TOP food.poisoning-victim-all-NOM
 NAni-o tabeta-ka] ma da kakunin-dekinai-no↑?
 what-ACC ate-COMP_{Wh} yet confirm-cannot-COMP_{Y/N}
 'Is the Department of Health yet to be able to confirm
 [**what** all of those who suffered from food poisoning ate]?'

Crucially, post-focal reduction in this sentence continues only up to the subordinate COMP, as the retention of the H tone in the matrix (ma da 'yet')

¹ Some recordings of EPDs can be heard by visiting "<http://www.iub.edu/~ykling/SoundGallery/index.html>". See Kitagawa (2005) for further arguments that EPD is a normal rather than exceptional prosodic pattern to be assigned to Wh-constructions in Japanese, contra Nishigauchi (1990). See also Maekawa (1991) for some phonetic experiments that support this point. The H tones involved in the unaccented words also undergo post-focal reduction, though I will not take them up for simplicity.

² In glosses of this and other examples, I will indicate each distinct function of complementizers in Japanese as COMP_{Wh} (Wh-scope maker), COMP_{Whr} (a polar-question complementizer), COMP_{Y/N} (yes/no question marker) or COMP_{That} (declarative complementizer).

indicates. When the same Wh-question is accompanied by *Global EPD* as in (3) below, on the other hand, matrix Wh-scope is obtained and the sentence is interpreted as a direct Wh-question. Note that post-focal reduction is extended to the matrix COMP in this case.^{3, 4}

- (3) Hokenzyo-wa [syokutyuudoku-kanzya-zen'in-ga
 health.department-TOP food.poisoning-victim-all-NOM
 NAni-o tabeta-ka] kakunin-siyoo-tositeiru-no↑?
 what-ACC ate-COMP_{whr} trying.to.confirm-COMP_{wh}
 'What₁ is such that the Department of Health is trying to confirm
 [whether all of those who suffered from food poisoning ate it₁]?'

Deguchi and Kitagawa (2002) also point out that multiple Wh-questions in Japanese exhibit their prosody-scope correlation in a very specific way — in the form of the correspondence between what we may call *Compound EPD*, in which more than one EPD ends at the same COMP and makes up a unit. As a result, more than one Wh-phrase takes synchronized scope and yields a "pair-wise" (or "set") interpretation. In (4a), Complex EPD ends at the subordinate COMP and both Wh-phrases must take subordinate scope, while in (4b-c), Compound EPD is extended to the matrix COMP and both Wh-phrases must

³ At least the seed of these observations can be found also in Tomioka (1997) on Japanese and Lee (1982) and Choe (1985) on Korean. Kubo (2001) also reports similar but somewhat different prosody-scope correlation in Wh-questions in the Fukuoka dialect of Japanese. Hirotani (2003) and Hirotani (2004), on the other hand, report that a sizable number of speakers in her perception experiment could interpret Wh-questions accompanied by *Global EPD* as indirect questions. Many of the example sentences used in her experiments, however, are biased, involving semantics and pragmatics that strongly encourage indirect question interpretations. See Kitagawa and Fodor (2003) and especially Kitagawa (2005) for the description of other factors that bias language users toward subordinate Wh-scope in this construction.

⁴ In *Local EPD*, the subordinate COMP also tends to be (though not necessarily) followed by a short pause while *Global EPD* is not. *Local EPD* and *Global EPD* were also called Short EPD and Long EPD, respectively, in Deguchi and Kitagawa (2002), Kitagawa and Deguchi (2002).

take matrix scope. The two instances of EPD terminating at the identical Comp in Compound EPD is indicated by an underscore and an overscore.

(4)

- a. Keesatu-wa [ano-ban DAre-ga DAre-to atteita-ka]
 police-TOP that-night who-NOM who-WITH seeing-COMP_{Wh}
 miN(na)-ni tazuneta-no↑?
 everyone-DAT asked-COMP_{Y/N}
 'Did the police ask everyone [**who** was with **whom** that night]?'
- b. Keesatu-wa [ano-ban Mary-ga
 police-TOP that-night -NOM
DAre-to atteitta-ka] DAre-ni tazuneta-no↑?
 who-WITH seeing-COMP_{Wthr} who-DAT asked-COMP_{Wh}
 '**Who**₁ is such that the police asked **whom** *whether* Mary was with
 him₁ that night?'
- c. Keesatu-wa [ano-ban
 police-TOP that-night
DAre-ga DAre-to atteitta-ka] kimi-ni tazuneta-no↑?
 who-NOM who-WITH seeing-COMP_{Wthr} you-DAT asked-COMP_{Wh}
 '**Who**₁ is such that the police asked you *whether* he₁ was with **whom**
 that night?'

1.2 Initial Grammaticalization

With these factual observations, we are now given the following mission. First, we must let the grammar of Japanese guarantee prosody-scope correlation in Wh-questions in one way or another. Second, the grammaticality of (3) and (4c) suggests that Wh-questions in Japanese are not constrained by the Subjacency Condition. If we maintain that this condition is imposed on movement, we are now obliged to determine scope of interrogative Wh-phrases in Japanese independent of movement, overt or covert.

Kitagawa and Deguchi (2002) proposed what they call "E-agreement" approach to fulfill both of these tasks at the same time. A remodeled version of

this analysis now postulates what is called "E-feature complex" of the form (E_{SEM} , E_{PHON}). This formal feature complex consists of an E-feature relevant to LF (E_{SEM}) and that relevant to PF (E_{PHON}), which are introduced under both COMP and a Wh-word (or any word that is focalized). We may consider that the E-feature complex introduced under COMP is uninterpretable while that introduced under a Wh-word is interpretable. Under Chomsky's Spell-Out analysis, only E_{PHON} would be stripped from the syntactic object and sent to PF, while E_{SEM} would be maintained through narrow syntax and the semantic component, and sent to LF. The E-feature complex induces the computational operation E-agreement between COMP and a Wh-word in the course of derivation to **both LF and PF**, and eventually uninterpretable E-features get deleted.⁵ When E-agreement takes place successfully in LF-computation under a c-command relation, it comes to identify a word containing E_{SEM} as the focus and the maximal projection of the COMP containing E_{SEM} as the domain of focus. As a result, E-agreement establishes at LF a domain for Wh-scope to be assigned at the Conceptual-Intentional (C-I) system. Successful E-agreement in PF-computation, on the other hand, identifies, in a linear fashion, the lexical item carrying E_{PHON} as the starting point of focus prosody and the COMP containing E_{PHON} as its endpoint. A prosodic domain marked this way comes to be phonetically interpreted as EPD at the Articulatory-Perceptual (A-P) system.⁶

⁵ Alternatively, we may consider that the E-feature complex on both of COMP and a Wh-word is uninterpretable and there exists asymmetrical assignment of some values between them just as in Case features. Postulation of a property that derives both semantic and phonetic effects can be traced back at least to the focus marker "F" of Jackendoff (1972: 240).

⁶ At this point, it is not clear if the E-agreement need to be translated into any hierarchical phonological phrasing which mediates syntax and phonetic interpretation of EPD, at least in the way proposed in the literature. Kubozono (2007: this volume) in fact reports his experimental results which indicate that Wh-focus does not reset the pitch range, which suggests that there exists no major phrase (or intermediate phrase) boundary starting there.

The two different cases of prosody-scope correlation observed in a "Wh-in-situ" sentence in (2) and (3) can be straightforwardly captured when we assume that an E-feature complex may appear in either subordinate or matrix COMP and undergo E-agreement in the course of both LF- and PF-computation, as illustrated in (5) and (6), respectively.

- (5) a. *Subordinate Scope*
- LF: [CP... [CP... **nani**₁[E]-o tabeta]-ka_{COMP}[E]] mada kakunin-dekinai-no_{COMP}]
 what-ACC
- b. PF: [CP... [CP... **NAni**[E]-o tabeta-ka_{COMP}[E]] (ma) da kakunin-dekinai-no_{COMP}↑]
Local-EPD
- (6) a. *Matrix Scope*
- LF: [CP... [CP... **nani**₁[E]-o tabeta-ka_{COMP}] kakunin-siyoo-tositeiru-**no**_{COMP}[E]]
 what-ACC
- b. PF: [CP... [CP... **NAni**[E]-o tabeta-ka_{COMP}] kakunin-siyoo-tositeiru-**no**_{COMP}[E]↑]
Global-EPD

In short, an E-feature complex (E_{SEM} , E_{PHON}) induces the computational operation E-agreement between COMP and a Wh-word simultaneously in LF- and PF-computation and yields a one-to-one grammatical correspondence between the domain of Wh-scope and EPD.

2 Elaborated Grammaticalization

Kitagawa (2006) elaborated on the computational process of LF E-agreement and claimed that the semantic E-feature (E_{SEM}) itself is complex and heterogeneous in nature and can be associated with any of *focused* Wh-phrases, *unfocused* Wh-phrases and *non-Wh focus* phrases, when it consists of a distinct combination of semantic features. First, a "Case-sensitive" labeling of distinct types of phrases and notions as summarized in (7) was established.

- (7) a. **WH**-P (to be referred to as "Big Wh-P") = *Focus Wh*-phrase
 b. **wh**-P (to be referred to as "Small Wh-P") = *Non-focus Wh*-phrase
 c. **FP** = *Non-Wh focus* phrase (both presentational and contrastive)
 d. **Wh** = Reference to Wh- in general as in "Wh-question, Wh-phrase, Wh-in-situ"

Then it was proposed that the scope of *WH*-Ps, *wh*-Ps and *FP*s is determined when each of them is associated with a specific formal feature of COMP as summarized in (8).

- (8) a. COMP [*wh*]: COMP with a **wh**-feature (= an interrogative feature) is *unselectively* associated with one or more **wh**-Ps.
 b. COMP [*F*]: COMP with an **F**-feature (= an emphatic feature) is associated with an **FP**.
 c. COMP [*WH*]: COMP with a **WH**-feature (= an interrogative and emphatic feature), is *unselectively* associated with one or more **WH**-Ps.

The association of COMPs with *wh*-Ps, *FP*s and *WH*-Ps takes place by means of *LF E-agreement* involving their shared feature [*wh*], [*F*], or [*WH*] (along with

PF E-agreement).⁷ I also assume that some economy principle prohibits any COMP feature from being redundantly introduced into a syntactic representation.⁸ Note that a *wh*-feature is characterized by its interrogative property and an *F*-feature by its emphatic property, and crucially, a *WH*-feature is regarded as a hybrid feature which has both interrogative and emphatic properties. This means that the introduction of the features [*WH*] and [*wh*] to a single COMP is prohibited due to the redundancy of an interrogative property. Likewise, the features [*WH*] and [*F*] cannot be introduced simultaneously to a single COMP because of the redundancy of an emphatic property. It was argued that this approach would permit us to discover and explain some puzzling scope phenomena, which would otherwise have remained unaccounted for or even unnoticed.

2.1 Novel Puzzles and Solutions #1

First, a paradigm involving multiple Wh-questions as in (9) was presented. Since prosody plays an essential role in the examples examined here, the readers must assign the prosodic pattern indicated on each example in interpreting them.

- (9) a. DAre-ga asoko-de NAni-o katta-no↑?
 who-NOM there what-ACC bought-COMP_{wh}
 'WHO bought WHAT there?'

- b. # DAre-ga asoko-de nani-o katta-no↑?

⁷ I tentatively assume that this LF-association yields Reinhart (1997)'s "choice function" as its semantic consequence. It is not clear to me if more than one *FP* may be also unselectively associated with a single COMP [*F*].

⁸ We may consider that this is a specific instance of the economy on lexical information argued for in Kitagawa (1999), which requires the grammar to examine and evaluate a reference set at LF in terms of the amount of lexical information involved in the representation.

c.

John-wa [**DAre-ga** asokode *nani-o* katta-*ka*] siritagatteiru-*no*↑?
 John-TOP who-NOM there what-ACC bought-COMP_{Wh} want.to.know-COMP_{Wh}
 'WHO₁ is such that John wants to know [*what* he₁ bought there]?'

As in (9a), the two Wh-phrases accompanied by Compound EPD exhibit synchronized scope and the sentence is grammatical with the resulting "paired foci" interpretation. When one of the Wh-phrases (*nani* 'what') fails to be focused in the same sentence as in (9b), on the other hand, the two Wh-phrases cannot synchronize in scope. The sentence in fact fails to provide any legitimate Wh-question interpretation, and is ungrammatical as a multiple or any other non-echo Wh-question. If (9b) is ever accepted, it must be interpreted as an echo question used in a dialogue between two speakers A and B (or A and B') as in (10).

(10) A: John-wa asoko-de **NA**n_i-o katta-**no**↑?
 John-TOP there what-ACC bought-COMP_{Wh}
 'What did John buy there?'

B: E? **DAre-ga** asoko-de *nani-o* katta-*ka-tte*↑?
 Huh who-NOM there what-ACC bought-COMP_{Wh}-COMP_{That}
 'Huh? What did who buy there?'

B':% E? **DAre-ga** asoko-de *nani-o* katta-*no*↑?
 Huh who-NOM there what-ACC bought-COMP_{Wh}
 'Huh? What did who buy there?'

Some speakers find both (10B) and (10B') possible as an echo question while others accept only (10B). Whichever may be accepted, the sentence would be answered with something like *John(-desu-yo)* '(It is) John,' which provides the identity of only the focused Wh-phrase *Dare-ga* 'who-NOM'. The interrogative

interpretation of the unfocused Wh-phrase *nani-o* 'what-ACC', in other words, must be suppressed.

A sentence like (9b) becomes grammatical, however, when we embed it in another Wh-interrogative clause as in (9c) and assign the scope interpretation of the two Wh-phrases as indicated there. One obvious difference between (9b) and (9c) is that the latter permits a focused Wh-phrase and a non-focused one each to take scope in a distinct CP while the former does not have any room for this option. Their contrast therefore suggests that a focused Wh-phrase (*WH*-P) and a non-focused Wh-phrase (*wh*-P) cannot synchronize in their scope even when they are located in the same CP. This phenomenon was referred to as *anti-scope-synchronization* between a *WH*-P and a *wh*-P.⁹

This anti-scope-synchronization phenomenon follows directly from the elaborated version of LF E-agreement introduced above. First, (9a) and (9b) come to involve COMP-Wh association as follows.

- (9) a. DAre-ga asoko-de NAni-o katta-no↑?
 who-NOM what-ACC -COMP_[WH]
 |.....|.....|
 'Who bought what there?'
 b. # DAre-ga asoko-de nani-o katta-no↑?
 who-NOM what-ACC -*COMP_{[WH][wh]}
 |.....|.....|

The contrast here arises because the unselective association of two *WH*-Ps with COMP [*WH*] as in (9a) is legitimate while association of a *WH*-P and a *wh*-P

⁹ One possibility is that the echo questions (10B/B') involve some abbreviated version of embedding and hence are assimilated to the case like (9c). The appearance of *-ka* 'COMP_{wh}' in (10B) is suggestive of this possibility.

with a single COMP as in (9b) would require the associated COMP to have both [*WH*] and [*wh*], which is illegitimate because of the redundancy of an interrogative property.

A *WH*-P and a *wh*-P can still co-occur, on the other hand, when they take distinct scope as in (9c).

- (9) c.
- | | | | | | | | |
|---------|---|---------|---------|----------------|-----------------------|---------------------------|----|
| John-wa | [| DAre-ga | asokode | <i>nani</i> -o | katta- <i>ka</i>] | siritagatteiru- <i>no</i> | ↑? |
| who-NOM | | | | what-ACC | -COMP _[wh] | -COMP _[WH] | |
| | | | | | | | |
- 'WHO₁ is such that John wants to know [*what* *he*₁ bought there]?'

We also predict this phenomenon, since each Wh-phrase (*WH*-P and *wh*-P) is legitimately associated with an appropriate type of COMP (COMP [*WH*] and COMP [*wh*], respectively) in its own clause. When we replace the subordinate COMP with a declarative COMP *-to* and use an appropriate predicate in the matrix as in (11) below, on the other hand, both *WH*-P and *wh*-P would be forced to be associated with the matrix COMP [*WH*][*wh*], which again is prohibited, and the sentence becomes uninterpretable.

- (11)
- | | | | | | | | | |
|-----------|---|---------|---------|----------------|-----------------------|-----------|--------------------------|----|
| # John-wa | [| DAre-ga | asokode | <i>nani</i> -o | katta- <i>to</i>] | omotteiru | - <i>no</i> | ↑? |
| who-NOM | | | | what-ACC | -COMP _{That} | think-* | COMP _{[WH][wh]} | |
| | | | | | | | | |
- 'WHO₁ is such that John still thinks [that (s)*he*₁ bought *what* there]?'

2.2 Novel Puzzles and Solutions #2

The "elaborated" LF E-agreement approach leads us to discover and solve another interesting interpretive puzzle when we extend our investigation from

Wh-focus to *non-Wh*-focus. We observe first that a non-*Wh*-focus phrase in Japanese is also generally accompanied by EPD, whether it is a presentational focus as in (12a-c) or a contrastive focus as in (13a-c).

- (12) a. **JO**hn-ga Mary-ni kagi-o watasimasita.
 -NOM -DATkey-ACC handed
 'It is John who handed a key to Mary.'
- b. John-wa **MA**ry-ni kagi-o watasimasita.
 -TOP
 'It is Mary to whom John handed a key.'
- c. John-wa Mary-ni ka**GI**-o watasimasita.
 'It is a key that John handed to Mary.'
- (13) a. **JO**hn-wa Mary-ni kagi-o watasimasita.
 -CONT(RASTIVE)
 'At least John handed a key to Mary.'
- b. John-wa **MA**ry-ni-wa kagi-o watasimasita.
 -TOP -DAT-CONT
 'John handed a key at least to Mary.'
- c. John-wa Mary-ni ka**GI**-wa watasimasita.
 -TOP -CONT
 'John handed at least a key to Mary.'

An interesting contrast arises when we introduce both *Wh*-focus and non-*Wh*-focus into a single sentence as in (14a-b).

(14) [Someone talking about a professional baseball team says:]

- a. Oonaa-wa [**DAre-ga tugi-no kantoku-ni-naru-ka**]
 owner-TOP who-NOM next manager.become-COMP_{Wh}
 // **SEnsyutati-ni osienakatta-no**↑? Sorya mazuine.
 players-DAT not.informed-COMP_{YN} that.is.unadvisable
 'Was it to the players that the owner did not inform who would be the
 next manager? That is inadvisable.'

- b. # Oonaa-wa [**DAre-ga tugi-no kantoku-ni-naru-to**]
 -COMP_{That}
 (//) **SEnsyutati-ni osienakatta-no**↑
 -COMP_{Wh}

(14a) contains a subordinate CP headed by an interrogative COMP *-ka*. When separate EPD is assigned to focus in each clause, the sentence is interpretable, presumably with each focus taking scope in a distinct CP. (// in (14a) indicates a little pause inserted to separate the two instances of EPD.) (14b), on the other hand, is quite awkward when two separate *Local* EPDs same as in (14a) are assigned, which would require the Wh-focus to be associated with the declarative COMP *-to* in the subordinate CP. What is puzzling is that (14b) still cannot be interpreted in any legitimate way even when it is assigned a (single) *Compound EPD* as indicated there, which should have permitted the subordinate Wh-focus to be successfully associated with the interrogative COMP in the matrix CP. This observation suggests that Wh-focus (*WH-P*) and non-Wh-focus (*FP*) are not interpretable when they are forced to take scope under the same CP. This phenomenon was again referred to as *anti-scope-synchronization* but this time involving a *WH-P* and a *FP*.

This contrast also follows under the "elaborated" LF E-agreement analysis:

(14) a.

Oonaa-wa [**DA**re-ga tugi-no kantoku-ni-naru-**ka**]
 who-NOM -COMP_[WH]
 // **SE**nsyutati-ni osienakatta-**no** ↑?
 players-DAT -COMP_[F]

'Was it to the players that the owner did not inform who would be the next manager? That is unadvisable.'

b.

Oonaa-wa [**DA**re-ga tugi-no kantoku-ni-naru-**to**] **SE**nsyutati-ni osienakatta-**no** ↑?
 who-NOM -COMP_{That} players-DAT -*COMP_{[WH][F]}

Since the *WH*-P and *FP* in (14a) can be associated with COMP [*WH*] and COMP [*F*], respectively in two distinct clauses, the sentence is legitimately interpreted. In (14b), on the other hand, the declarative COMP in the subordinate CP forces the *WH*-P and *FP* to be associated with a single, illegitimate COMP with both [*WH*] and [*F*] in the matrix CP. For the same reason, (14a) would not permit the matrix scope interpretation of the *WH*-P even when *Global EPD* is assigned and the subordinate COMP *-ka* is to be interpreted as *whether*.

The anti-scope-synchronization effects involving *WH*-Ps, *wh*-Ps, and *FP*s thus follow from the "elaborated" LF E-agreement, whose analyses are summarized in (15).

(15) a. Permitted:

Type of COMP	COMP [<i>wh</i>]	COMP [<i>F</i>]	COMP [<i>WH</i>]	COMP [<i>F</i>][<i>wh</i>]
Feature properties	interrogative	emphatic	interrogative & emphatic	interrogative & emphatic
Associated phrases	<i>wh</i> -P	<i>FP</i>	<i>WH</i> -P	<i>FP</i> & <i>wh</i> -P

b. Prohibited:

Type of COMP	*COMP [<i>F</i>] [<i>WH</i>]	*COMP [<i>WH</i>][<i>wh</i>]
Feature properties	emphatic & [emphatic & interrogative]	[emphatic & interrogative] & interrogative
Associated phrases	* <i>FP</i> & <i>WH</i> -P	* <i>WH</i> -P & <i>wh</i> -P

Note that, as described in the last column in (15a), the proposed system also predicts that the features [*F*] and [*wh*] can be simultaneously introduced under a single COMP and be associated with an *FP* and a *wh*-P at the same time since no conflict or redundancy should arise in this situation. Such a construction indeed seems to be possible and a sentence like (16) can be properly interpreted, in which a contrast phrase *JOhn-wa* as an *FP* and a *wh*-P may co-occur.

(16)

Zyaa	JOhn-wa	<i>nani</i> -o	eranda	-no↑?
then	-CONT	what-ACC	selected-COMP	[<i>F</i>][<i>wh</i>]

'Then, what did **JOHN** select?'

To sum up, the "elaborated" LF E-agreement approach offers much finer and precise feature analysis of complementizers, Wh-words and other focused items in the grammar. By requiring us to pay close attention to the prosody and information structure of these syntactic elements, this approach shows us a simple way to account for various scope restrictions imposed on them. Without this approach, we would not have even noticed the existence of such empirical problems.¹⁰ This in turn makes us realize that we have long been trying to build a syntactic theory of Wh-constructions in Japanese based upon quite limited empirical observations.

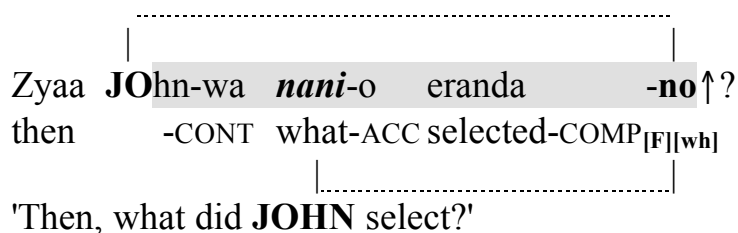
3 Unexpected Restrictions:

When we proceed with the "elaborated" LF E-agreement approach further extending our observations, we encounter even more puzzles. They first appear to contradict with this approach but come to reveal themselves to involve problems that go beyond grammar when we examine them closely. It will be pointed out that the problems involve two distinct extra-grammatical factors — information structure and sentence processing. I will examine them in turn.

3.1 Informational Dead End

First, we have seen at the end of the previous section that an *FP* and a *wh*-P can co-occur and their scope may synchronize as in (17) (= (16)).

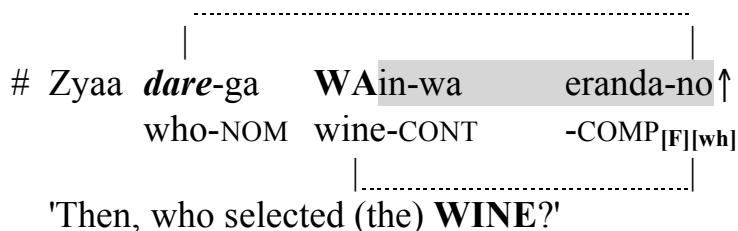
(17)



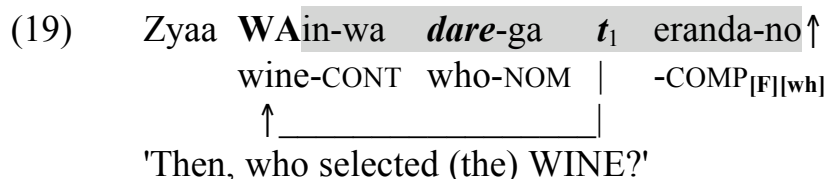
¹⁰ See Kitagawa (2006) for further arguments for this approach.

When an *FP* and a *wh*-P appear in the opposite order in the same construction and pronounced accordingly as in (18), however, the sentence is noticeably awkward.

(18)



As indicated in each example, the association between COMP and *wh*-P/*FP* are identical and legitimate in both cases, involving *[F]* and *[wh]*. The computational process of LF E-agreement, in other words, is successfully carried out and no other grammatical problem appears to arise in (18). An interesting and appropriate observation here is that this sentence becomes acceptable again when scrambling reorders the *FP* and *wh*-P as in (19).



When we compare (17), (18) and (19), we notice that no problem arises when a *wh*-P appears between an *FP* and its associated COMP while the sentence becomes awkward when a *wh*-P fails to do so. A *wh*-P, in other words, is parasitic to the E-agreement domain of an *FP*, so to speak. Let me capture this observation as a theorematic requirement imposed on *wh*-Ps as in (20).

(20) The *wh*-P Theorem:

A *wh*-P is parasitic to the LF E-agreement involving a focus property (= [*F*] or [*WH*]) in that the former must be provided its interpretation within the domain of the latter.

It is not clear at this point how exactly this theorem should be derived but it makes perfect sense when we consider the information structure involved in the paradigm (17)-(19). First, as Vallduvi (1992) and Vallduvi (1995) suggested, information packaging of an utterance involves its three primitives typically appearing in the order indicated in (21).

(21) Link [**FOCUS** tail]

Roughly speaking, *link* represents part of the background information that is discoursally or pragmatically anaphoric and hence signals the connection point of background and prominent information.¹¹ *Focus* introduces the most prominent piece of information to be conveyed and *tail* represents a truly non-salient background context into which this information is introduced. *Focus* and *tail* make up a unit that corresponds to our LF E-agreement domain, which is also prosodically realized as EPD, as indicated in (21) with our notation. In (18), the *wh*-P appears as if it were to serve as *link*, but as Tomioka (2004) points out, a Wh-word by its very nature is incapable of being anaphoric and hence cannot serve as *link* (being an "Anti-Topic Item" in Tomioka's terms). The information structure involved in (18) therefore is inappropriate. As a result, the sentence becomes awkward even if it involves legitimate LF E-agreement and is grammatical. This is a likely source of the *wh*-P Theorem in (20). When we

¹¹ *Link* itself can be also pragmatically accommodated. cf. Heycock (1994)

replace the *wh*-P in (18) with an item that can serve as a link as in (22), the sentence comes to be interpretable again with the same prosodic pattern.

- (22) Zyaa John-wa **WA**in-wa eranda-no ↑
 John-TOP wine-CONT -COMP_[F]
 'Then, did John select (the) **WINE**?'

The *wh*-P *Theorem* as stated in (20) can also account for another scope restriction. First, with example (9c), we have seen above that a *WH*-P and a *wh*-P can co-occur as long as they are associated with distinct COMPs and take distinct scope.

- (9) c.
- | | | | | | |
|---|----------|---------|----------------|-----------------------|------------------------------|
| John-wa | [DAre-ga | asokode | <i>nani</i> -o | katta- <i>ka</i>] | siritagatteiru- <i>no</i> ↑? |
| who-NOM | | | what-ACC | -COMP _[wh] | -COMP _[WH] |
| | | | | | |
| 'WHO ₁ is such that John wants to know [<i>what</i> <i>he</i> ₁ bought there]?' | | | | | |

What is puzzling in this regard is that the opposite scope relation of the same *WH*-P and *wh*-P as indicated in (23) is not permitted when we assign *Local* EPD — we can never let the *WH*-P take the subordinate scope and the *wh*-P take the matrix scope here.

- (23)
- | | | | | | |
|---|----------|---------|----------------|-----------------------|---------------------------------|
| # John-wa | [DAre-ga | asokode | <i>nani</i> -o | katta- <i>ka</i>] | siRITA (ga)tteiru- <i>no</i> ↑? |
| who-NOM | | | what-ACC | -COMP _[WH] | -COMP _[wh] |
| | | | | | |
| 'WHAT ₁ is such that John wants to know [<i>who</i> bought <i>it</i> ₁ there]?' | | | | | |

Note that the involved LF E-agreement is legitimate here and grammar should permit this scope relation. This scope restriction follows, however, from the *wh*-

P Theorem as stated in (20) since the *wh*-P in (23) is being associated with the matrix COMP and hence is not provided its interpretation within the LF E-agreement domain involving a focus property ([*WH*] in this case).¹²

Presenting paradigm (1), we also pointed out above that Wh-questions in general require EPD.

- (1) a. **DA**re-ga itumo ohiru-ni piza-o taberu-**no**↑?
 who-NOM always lunch-for pizza-ACC eat-COMP_{wh}
 'Who always eats pizza for lunch?'
- b. # **da**re-ga **i**tumo o**hi**ru-ni **pi**za-o taberu-no↑?
 who-NOM

¹² We can account for a similar contrast between the two sentences in (i). Here, the multiple *wh*-Ps can take subordinate scope within the LF E-agreement domain of a *WH*-P in (ia) but are disallowed to take matrix scope in (ib).

- (i) a. Kimi-wa [John-ga
 you-TOP -NOM

 | **DA**re-ni *nani*-o *ikura*-de utta-**ka**] mada oboeteiru-**no**↑?
 who-DAT what-ACC how.much-for sold-COMP_[wh] stillremember-COMP_[wh]
 |-----|-----|
 'WHO₁ is such that you still remember [*what* John sold to her₁ for *how much*]?'
- b. Kimi-wa [John-ga
 you-TOP -NOM

 | **#DA**re-ni *nani*-o *ikura*-de utta-**ka**] **ma**da oboeteiru-**no**↑?
 who-DAT what-ACC how.much-for sold-COMP_[wh] still remember-COMP_[wh]
 |-----|-----|
 'What₁ is such that you still remember [**to WHOM** John sold *it*₁ for *how much*]?'

The LF E-agreement involved in (ib) is legitimate and also satisfies the Relativized Opacity Condition discussed in Kim and Kitagawa (2002) and Kitagawa (2006). It, however, dissatisfies the *wh*-P *Theorem*.

- c. ①ohn-wa ①tumo o①hi-ru-ni ①pi-za-o taberu.
 John-TOP
 'John always eats pizza for lunch.'

Having introduced the distinction between *WH*-Ps and *wh*-Ps, we now can reassess the restriction observed in (1b) as the indication that a *wh*-P is incapable of making up a matrix Wh-question by itself. This restriction also follows from the *wh*-P *Theorem* since the *wh*-P here certainly is not parasitic to any LF E-agreement domain involving a focus property.

Although how exactly the *wh*-P *Theorem* should be derived remains to be worked out, acknowledging it permits us to capture various restricted behaviors of Wh-words, which otherwise would remain to be mysterious. When we recognize the information structural nature of this theorematic requirement, we can also recognize the extra-grammatical character of the observed restrictions and maintain the integrity of the elaborated LF E-agreement analysis in our grammar.

3.2 Processing Dead End

When we extend our observations to a construction containing more than one *WH*-P, we notice a curious absence of a certain type of scope interpretations. To begin with, suppose that we try to interpret a Wh-construction as in (24), letting both *WH*-Ps there be associated with the matrix COMP as indicated:

- (24)
- | | | | | | | | | | |
|----------------------------------|--------|----------|----------------------------|-------------------------------|-----------|---|-----------|-----|----|
| D | Are-ga | Mary-ga | N | Ani-o | eranda-ka |] | oboeteiru | -no | ↑? |
| who-NOM | -NOM | what-ACC | chose-COMP _{Wthr} | remember-COMP _[WH] | | | | | |
| 'Who remembers Mary chose what?' | | | | | | | | | |

of LF E-agreement involved here is legitimate, and no other grammatical problem seems to arise in the attempt to derive this interpretation.¹⁴ We thus face again a situation that is not anticipated in the "elaborated" LF E-agreement.

One thing we notice about (25c), however, is that the prosodic information assigned there is not distinguishable from that assigned to (24). In particular, since the lexical accent of *oBOeteiru* 'remember' in the matrix clause of (25c) is reduced in the post-focal reduction domain of the *Global* EPD starting from *DAre-ga*, this representation fails to provide any prosodic cue to mark the end of the *Local* EPD involved in its subordinate clause. The prosodic information assigned to (25c) therefore can be easily — perhaps inevitably — misinterpreted as a phonetic realization of *Global Compound* EPD involving both *DAre-ga* and *NAni-o*, which encompasses the entire utterance just as in (24). If so, when we perceive the prosody provided here and attempt to process the sentence, we are forced to analyze it as involving the E-agreement as in (24) rather than (25c), making the scope interpretation in question unavailable. This analysis of the interpretive restriction in (25c) is well in accordance with the processing principle in (26) argued for by Kitagawa and Fodor (2003).¹⁵

(26) Maximize Prosody-Syntax Congruence (Max PSC):

Attribute a prosodic property of a sentence to a syntactic property, and vice versa, whenever possible in processing a sentence.

Max PSC is designed to capture a very general preference for congruence between syntactic and prosodic structure in parsing, which encourages perceivers to assume a simple transparent relationship between prosody and

¹⁴ The opposite COMP-*WH*-P association is disallowed since the matrix *WH*-P (*DAre-ga* 'who-NOM') is not c-commanded by the subordinate COMP. See Section 1.2 above.

¹⁵ (26) is a slightly refined version of the Structural Interpretation of Prosody Principle proposed by Fodor (2002b).

syntax wherever possible, that is, unless they encounter evidence to the contrary. It then should also force the perceiver to derive synchronized matrix scope of *WH*-Ps in both (24) and (25c) based upon the prosody perceived there.

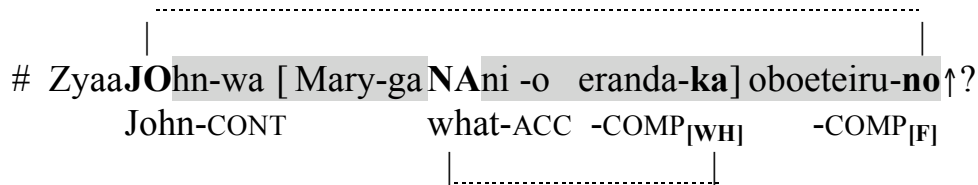
We can also extend this approach to the analyses of the constructions involving both *WH*-Ps and *FP*s. Observe first that we simply fail to come up with any legitimate interpretation when we combine these two types of focused items even when they appear in distinct clauses in whichever order to represent whatever scope relation, as shown in (27a-b). The general impression of the source of problem here is that both sentences involve an excessive focused item.

- (27) a. # Zyaa**JO**hn-wa [Mary-ga **NA**ni -o eranda-**ka**] oboeteiru- **no**↑?
 then -CONT -NOM what-ACC chose-COMP_[WH] remember-COMP_[F]
 (i) 'Then, **what** does **at least** John remember Mary selected?'
 (ii) 'Then, does **at least** John remember **what** Mary selected?'
 b. # Zyaa**DA**re-ga [Mary-ga **WA**in-wa eranda-**ka**] oboeteiru-**no**↑?
 then who-NOM -NOM wine-CONT chose-COMP_[F] remember-COMP_[WH]
 (i) 'Then, is **at least** wine such that **who** remembers whether Mary selected it?'
 (ii) 'Then, **who** remembers whether Mary selected **at least** wine?'

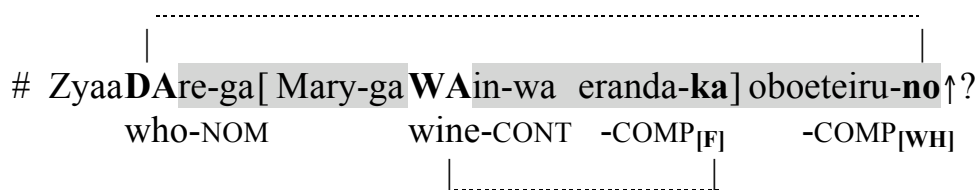
We can also explain the interpretive restrictions here, making an appeal to the Max PSC combined with the LF E-agreement approach. First, as we have already seen in (14) in Section 2.2 above, a *WH*-P and an *FP* cannot synchronize their scope because that would require an illegitimate type of COMP [*WH*][*F*], which would involve redundancy of a focus property. This would disallow the *WH*-P and *FP* in both these examples to take synchronized matrix scope that leads to the interpretation in (27a-i) and (27b-i). What is

puzzling is why the *WH*-P and *FP* cannot involve E-agreement as in (28a-b) and take distinct scope as in (27a-ii) and (27b-ii).¹⁶

(28) a.

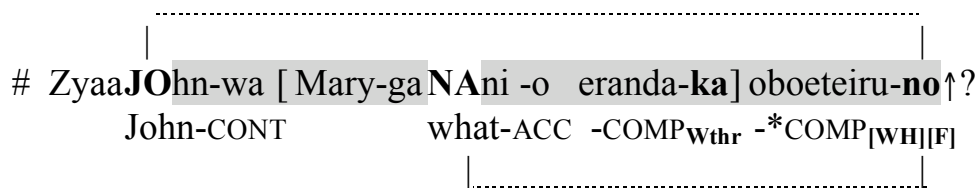


b.

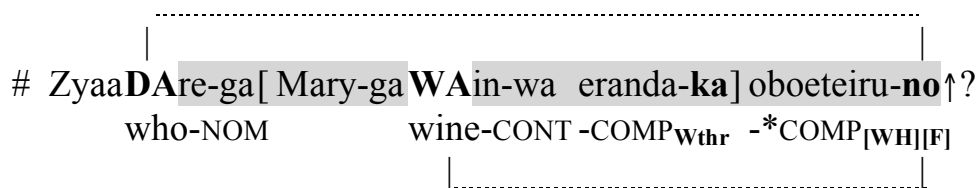


Here again, each E-agreement is legitimate and no other grammatical problem seems to arise, but the intended interpretation is not available. This interpretive restriction, however, can follow from the Max PSC since the prosodic information carried by (28a-b) is indistinguishable from that for *Global Compound* EPD, which we know will lead the perceivers to the parsing requiring an illegitimate type of COMP[*WH*][*F*] as in (29a-b).

(29) a.



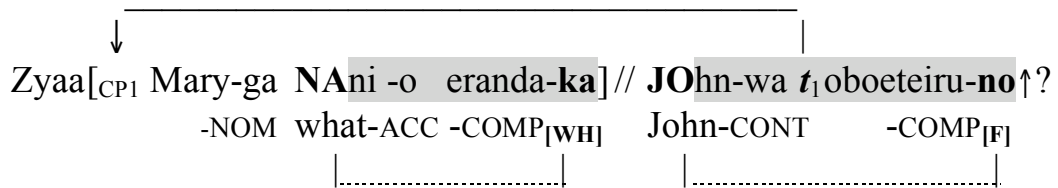
b.



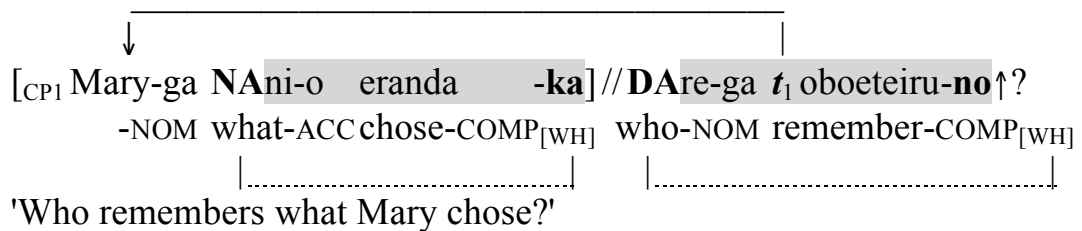
¹⁶ Again, the matrix *WH*-P can never take subordinate scope, failing to be c-commanded by the subordinate COMP.

Quite interestingly, when we scramble the embedded CP over the matrix focus as in (30a), the sentence comes to permit distinct scope for each focused item, as pointed out to me by Satoshi Tomioka (personal communication). As shown in (30b), the unacceptable multiple *WH*-P construction in (25c) also becomes interpretable with distinct scope.

(30) a.



b



One notable effect of scrambling here is that the emphatic accent of the matrix focus now comes to follow the embedded clause, which can be interpreted as a phonetic cue for the termination of the subordinate *Local* EPD, especially when a short pause is also added there. This prosodic pattern then can be interpreted as the phonetic realization of two distinct *Local* EPDs, one in the subordinate and the other in the matrix clause as indicated by // in (30a-b). This successfully leads to the distinct scope of the two focused items in these examples.¹⁷ Since all the grammatical operations involved in (29a-b) and (30a-b) remain the same,

¹⁷ Note that the same prosodic information in (30b) (without any pause) can be also regarded as the realization of *Global Compound* EPD for a matrix multiple Wh-question similar to (4b) above. This interpretation is indeed possible in (30b). Such an analysis, on the other hand, is not permitted in (30a) due to the illegitimacy of COMP[*WH*][*F*].

the contrast between the two cases strongly suggests that the interpretive restrictions observed in (29) does not originate in grammar and the solution appealing to processing as proposed here seems quite appropriate.

Finally, I would like to show how the proposed approach handles the same sentences when they are presented in writing without their prosody indicated, for instance, as in (31) (for (24) above).

- (31) Dare-ga [Mary-ga nani-o eranda-ka] oboeteiru-no?
 who-NOM -NOM what-ACC chose-COMPremember-COMP_[WH]
 (i) '**WHO** remembers **what** Mary selected?'
 (ii) '**WHO** remembers Mary selected **WHAT**'?

It might be thought that reading – especially silent reading – of a written example is immune to prosodic influences, but recent psycholinguistic findings suggest that this is not so. Sentence parsing data for languages as diverse as Croatian and Japanese are explicable in terms of the *Implicit Prosody Hypothesis* (IPH: Fodor (2002a), Fodor (2002b)), which explores the idea that prosody is always present in the processing of language, whether by ear or by eye.¹⁸ In the E-agreement approach pursued in this paper, this amounts to the claim that even the silent reading of sentences are interpreted based upon E-agreement taking place both at PF and LF. I would like to point out here that the silent reading of (31) and other written sentences permits a different range of scope interpretations from the pronounced examples we have examined above, but that range is still controlled by prosody in a very subtle way. First, (31) in silent reading permits the second Wh-phrase *nani-o* 'what-ACC' to be interpreted

¹⁸ IPH also pursues the hypothesis that a default prosodic contour is projected onto the stimulus in silent reading, which biases the parser toward the syntactic analysis associated with it. I will not discuss in this work the implication of this aspect of IPH to Wh-constructions in Japanese. See Kitagawa and Fodor (2003), Kitagawa and Fodor (2006) and Kitagawa (2005) for relevant discussion.

as a *wh*-P taking scope within the subordinate clause, just as in the way its pronounced version (32) is interpreted. This leads us to the interpretation (31-i).

(32)

D	A	re-ga	[Mary-ga		<i>nani-o</i>		eranda-ka]	oboeteiru- <i>no</i>	↑?
who-NOM						what-ACC		-COMP _[wh]		-COMP _[WH]	
.....											
'WHO remembers [<i>what</i> Mary selected]?' (= (31-i))											

In silent reading, we can also analyze both of the Wh-phrases in (31) as *WH*-Ps and interpret the sentence as a matrix multiple Wh-question, just as in the way its pronounced version (24) is interpreted. This leads us to the translation in (31-ii).

(24)

D	A	re-ga	[Mary-ga		N	A	ni-o		eranda-ka]		oboeteiru		-no	↑?
who-NOM						-NOM		what-ACC		chose-COMP _{Wthr}			remember-COMP _[WH]			
.....																
'WHO remembers Mary selected WHAT?' (= (31-ii))																

It probably is true that the distinction of these two interpretations can be sensed in silent reading only when we succeed in mentally associating them with distinct prosodic patterns as in (32) and (24). The reader can try to distinguish the two interpretations while forcing themselves not to assign any such prosodic contours and see how difficult it is. One thing we cannot do in (31), however, is to interpret both of Wh-phrases as foci, i.e., as *WH*-Ps, and let them take distinct scope, just as in the way its pronounced version (25c) is to be interpreted.

(25) c.

.....
 |-----|
 # **D**Are-ga [Mary-ga **N**Ani-o eranda **-ka**] oboeteiru **-no**↑?
 who-NOM -NOM what-ACC chose-COMP_[WH] remember-COMP_[WH]
 |-----|
 'Who remembers what Mary chose?'

We can ascribe this interpretive restriction in silent reading to the Max PSC just as we did above if we assume that the parsing of (31) is controlled by implicit prosody even when it is not accompanied by any overt prosody.¹⁹

Let us now examine the silent reading of (33) (for (27a) above).

- (33) Zyaa John-wa [Mary-ga nani-o eranda-ka] oboeteiru-no↑?
 then -TOP/CONT -NOM what-ACC chose-COMP remember-COMP
 (i) 'Then, **what**₁ is such that John remembers if Mary selected it₁?'
 (ii) 'Then, does John remember **what** Mary selected?'
 (iii) 'Then, **what** does **at least** John remember Mary selected?'
 (iv) 'Then, does **at least** John remember **what** Mary selected?'

When no discourse or pragmatic contexts are provided, the silent reading of this sentence leaves room for *John-wa* in the matrix to be analyzed either as a topic phrase or a contrast phrase. When it is interpreted as a topic phrase, the Wh-phrase must be analyzed as a *WH-P*. This *WH-P* may take either matrix or subordinate scope yielding (33-i) or (33-ii), but the choice depends on which of *Global EPD* and *Local EPD* is assigned to it implicitly, as described in (34a-b).

(34) a.

.....
 |-----|
 ZyaaJohn-wa [Mary-ga**N**Ani -o eranda-ka] oboeteiru-**no**↑?
 John-TOP what-ACC -COMP_{Wthr} -COMP_[WH]
 'Then, **what**₁ is such that John remembers if Mary selected it₁?' (= (33-i))

¹⁹ The first Wh-phrase in (31) cannot be interpreted as an unfocused Wh-phrase, i.e., as a *wh-P*, due to the *wh-P* Theorem, as we pointed out in Section 3.1.

- b. ZyaaJohn-wa [Mary-ga NAni -o eranda-ka] o (b)eteiru-no↑?
 John-TOP what-ACC -COMP_[WH] -COMP_{Y/N}
 |.....|
 'Then, does John remember **what** Mary selected?' (= (33-ii))

Again, it would be rather difficult in silent reading to distinguish these two scope interpretations without implicitly assigning the distinct prosodic patterns.

When *John-wa* in (33) is analyzed as a contrast phrase, the sentence also permits both matrix and subordinate scope reading of the Wh-phrase as in (33-iii) and (33-iv), but the Wh-phrase must be demoted to a non-focus, i.e., to a *wh*-P, and assigned implicit prosody as in (35a-b).

- (35) a.
- Zyaa**J**Ohn-wa [Mary-ga nani -o eranda-ka] oboeteiru-**no**↑?
 John-CONT what-ACC -COMP_{Wthr} -COMP_{[F][wh]}
 |.....|
 'Then, **what** does **at least** John remember Mary selected?' (= (33-iii))
- b.
- Zyaa**J**Ohn-wa [Mary-ga nani -o eranda-ka] oboeteiru-**no**↑?
 John-CONT what-ACC -COMP_[wh] -COMP_[F]
 |.....|
 'Then, does **at least** John remember **what** Mary selected?' (= (33-iv))

Since prosody here, implicit or explicit, indicates only the scope of the *FP*, and *wh*-Ps in general are not accompanied by EPD, it is not an easy task to detect the two distinct scope readings of the *wh*-P arising from two distinct modes of E-agreement indicated here. The distinction, nonetheless, can be made when we can consider two distinct situations as follows. First, Mary is known to have selected two items — one item when Bill was with her and the other item when John was with her. After learning that Bill remembers what she selected when

he was with her, the speaker utters (35a), asking for the identity of the other item Mary selected when John was with her. This involves the matrix scope interpretation of the *wh*-P as in (33-iii). In a similar situation, after learning that Bill did not remember what Mary selected, the speaker utters (35b), asking whether John, unlike Bill, remembers what Mary selected. This involves the subordinate scope interpretation of the *wh*-P as in (33-iv). It seems that the direct Wh-question reading can be made more easily available by adding *-no hoo* 'as a choice' before *-wa* in (35a) and the indirect Wh-question reading can be forced by using *-nara* or *-dattara* 'if ... is the choice' instead of *-wa* in (35b).

One thing we cannot do in the silent reading of (33), on the other hand, is to analyze *John-wa* as a contrast phrase and at the same time to interpret the Wh-phrase there as focused, i.e., as a *WH*-P, either with its matrix or subordinate scope. This interpretive restriction directly follows from our analyses above when we assume that prosody is implicitly assigned in silent reading, as *Global Compound* EPD as in (29a) and as co-occurrence of a *single Global* EPD and another *single Local* EPD as in (28a) repeated here.

(29) a.

# Zyaa	JO hn-wa	[Mary-ga	NA ni -o	eranda- ka]	oboeteiru- no ↑?
John-CONT		what-ACC	-COMP _{Wthr}	- [*] COMP _{[WH][F]}	

(28) a.

# Zyaa	JO hn-wa	[Mary-ga	NA ni -o	eranda- ka]	oboeteiru- no ↑?
John-CONT		what-ACC	-COMP _[WH]	-COMP _[F]	

As we argued above with the sentences involving overt prosody, E-agreement in (29a) would require an illegitimate COMP with the redundant feature

specification [*WH*][*F*], and the parsing as in (28a) would not be available due to the Max PSC. Both the information structural restrictions imposed by the *wh*-P Theorem and the processing restrictions imposed by the Max PSC, in other words, are observed even when the sentences are processed in silent reading.

4 Conclusions

In Sections 1 and 2, I sketched out a general research method incorporating prosodic and other extra-syntactic/extra-grammatical analyses into the formal study of syntax, and some details of its "elaborated" LF E-agreement approach. In Sections 3.1 and 3.2, I argued that certain representations permitted by grammar with the appropriate application of E-agreement may become unacceptable when some extra-grammatical problems arise. One such case involves an information structural problem, which we proposed to capture with the *wh*-P Theorem. Another case involves processing restrictions imposed by the specific prosodic pattern assigned to a sentence, whether or not the sentence is presented with overt prosody or is assigned implicit prosody by a parser when it is processed in silent reading. I argued that they can be captured by the Max PSC, a prosody-sensitive processing constraint.

References

- Choe, Jae-Woong. 1985. Pitch-accent and q/wh Words in Korean. In *Proceedings of 1985 Harvard Workshop on Korean Linguistics (WOKL)*, 113-123.
- Deguchi, Masanori and Yoshihisa Kitagawa. 2002. Prosody and Wh-questions. In *Proceedings of the Thirty-second Annual Meeting of the North-Eastern Linguistic Society*, 73-92. GLSA, University of Massachusetts at Amherst.
- Fodor, Janet Dean. 2002a. Prosodic Disambiguation In Silent Reading. In *Proceedings of the Thirty-second Annual Meeting of the North-Eastern*

-
- Linguistic Society*, 113-137. GLSA, University of Massachusetts at Amherst.
- Fodor, Janet Dean. 2002b. Psycholinguistics Cannot Escape Prosody. In *Proceedings of the Speech Prosody 2002 Conference*, 83-88. Aix-en-Provence, France.
- Heycock, Caroline. 1994. Focus Projection in Japanese. In *Proceedings of NELS 24*, 157-171. GLSA, University of Massachusetts at Amherst.
- Hirotsu, Masako. 2003. Prosodic Effects on the Interpretation of Japanese Wh-questions, Alonso-Ovalle, Luis ed. *University of Massachusetts Occasional Papers in Linguistics 27 - On Semantic Processing*. 117-137.
- Hirotsu, Masako. 2004. *Prosody and LF: Processing of Japanese Wh-questions*. Doctoral dissertation, University of Massachusetts at Amherst.
- Ishihara, Shinichiro. 2002. Invisible but Audible Wh-scope Marking: Wh-constructions and Deaccenting in Japanese. In *Proceedings of the Twenty-first West Coast Conference on Formal Linguistics*, 180-193. Cascadilla Press.
- Ishihara, Shinichiro. 2003. *Intonation and Interface Conditions*. Doctoral dissertation, Massachusetts Institute of Technology.
- Jackendoff, Ray. 1972. *Semantic Interpretation in Generative Grammar*. Cambridge, Massachusetts: The MIT Press.
- Kim, Ae-ryung and Yoshihisa Kitagawa. 2002. Opacity in Japanese and Korean. In *Akatsuka, Noriko M. and Susan Strauss (eds.) Japanese/Korean Linguistics 10*, 601-614. CSLI Publications.
- Kitagawa, Yoshihisa. 1999. Economy of Lexical Selection. Paper presented at LICSSOL 1 — Economy in Language Design, Computation and Use, The Lyon Institute for Cognitive Science,
- Kitagawa, Yoshihisa. 2005. Prosody, Syntax and Pragmatics of Wh-questions in Japanese. *English Linguistics*. 22:
- Kitagawa, Yoshihisa. 2006. Wh-scope Puzzles. In *Proceedings of the Thirty-fifth Annual Meeting of the North-Eastern Linguistic Society*, GLSA.
- Kitagawa, Yoshihisa and Masanori Deguchi. 2002. Prosody in Syntactic Analyses. ms., Indiana University.

-
- Kitagawa, Yoshihisa and Janet Dean Fodor. 2003. Default Prosody Explains Neglected Syntactic Analyses of Japanese. In *Japanese/Korean Linguistics 12*, 267-279. CSLI Publication.
- Kitagawa, Yoshihisa and Janet Dean Fodor. 2006. Prosodic Influences on Syntactic Judgments. In *Gradience in Grammar: Generative Perspectives*, ed. Fanselow, Gisbert, et al., Oxford, UK: Oxford University Press.
- Kubo, Tomoyuki. 2001. Syntax-Phonology Interface in the Fukuoka Dialect. *Journal of the Phonetic Society of Japan*. 5: 27-32.
- Kubozono, Haruo. 2007. Focus and Intonation in Japanese. In *Interdisciplinary Studies on Information Structure 9: Working Papers of SFB632*, ed. Ishihara, Shinichiro, Potsdam: University of Potsdam.
- Lee, Hyo Sang. 1982. Asymmetry in Island Constrains in Korean. ms., University of California at Los Angeles.
- Maekawa, Kikuo. 1991. Perception of intonation characteristics of WH and non-WH questions in Tokyo Japanese. In *Proceedings of the XXIIInd International Congress of Phonetic Science 4*, 202-205. Ain-en-Provence, France: Universite de Provence.
- Nagahara, Hiroyuki. 1994. *Phonological Phrasing in Japanese*. Doctoral dissertation, University of California, Los Angeles.
- Nishigauchi, Taisuke. 1990. *Quantification in the Theory of Grammar*. Dordrecht: Kluwer Academic Publishers.
- Pierrehumbert, Janet and Mary Beckman. 1988. *Japanese Tone Structure*. Cambridge, Massachusetts: The MIT Press.
- Reinhart, Tanya. 1997. Quantifier Scope: How Labor is Divided between QR and Choice Functions. *Linguistics and Philosophy*. 20: 335-397.
- Sugahara, Mariko. 2003. *Downtrends and Post-focus Intonation in Tokyo Japanese*. Doctoral dissertation, University of Massachusetts at Amherst.
- Tomioka, Satoshi. 1997. Wh-in-situ, Subjacency, and LF Syntax. ms., Cornell University.
- Tomioka, Satoshi. 2004. Pragmatics of LF Intervention Effects: Japanese and Korean Interrogatives. ms., University of Delaware.

Truckenbrodt, Hubert. 1995. *Phonological Phrases: Their Relation to Syntax, Focus, and Prominence*. Doctoral dissertation, Massachusetts Institute of Technology.

Vallduvi, Enric. 1992. *The Information Component*. New York: Garland.

Vallduvi, Enric. 1995. Structural Properties of Information Packaging in Catalan. In *Discourse Configurational Languages*, ed. E Kiss, Katalin, Oxford: Oxford University Press.

Yoshihisa Kitagawa
Memorial Hall 325
Department of Linguistics
Indiana University
Bloomington, IN 47405
U.S.A.
kitagawa@indiana.edu
<http://www.iub.edu/~ykling/>

Intonation of Sentences with an NPI*

Shinichiro Ishihara

University of Potsdam

This paper presents the results of a production experiment on the intonation of sentences containing a *negative polarity item (NPI)* in Tokyo Japanese. The results show that NPI sentences exhibit a *focus intonation*: the F_0 -peak of the word to which an NPI is attached is raised, while the pitch contour after the NPI-attached word is compressed until the negation. This intonation pattern is parallel to that of *wh*-question, in which the F_0 of the *wh*-phrase is raised while the post-*wh*-contour is compressed until the question particle.

Keywords: Japanese, negative polarity item (NPI), focus intonation, wh-question

1 Introduction

This paper presents the results of a production experiment on the intonation of sentences containing a *negative polarity item* (henceforth, *NPI*) in Tokyo Japanese¹. The experiment will examine how sentences with an NPI are phonetically realized.

NPIs are a group of words that can only appear in the scope of negation.² In the production experiment to be reported here, the NPI *sika* was used. *Sika*, together with the negation, means ‘only / nothing but...’, as shown below.

* I would like to thank Gisbert Fanselow, Ingo Feldhausen, Caroline Féry, Haruo Kubozono, and the participants of WPSI 2 for their comments and discussion. Special thanks go to Shrvan Vasishth for his help on statistic analysis, as well as to Felix Engelmann for his assistance on data analysis. All the errors are of course my own.

¹ In this paper, we will only discuss intonation of Tokyo Japanese. For brevity, I will call it ‘Japanese’ for the rest of the paper.

² There are some kinds of NPIs which appears non-negative environments as well (Ladusaw 1979). In this paper, however, we only use the so-called ‘strong NPIs’, which can only appear in the scope of negation. See Vasishth (1998) for other types of NPI in Japanese.

Interdisciplinary Studies on Information Structure 9 (2007): 65–96

Ishihara, Shinichiro (ed.):

Proceedings of the 2nd Workshop on Prosody, Syntax, and Information Structure (WPSI 2)

©2007 Shinichiro Ishihara

- (1) Mária-**sika** rámu-o nomíya-de nomá-**nakat**-ta
 Mária-SIKA rum-ACC bar-LOC drink-NEG-PST
 ‘Only Mari drank rum at the bar./No one but Mari drank rum at the bar.’

There has been a claim that NPI sentences has a certain prosodic constraint: the NPI and the negation must be within the same prosodic phrase (Hirotani 2005; Lee and Tomioka 2001; Tomioka 2004). The main goal of the paper is to examine the intonation of sentences like (1) and see if the claim is supported experimentally. As we will see below, the results of the experiment actually confirms the claim. Three phonetic phenomena are observed in the sentences with a *sika*-phrase: (i) F_0 -rise of the word to which *-sika* is attached, (ii) the F_0 -downtrend of the post-NPI material, and (iii) the pitch reset after the negation.

This intonation pattern of NPI sentences is parallel to that of *wh*-questions, in which the F_0 of the *wh*-phrase is raised while the post-*wh*-contour is compressed until the question particle that binds the *wh*-phrase. Following Ishihara (2005, 2007b), we will call this intonation *focus intonation (FI)*. An FI is characterized by three phonetic phenomena: (i) an F_0 -rise of the focused phrase, (ii) a F_0 -downtrend of the post-focal material, and (iii) the pitch reset after the scope of the focus. The results of the experiment suggest that an NPI, together with its licenser (i.e., negation), induces an FI within the scope of the negation.

The paper is organized as follows. In the next section (§2), we will briefly review the intonation of *wh*-questions (§2.1), a previous claim about the intonation of NPI sentences (§2.2), and the assumptions about FI taken in this paper (§2.3). §3 explains the details of the production experiment. The result of the experiment will be presented in §4, followed by discussion in §5.

2 Background

2.1 Focus Intonation in *Wh*-questions

It has been observed that a Japanese *wh*-question sentence obligatorily exhibits an FI: The F_0 -peak of the *wh*-phase is raised (*focal F_0 -rise*) while the F_0 -peaks of the post-*wh*-phrases are lowered (*post-focal F_0 -downtrend*) (Maekawa 1991, 1997). Furthermore, Deguchi and Kitagawa (2002) and Ishihara (2002, 2003) claim that the phonological domain of the FI (henceforth, *FI domain*) and the semantic scope of *wh*-question shows a correspondence. A post-focal F_0 -downtrend in a *wh*-question continues until the end of the scope of the *wh*-question, where the question particle that binds the *wh*-phrase appears.

For example, in a matrix *wh*-question like (2a), the post-focal downtrend continues until the end of the matrix clause, where the matrix question particle *no* appears (Figure 1), while in an indirect *wh*-question like (2b), the post-focal downtrend stops at the end of the embedded clause, where the question particle *ka* appears, and the pitch range is reset to the original, non-compressed level thereafter (Figure 2). This essentially means that the FI domain indicates the scope of the *wh*-question. (See Ishihara (2003, 2005) for explanation how this FI-*wh*-scope correspondence is derived.)

(2) a. *Matrix wh-question*

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

ímademo obóeteru
 even.now remember
 ‘Naoya still remembers what_i Mari drank t_i at the bar.’

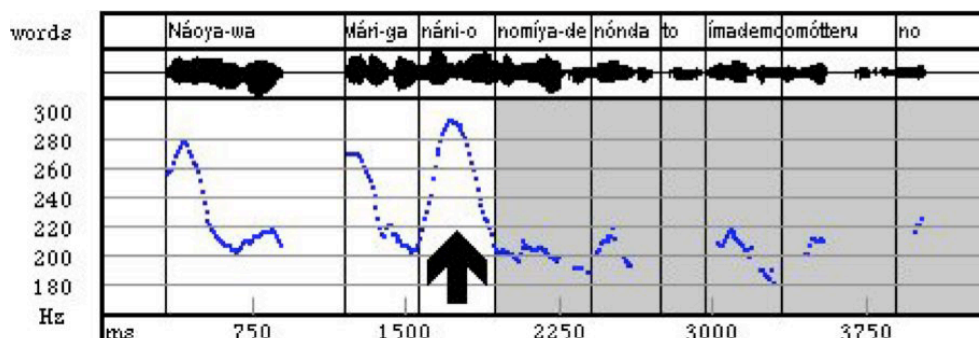


Figure 1: Matrix *wh*-question

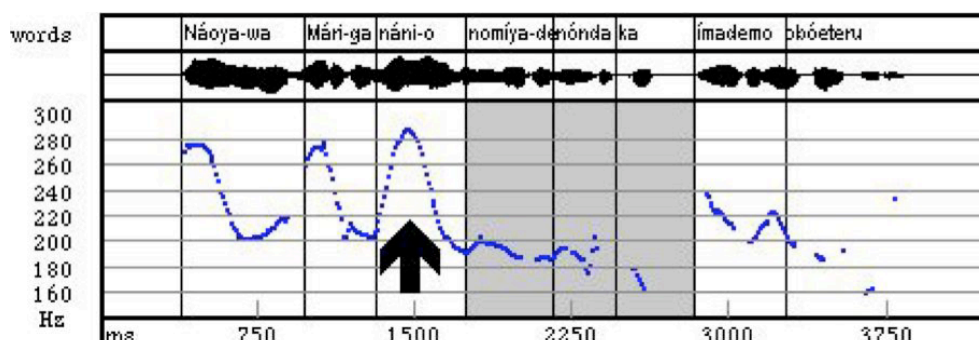


Figure 2: Indirect *wh*-question

A similar claim has been made from a processing point of view. Hirotani (2005) claims that a processing principle called *Scope-Prosody Correspondence (SPC)* requires that a *wh*-phrase and the question particle binding it be in the same prosodic phrase, namely *Major Phrase (MaP)*, in order for the *wh*-scope to be interpreted properly. According to the standard assumption about Japanese FI (e.g., Pierrehumbert and Beckman 1988; Nagahara 1994; Truckenbrodt 1995), which is adopted by Hirotani, but not in this paper (see §2.3 below), MaP is the

domain of FI. Therefore it is equivalent to say in our terms that SPC requires that the *wh*-phrase and the question particle be in a single FI domain.

Strictly speaking, there is one difference between Hirotani's claim and the one proposed by Deguchi and Kitagawa (2002) and Ishihara (2002, 2003). The requirement of Hirotani's SPC is weaker than that of the other proposals in that the pitch reset after negation is not obligatorily expected in Hirotani's SPC. We will discuss pitch reset in the results and the discussion sections (§4.3, §5.2). In any case, it is a well-observed fact that *wh*-phrase, together with a question particle, triggers an FI, so that they are grouped prosodically into a single FI domain.

2.2 Negative Polarity Items (NPI) and FI

A similar claim has been made for sentences containing an NPI (cf. Hirotani 2005; Lee and Tomioka 2001; Tomioka 2004). Hirotani (2005) claims, extending her analysis of *wh*-question, that SPC requires that an NPI and the negation binding it be in the same MaP.

If a processing principle like SPC expects such a prosodic marking for a NPI-NEG relation, we would also expect in terms of production that an NPI and a negation trigger an FI to be included in the same prosodic domain, just like a *wh*-phrase and a question-particle trigger one. I will call this hypothesis *NPI-FI Hypothesis*:

(3) *NPI-FI Hypothesis*

An NPI triggers an FI within the domain of negation.

- a. A *focal F₀-rise* of the phrase to which an NPI attaches.
- b. A *post-focal downtrend* on all the material following the NPI until the negation that binds the NPI.
- c. A *pitch reset* after the negation.

For example, if an NPI and a negation are in the embedded clause as in (4a), an FI would appear only within the embedded clause, starting from the phrase to which the NPI is attached (*Mari*) until the verb to which the negation *-nakat-* is attached (*noma-* ‘drink’). The pitch range will be reset after the embedded clause.

On the other hand, if the NPI and the negation are in the matrix clause as in (4b), the FI would appear on the matrix clause (and contain the entire embedded clause in its domain). In (4b), the F_0 of the matrix subject *Naoya* will be raised, while all the F_0 -peaks thereafter will be lowered until matrix verbal complex head containing negation *iwa-nakat-ta* ‘say-NEG-PST’.

(4) a. *NPI in the embedded clause*

Náoya-wa [Mári-**sika** rámu-o nomíya-de nomá-**nakat**-ta to]
 Naoya-TOP Mári-SIKA rum-ACC bar-LOC drink-NEG-PST that
 Yúmi-ni itta
 Yumi-DAT said
 ‘Naoya said to Yumi that only Mari drank rum at the bar.’

b. *NPI in the matrix clause*

Náoya-**sika** [Mári-ga rámu-o nomíya-de nónda to]
 Náoya-SIKA Mari-NOM rum-ACC bar-LOC drank that
 Yúmi-ni iwa-**nákat**-ta
 Yumi-DAT say-NEG-PST
 ‘Only Naoya said to Yumi that Mari drank rum at the bar.’

This FI-NPI hypothesis, as far as I know, has never been experimentally examined in terms of production.³ In this paper, therefore, I will present the results of the production experiment testing the FI-NPI hypothesis.

³ For a perception experiment, see Hirotani (2005).

2.3 Definitions

Before going into the details of the experiment, let us make clear the definitions of the phonetic phenomena to be examined in the experiment. I will assume that FI can be detected by the three phonetic phenomena listed in (5). They are schematically illustrated in Figure 3 and 4:

- (5)
- a. F_0 -rise on the focused phrase (e.g., *wh*-phrase, NPI)
 - b. post-focal F_0 -downtrend
 - c. pitch reset after FI domain.

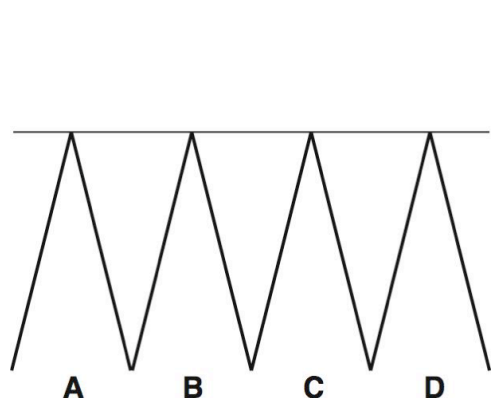


Figure 3: Default contour (No FI)

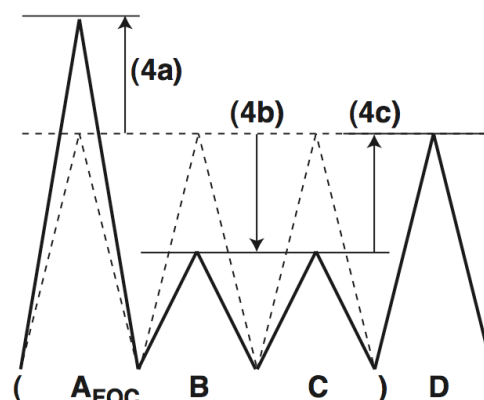


Figure 4: FI (Focus A; FI domain A–C)

I will assume that focus F_0 -rise (5a) is a phonetic effect that raises the F_0 -peak of the phrase bearing (semantic) narrow focus, and that *post-focal F_0 -downtrend* (5b) is a phonetic effect that compresses the pitch range of the post-focal material. In other words, an FI is created as a result of direct manipulation of pitch range. In the schematic illustration in Figure 4, the pitch range of the focused phrase (A) is expanded, while that of post-focal elements (B and C) is compressed, resulting in lower F_0 -peaks for these phrases. *FI domain* is the phonological domain in which (5a) and (5b) apply. In Figure 4, the FI domain, indicated by brackets (), contains A, B, and C.

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 boundaries. Under these analyses, focus F_0 -rise is explained as an insertion of MaP boundary on the left of focused phrase,⁴ and post-focal downtrend as *downstep* as a result of MaP boundary deletion at the post-focal area. In other words, in the standard analyses, a MaP 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 (5). In other words, the domain of downstep (MaP) and the domain of the FI (FI domain) are not necessarily identical. See Ishihara (2007a,b) for arguments for this assumption about FI.⁵ See also Kubozono (this volume) for experimental evidence against the standard “FI = MaP” analysis.

Pitch reset (5c) is a phenomenon which cancels the effect of post-focal downtrend after the FI domain. In Figure 4, where the FI domain is assumed to be (A B C), the compressed pitch range of the post-focal 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 (D) has the non-compressed pitch height.

This means that a pitch reset after the post-focal downtrend will indicate the end point of the FI domain. In the indirect *wh*-question in (2b) above, for example, an FI is observed in the embedded clause: Focus F_0 -rise raises the F_0 -peak of the *wh*-phrase *nani-o*; the post-focal downtrend compresses the pitch

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

⁵ How the prosodic phrasing and FI (under the assumption adopted here) interact with each other is discussed in Ishihara (2007b).

contour after the *wh*-phrase until the end of the embedded clause, where the question particle *ka* appears; and the pitch range is reset to the original height after the question particle. The FI domain in this case is between the *wh*-phrase and the question particle.

As it turned out from the results of the experiment, there are two different ways in which pitch reset is realized. In some utterances a high peak is observed only on the phrase after the embedded clause, while in other utterances another sharp F_0 -rise appears on the complementizer of the embedded clause, and creates a higher peak than that of the following phrase. Sample pitch contours are given in Figure 5 and 6. These two samples are taken from the recordings of the same sentence from a single speaker.

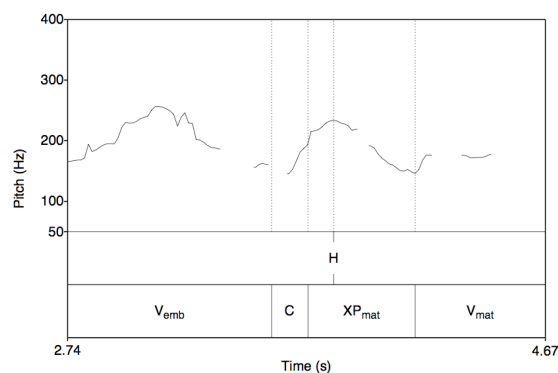


Figure 5: XP-type pitch reset

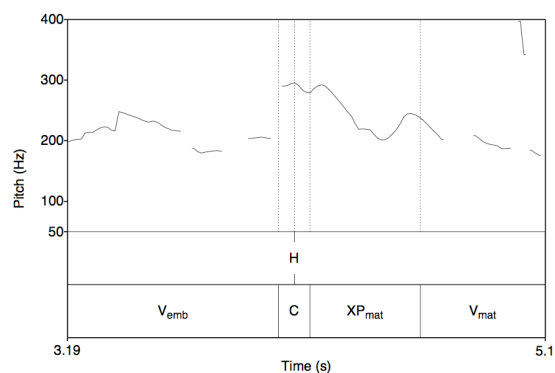


Figure 6: COMP-type pitch reset

In Figure 5, the last mora of the embedded clause, which is the complementizer (C), is realized low, and a high peak appears on the following phrase (XP_{mat}). In Figure 6, on the other hand, the complementizer bears a higher peak than the following phrase.

From this fact, I tentatively assume that the pitch reset is realized either on the first high peak of the next prosodic domain, or at the end of the F_0 -lowered prosodic phrase as some kind of boundary tone. For expository purposes, I will call the first type of pitch reset *XP-type* (pitch reset realized on the XP following

the embedded clause), and the latter *Comp-type* pitch reset (pitch reset realized on the embedded clause complementizer).

Although some speakers seemed to have a tendency to use *Comp-type* more frequently than reset in a consistent manner, it appears that both variations are available for everyone. But crucially, it seems that the choice of *XP-type/Comp-type* is also strongly dependent on the experiment conditions. We will discuss this phenomenon more in detail in §5.2. It is sufficient here just to keep in mind that there are two different places where the pitch reset may be realized.

3 Experiment

3.1 Goal

The goal of this experiment is to examine the validity of the NPI-FI hypothesis in (3). More specifically, it is designed to check whether the following phenomena listed in (6), repeated below, are actually observed. If NPI sentences are to trigger FIs just like *wh*-questions, these phenomena are expected in their pitch contours.

- (6) a. F₀-rise on NPI
- b. F₀-lowering on post-NPI material
- c. F₀-reset on post-negation material

3.2 Method

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

Stimuli 8 sets of 3 types of target sentences (24 total, see §3.3 and Appendix A for detail)

Presentation of the stimuli Stimuli are mixed with 112 filler sentences (used as stimuli for other experiments), 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 (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 from 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.⁶ 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. (See Appendix B for the individual results.)

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, using a normalization method adopted in Truckenbrodt (2004). All the measured values are transformed according to the following linear transformation:

$$\text{transformed_value} = (\text{original_value} - \text{Av}_S(R_2)) / (\text{Av}_S(R_1) - \text{Av}_S(R_2))$$

⁶ This tendency of NM's data has been consistently observed for other experiments as well (cf. Ishihara 2003).

where $Av_S(R_n)$ is the speaker-specific mean F_0 -value of the two reference point R_1 and R_2 . This formula rescales the mean of R_1 measurements to 1 and the mean of R_2 measurements to 0, for each speaker. The following two values are chosen as the reference points (R_1, R_2) for the normalization:

(7) *Reference points for the normalization formula*

R_1 : Mean highest F_0 -value of the embedded clause subject (P1 in (9)).

R_2 : Mean lowest F_0 -value of the phrase immediately following the embedded clause (i.e., L tone immediately after P3 in (9))⁷

Equipment The recorded data was digitized using SimpleSound on a Macintosh PowerBook G3. Segmentation and F_0 measurement was done using Praat (Boersma and Weenink 1992–2007) with the help of Praat scripts. After the half-automated measurement, I checked the data using Praat one by one to make sure that the measurements were done appropriately by the scripts. When some wrong measurement points were found, I modified them by hand and updated the results. Statistic data analysis was done using R.

3.3 Stimuli

In the experiment, I used the NPI *-sika*, which, together with negation, means ‘only’. Three sentence types are compared in the experiment. Below is one of the eight stimulus sets used in the experiment. (See Appendix A for the complete stimulus sets). A is the control sentence with no NPI. B has an NPI and a negation in the embedded clause, while C has an NPI and a negation in the matrix clause.

⁷ There are cases where the highest peak of the phrase is realized at the end of the phrase (i.e., on the PP/case-marker). In such a case, the lowest point before P3 is measured.

(8) A. *No NPI (Control)*

Náoya-wa [Mári-ga rámu-o nomíya-de nomá-**nakat**-ta to]
 Naoya-TOP Mari-NOM rum-ACC bar-LOC drink-NEG-PST that
 Yúmi-ni itta
 Yumi-DAT said
 ‘Naoya said to Yumi that Mari didn’t drink rum at the bar.’

B. *NPI in the embedded clause*

Náoya-wa [Mári-**sika** rámu-o nomíya-de nomá-**nakat**-ta to]
 Naoya-TOP Mári-SIKA rum-ACC bar-LOC drink-NEG-PST that
 Yúmi-ni itta
 Yumi-DAT said
 ‘Naoya said to Yumi that only Mari drank rum at the bar.’

C. *NPI in the matrix clause*

Náoya-**sika** [Mári-ga rámu-o nomíya-de nónda to]
 Náoya-SIKA Mari-NOM rum-ACC bar-LOC drank that
 Yúmi-ni iwa-**nakat**-ta
 Yumi-DAT say-NEG-PST
 ‘Only Naoya said to Yumi that Mari drank rum at the bar.’

In order to check the three prosodic phenomena listed in (6), the F_0 -peaks of the following three phrases in each stimulus sentence are measured. They are labeled P(eak)1, P2, and P3, respectively, as shown in (9) below. As mentioned in §2.3, there are two places where a pitch reset is realized: the embedded clause complementizer (C_1) or the phrase following it (YP). Therefore I decided to measure the F_0 of both words and used whichever higher as the value for P3.

(9) *Labels of the relevant F_0 peaks*

[CP_2 Subj₂ [CP_1 Subj₁ XP ... V₁-NEG C_1] YP V₂(-NEG)]
 P1 P2 P3* P3

*Only when this peak is higher than that of YP.

P1: Embedded clause subject (Subj₁)

P2: Material immediately following the embedded clause subject (XP)

P3: The matrix phrase immediately following the embedded clause (YP)
(Or the embedded clause Complementizer (C₁), if its F₀ is higher than YP)

The expected FIs in the stimuli in (8) is schematically illustrated in (10) (Box indicates the F₀-rise, and underline indicates the F₀-downtrend). In the control stimulus A, no FI is expected. The pitch contour of this sentence would be a default pitch contour. B has an NPI-NEG pair in the embedded clause. Accordingly an FI is expected between the NPI-attached word, i.e., the embedded clause subject (Subj₁) and the embedded clause verbal complex (V-NEG). After the FI, F₀-downtrend effect should be cancelled by pitch reset. In C, an NPI-NEG pair is in the matrix clause. Therefore F₀-rise is expected on the matrix subject (Subj₂), and F₀-downtrend is expected until the end of the sentence.

(10) *Schematic representation of (8)*

	P1	P2	P3
A.	[CP ₂ Subj ₂	[CP ₁ Subj ₁ XP ... V ₁ -NEG C ₁]	YP V ₂]
B.	[CP ₂ Subj ₂	[CP ₁ <u>Subj₁</u> -NPI XP ... V ₁ -NEG C ₁]	YP V ₂]
C.	[CP ₂ <u>Subj₂</u> -NPI	[CP ₁ Subj ₁ XP ... V ₁ C ₁]	YP V ₂ -NEG]

3.4 Predictions

From (10), we can make predictions regarding the three peaks P1, P2, and P3.

P1 At P1 (embedded clause subject), we expect to observe a F₀-rise effect in B sentence, because in B an NPI is attached to this phrase. Also, we expect a post-focal F₀-downtrend in C sentence, because in C the matrix subject is attached

an NPI and accordingly triggers an F_0 -downtrend on the following phrases. As a result, P1 in B and C are expected to show difference in terms of F_0 -height compared to the control stimulus A: B is higher than A (= (11a)), and C is lower than A (= (11b)).

(11) *Predictions for P1*

- a. $A < B$ (due to F_0 -rise in B)
- b. $A > C$ (due to F_0 -lowering in C)

P2 At P2, F_0 -downtrend is expected both in B and C, because in both sentences, P2 follows the NPI-attached phrase. Accordingly, P2 in B and C is lower than that of the control sentence A, where no F_0 -lowering is expected.

(12) *Predictions for P2*

- a. $A > B$ (due to F_0 -lowering in B)
- b. $A > C$ (due to F_0 -lowering in C)

P3 Lastly, at P3 we expect a pitch reset in B. According to the NPI-FI hypothesis in (3), the FI in B should be found only within the embedded clause. Therefore, the F_0 -peak of the phrase after the embedded clause should exhibit a pitch reset. If that's the case, P3 should become as high as the control case, A.

In C, on the other hand, FI is expected in the matrix clause. Therefore the F_0 -downtrend is expected to continue on P3. As a result, we predict that P3 is lower than that of A and B.

(13) *Predictions for P3*

- a. $A = B$ (due to pitch reset in B)
- b. $A > C$ (due to F_0 -lowering in C)
- c. $B > C$ (due to pitch reset in B and F_0 -lowering in C)

4 Results

Figure 7 shows the normalized means of P1, P2, and P3, with 95% confidence interval. (See Appendix B for individual results.) As will be shown below, all the predictions are supported by the results.

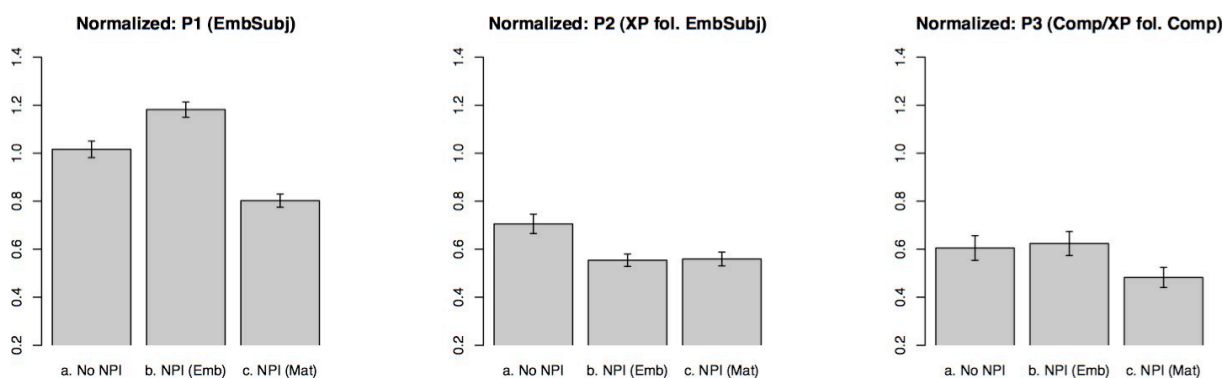


Figure 7: Normalized Means of P1, P2, and P3, with 95% CIs

4.1 P1

The predictions for P1 (11) are repeated below:

(11) *Predictions for P1*

- a. $A < B$ (due to F_0 -rise in B)
- b. $A > C$ (due to F_0 -lowering in C)

As we can see in Figure 7, the two predictions for P1 are borne out. B is significantly higher than A (1 sided t-test, $t(190) = -6.9697$, $p < 0.0001$), and C is significantly lower than A ($t(181.078) = 9.6701$, $p < 0.0001$).⁸

As for the individual results, the expected F_0 -rise in B (i.e., (11a)) were observed in all subjects except one marginal result from KS (1-sided t-test, $t(46)$

⁸ For the t-tests here and hereafter, the F test is done to check the variance of the two samples. If the two variances are not equal, Welch's correction is made on t-test.

= -1.8067 , $p = 0.03868$).⁹ From these results, we can conclude that the F_0 -rise on the NPI-attached word is a quite steady phenomenon.

The F_0 -lowering effect expected in C (i.e., (11b)) was statistically significant in three subjects' (AH, CS, YY), but not in KS's data (1-sided t-test, $t(38.671) = 0.7764$, $p = 0.2211$). In fact, KS did not show any F_0 -lowering effect in P2 and P3, either. Given that KS's F_0 -rise effect was also only marginally significant, it may be the case that she does not exploit FI for prosodic marking of NPI sentences. It will be shown later, however, that she uses a particular way of NPI-domain marking, namely strong 'upstep' after the negation.

All in all, the F_0 -rise effect and F_0 -lowering effect expected in P1 were both confirmed by the results (except KS's).

4.2 P2

The predictions for P2 (12) are repeated below:

(12) *Predictions for P2*

- a. $A > B$ (due to F_0 -lowering in B)
- b. $A > C$ (due to F_0 -lowering in C)

Again, both predictions are borne out in the normalized results, as shown in Figure 7. B and C are both significantly lower than A (A vs. B: 1 sided t-test, $t(160.981) = 6.2665$, $p < 0.0001$; A vs. C: 1 sided t-test, $t(171.153) = 5.853$, $p < 0.0001$).

Individually, KS did not show any clear sign of F_0 -lowering effect, as mentioned above. Therefore neither of the contrasts in (12) are statistically significant in her results. The other three subjects (AH, CS, YY) showed statistically significant contrasts both for (12a) and (12b).

⁹ In fact, this contrast was also statistically significant in NM's result, which was excluded in the final analysis. This contrast, however, is the only significant contrast found in NM's data.

4.3 P3

The predictions for P3 (13) are repeated below:

(13) *Predictions for P3*

- a. $A = B$ (due to pitch reset in B)
- b. $A > C$ (due to F_0 -lowering in C)
- c. $B > C$ (due to pitch reset in B and F_0 -lowering in C)

First of all, it is clear from Figure 7 that C is lower than A and B. The contrasts were both statistically significant in the normalized result (A vs. C: 1-sided t-test, $t(182.464) = 3.6626$, $p < 0.001$; B vs. C: 1-sided t-test, $t(190) = 4.2952$, $p < 0.0001$). This means that (13b) and (13c) were supported by the normalized data.

As for (13a), in order to check the equivalence of the mean F_0 of A and that of B, I used the Two One-Sided T-tests (TOST) method (Hönig and Heisey 2001; Berger and Hsu 1996). In this method, we will check whether the F_0 -mean difference between A and B (d) will fall within the range of a certain equivalent threshold ($\pm\Theta$). Here I set the threshold as $\pm 10\%$ of the mean F_0 of the control stimuli (A). This essentially means that if the mean difference between A and B (d) is within the range of -10% and $+10\%$ of the F_0 -mean of A (i.e., $-\Theta < d < \Theta$), we will conclude that A and B are equivalent. This can be checked by running two one-sided t-tests, with the following null and alternative hypotheses:

(14) *Null / alternative hypotheses tested by TOST*

- a. Test 1: $\mathcal{H}_0 : d \leq -\Theta$
 $\mathcal{H}_A : -\Theta < d$
- b. Test 2: $\mathcal{H}_0 : \Theta \leq d$
 $\mathcal{H}_A : d < \Theta$

where

$$d = \text{Mean}(B) - \text{Mean}(A)$$

$$\Theta = \text{Mean}(A) \times 0.1$$

As it turned out, although the null hypothesis of Test 1 was rejected as predicted ($t(96.444) = 2.6589, p < 0.01$), the null hypothesis of Test 2 was not ($t(96.444) = -1.3993, p = 0.08246$). Therefore our prediction (13a) was not fully confirmed from the normalized result.¹⁰

This result, however, seems to be due to one subject's result that has an extremely different tendency from the others. KS's result was different from the others in that B is significantly higher than A at P3 (see Appendix B.3). In other subjects' data (AH, CS¹¹, YY), there is no such big difference between A and B. In fact, the results of TOST show that A and B are the same in these speakers' data. Therefore, if we exclude KS's data on P3, we can conclude that A and B are actually the same. We will consider possible explanation for KS's unexpected result in the discussion section (§5).

As for the other subjects' data, AH and YY's results were basically parallel to that of normalized data, namely, A and B are at the same height (= (13a)), and C is significantly lower than A and B (= (13b), (13c)). Therefore these subjects' data basically confirms all the three predictions for P3.

CS's data did not show any significant contrasts among A, B, and C. This appears to contradict (13b) and (13c). However, it was a general tendency in CS's utterances that the pitch range is strongly narrowed down toward the end the utterance, so that all the expected contrasts (not only for this experiment, but also for other experiments, whose stimuli are inserted in the recordings as fillers)

¹⁰ If we set the threshold as $\pm 15\%$ of the F_0 -mean of A, the both null hypotheses were both rejected. (Test 1: $t(98.248) = 3.6563, p < 0.001$; Test 2: $t(98.248) = -2.4026, p < 0.01$)

¹¹ As mentioned below, CS's data did not show the contrasts expected at P3 (= (13b) and (13c)) either. Therefore the lack of difference between A and B in her data does not necessarily confirm their equivalence. See below.

were unable to detect. Given that, the lack of expected difference between C and the other two conditions is presumably due to an independent reason, most likely relatively strong declination effect, and hence would not necessarily falsify the predictions.

In sum, predictions for P3 was generally confirmed by the normalized as well as the individual results, except a couple of cases (KS's unexpected F_0 -rise in B; the lack of contrast in CS's utterance).

5 Discussion

5.1 NPI-FI hypothesis

In the previous section, I presented the results of the experiment. The results generally confirmed the predictions for P1, P2, and P3. These predictions concern the prosodic phenomena listed in (6), repeated below, which are the indications of FI in NPI sentences.

- (6) a. F_0 -rise on NPI
- b. F_0 -lowering on post-NPI material
- c. F_0 -reset on post-negation material

Given that all these predictions are confirmed by the results, we can conclude that the NPI-FI hypothesis in (3) is supported by the result of this experiment. This means that NPI sentences exhibit an FI, just like *wh*-questions. This conclusion suggests that *wh*-questions and NPI sentences share the same kind of phonological process (or, in constraint-based terms, they are subject to the same sets of prosodic constraints).

This means that the results confirm the claim made by Deguchi and Kitagawa (2002) and Ishihara (2002, 2003) (cf. §2.1), which predicts a correlation between the scope of focus (in this case, the scope of negation that binds the

NPI) and the prosodic domain of the post-focal downtrend. A pitch reset was consistently observed when the NPI and the negation appear in the embedded clause, while such a reset is absent if they appear in the matrix clause. On the other hand, Hirotsu's (2005) SPC analysis, which does not predict the one-to-one correlation between prosodic phrasing and semantic scope, would require an additional explanation for the consistency correlation found in the experiment presented here.

5.2 XP-type and Comp-type Pitch Reset

As mentioned in §2.3, I used two different measurement points for P3, assuming that there are two types of pitch reset, XP-type (P3 in (9), repeated below) and Comp-type (P3*).

- (9) *Labels of the relevant F₀ peaks*
- | | | | | | | | |
|------------------------------------|------------------------------------|--------|---------------------|----------------|------|-----------------------|---|
| [CP ₂ Subj ₂ | [CP ₁ Subj ₁ | XP ... | V ₁ -NEG | C ₁ |] YP | V ₂ (-NEG) |] |
| | P1 | P2 | | | P3* | P3 | |

The frequency of the use of Comp-type varies from subject to subject (AH: 80.6%; CS: 43.1%; KS: 40.3%; YY: 63.9%). If we compare the use of Comp-type according to the stimuli types, however, we find an interesting tendency common to all the subjects. That is, the Comp-type pitch reset is used more frequently in B, and less frequently in C, than A (see Table 1). CS's data did not show these contrasts, but this is not surprising given that her data generally do not show any significant contrast expected in P3, as mentioned in §4.3.

This fact suggests that when a pitch reset of an FI is expected, speakers tend to realize the high peak on the complementizer more frequently than in the cases where it is not expected. If we consider the Comp-type pitch reset as some kind of phrase-boundary tone, this tendency seems to make sense. Given that NPI-sentences trigger an FI and creates an FI domain between NPI-attached word

Subject	A	B	C	Total
AH	87.5	95.8	58.3	80.6
CS	41.7	45.8	41.7	43.1
KS	50.0	54.2	16.7	40.3
YY	62.5	83.3	45.8	63.9
Total	60.4	69.8	40.6	56.9

Table 1: Frequency of Comp-type Pitch Reset (%)

and NEG-attached verbal complex, pitch reset is realized as a boundary tone at the end of the FI-domain.

5.3 KS's unexpected contour on P3

We saw in §4.3 that KS's F_0 -mean of P3 in B sentence is raised much higher than expected. This F_0 -rise is of different kind from the focus F_0 -rise, which was only marginally significant at P1 in her data. It seems also different from pitch reset, which is supposedly reset the pitch range to the original height, i.e., supposedly as high as the control stimulus A.

I tentatively suggest it is upstep (Truckenbrodt 2002). Caroline Féry (p.c.) pointed out that the amount of F_0 -rise expected for pitch reset and that for upstep are predicted to be different. The pitch reset resets the pitch-range relative to the one set by the prominence of the previous domain (i.e., in (9), relative to the pitch range set by the NPI-attached phrase, Subj_1). Upstep, on the other hand, resets the pitch range relative to the topmost pitch range (i.e., relative to matrix subject, Subj_2).

It is plausible to consider that KS uses upstep, instead of F_0 -lowering and pitch reset, after the embedded clause to mark the domain of NPI-NEG relation. Recall that KS did not show clear FI effects, especially in terms of F_0 -lowering. Since she exploits no F_0 -lowering to mark the domain of NPI-NEG relation, she indicates the end point by raising the F_0 -peak at the end of the domain. Further

research is needed to find out whether this F_0 -rise is an upstep phenomenon, or how often such a pattern can be found in other speakers.

5.4 F_0 -peak on the NEG-attached V-complex

Although we did not discuss at all in this paper, the F_0 -realization of the verbal complex (which includes negation) could have been a point of discussion. It may be the case that the pitch reset takes place not after the negation, but on the negation. Unfortunately, however, the F_0 of the verbal complex was not systematically measured in this experiment. I had an impression during the measurement that this peak seems to be consistently raised to some extent. This means that negation might be outside the FI triggered by NPI. Since I did not measure this peak, no definite statement can be made regarding this F_0 -peak. Therefore it is not clear whether the post-NPI F_0 -lowering ends before this phrase or after it. I leave this question for future research.

6 Conclusion

In this paper, I presented the experimental results that shows the existence of FI in NPI sentences. The characteristic phonetic phenomena of FI, i.e., (i) F_0 -rise on the focused phrase, (ii) post-focal F_0 -downtrend, and (iii) pitch reset after the FI domain, are all attested in the data. This result indicates the parallelism between NPI sentences and *wh*-question, both of which exhibit FI to mark the semantic relation between the two elements (NPI-NEG for the former, *wh*-phrase and question particle for the latter).

Aside from the main concern of the paper, we also discussed two types of pitch reset realization, which we called XP-type and Comp-type, as well as a sharp F_0 -rise used by one subject to mark the end of NPI-NEG domain, which I tentatively consider as an upstep phenomenon. The exact properties of these variations still need to be examined. I will leave this question for future research.

Appendix A Stimulus Sets

A.1 *Nomiya* set

- (1A) Náoya-wa [Mári-ga rámu-o nomíya-de nomá-nakat-ta to] Yúmi-ni itta
Naoya-TOP Mari-NOM rum-ACC bar-LOC drink-NEG-PST that Yumi-DAT said
'Naoya said to Yumi that Mari didn't drink rum at the bar.'
- (1B) Náoya-wa [Mári-sika rámu-o nomíya-de nomá-nakat-ta to] Yúmi-ni itta
Naoya-TOP Mári-SIKA rum-ACC bar-LOC drink-NEG-PST that Yumi-DAT said
'Naoya said to Yumi that only Mari drank rum at the bar.'
- (1C) Náoya-sika [Mári-ga rámu-o nomíya-de nónda to] Yúmi-ni iwa-nakat-ta
Náoya-SIKA Mari-NOM rum-ACC bar-LOC drank that Yumi-DAT say-NEG-PST
'Only Naoya said to Yumi that Mari drank rum at the bar.'

A.2 *Roommate* set

- (2A) Náoya-wa [Mári-ga Yúmi-o ruumuméito-ni erabá-nakat-ta to] Yúuzi-ni osieta
N.-TOP M.-NOM Y.-ACC roommate-DAT choose-NEG-PST that Yuji-DAT told
'Naoya told Yuji that Mari didn't choose Yumi as a roommate.'
- (2B) Náoya-wa [Mári-sika Yúmi-o ruumuméito-ni erabá-nakat-ta to] Yúuzi-ni osieta
N.-TOP M.-SIKA Y.-ACC roommate-DAT choose-NEG-PST that Yuji-DAT told
'Naoya told Yuji that only Mari chose Yumi as a roommate.'
- (2C) Náoya-sika [Mári-ga Yúmi-o ruumuméito-ni eránda to] Yúuzi-ni osie-nakat-ta
N.-SIKA M.-NOM Y.-ACC roommate-DAT chose that Yuji-DAT tell-NEG-PST
'Only Naoya told Yuji that Mari chose Yumi as a roommate.'

A.3 *Erimaki* set

- (3A) Mári-wa [Yúmi-ga Náoya-ni erímaki-o amá-nakat-ta to] Yúuko-ni osieta
Mari-TOP Yumi-NOM Naoya-DAT scarf-ACC knit-NEG-PST that Yuko-DAT told
'Mari told Yuko that Yumi didn't knit a scarf for Naoya.'
- (3B) Mári-wa [Yúmi-sika Náoya-ni erímaki-o amá-nakat-ta to] Yúuko-ni osieta
Mari-TOP Yumi-SIKA Naoya-DAT scarf-ACC knit-NEG-PST that Yuko-DAT told
'Mari told Yuko that only Yumi knitted a scarf for Naoya.'

- (3C) **Mári-sika** [Yúmi-ga Náoya-ni erímaki-o ánda to] Yúuko-ni osie-**nákat**-ta
 Mari-SIKA Yumi-NOM Naoya-DAT scarf-ACC knitted that Yuko-DAT tell-NEG-PST
 ‘Only Mari told Yuko that Yumi knitted a scarf for Naoya.’

A.4 Boston set

- (4A) áru razió-kyoku-ga [Bósuton-wa gógo áme-ga furá-nai to]
 some radio-station-NOM Boston-TOP afternoon rain-NOM fall-NEG that
 tenki-yóhoo-de tutaeta
 weather-forecast-at reported
 ‘Some weather forecast reported that it won’t rain in Boston in the afternoon.’
- (4B) áru razió-kyoku-ga [Bósuton-**sika** gógo áme-ga furá-**nai** to]
 some radio-station-NOM Boston-SIKA afternoon rain-NOM fall-NEG that
 tenki-yóhoo-de tutaeta
 weather-forecast-at reported
 ‘Some weather forecast reported that it will rain only in Boston in the afternoon.’
- (4C) áru razió-kyoku-**sika** [Bósuton-wa gógo áme-ga fúru to]
 some radio-station-SIKA Boston-TOP afternoon rain-NOM fall that
 tenki-yóhoo-de tutae-**nákat**-ta
 weather-forecast-at report-NEG-PST
 ‘Only a certain weather forecast reported that it will rain in Boston in the afternoon.’

A.5 Aisiteru set

- (5A) Yúmi-wa [Yúuzi-ga Yúuko-o áisite-nái to] Mári-ni itta
 Yumi-TOP Yuji-NOM Yuko-ACC love-NEG that Mari-DAT told
 ‘Yumi told Mari that Yuji doesn’t love Yumi.’
- (5B) Yúmi-wa [Yúuzi-**sika** Yúuko-o áisite-**nái** to] Mári-ni itta
 Yumi-TOP Yuji-SIKA Yuko-ACC love-NEG that Mari-DAT told
 ‘Yumi told Mari that only Yuji loves Yumi.’
- (5C) Yúmi-**sika** [Yúuzi-ga Yúuko-o áisiteru to] Mári-ni iwa-**nákat**-ta
 Yumi-SIKA Yuji-NOM Yuko-ACC love that Mari-DAT tell-NEG-PST
 ‘Only Yumi told Mari that Yuji loves Yumi.’

A.6 *Maneita set*

- (6A) Yúuzi-wa [Yúmi-ga Náoya-o ié-ni manéita to] Yúuko-ni morásita¹²
 Yuji-TOP Yumi-NOM N.-ACC house-DAT invited that Yuko-DAT divulged
 ‘Yuji divulged to Yuko that Yumi invited Naoya to her house.’
- (6B) Yúuzi-wa [Yúmi-**sika** Náoya-o ié-ni maneká-**nakat**-ta to] Yúuko-ni
 Yuji-TOP Yumi-SIKA N.-ACC house-DAT invite-NEG-PST that Yuko-DAT
 morásita
 divulged
 ‘Only Yuji divulged to Yuko that Yumi invited Naoya to her house.’
- (6C) Yúuzi-**sika** [Yúmi-ga Náoya-o ié-ni manéita to] Yúuko-ni morasá-**nakat**-ta
 Yuji-SIKA Y.-NOM N.-ACC house-DAT invited that Yuko-DAT divulge-NEG-PST
 ‘Only Yuji divulged to Yuko that Yumi invited Naoya to her house.’

A.7 *Ookina mi set*

- (7A) Náoya-wa [Mári-no kí-ni óokina mi-ga nará-nakat-ta no]-o
 Naoya-TOP Mari-GEN tree-LOC big fruit-NOM be.borne-NEG-PST NL -ACC
 nobotte tasikámeta
 by.climbing checked
 ‘Naoya checked Mari’s tree didn’t bare a big fruit by climbing.’
- (7B) Náoya-wa [Mári-no kí-ni-**sika** óokina mi-ga nará-**nakat**-ta no]-o
 Naoya-TOP Mari-GEN tree-LOC-SIKA big fruit-NOM be.borne-NEG-PST NL -ACC
 nobotte tasikámeta
 by.climbing checked
 ‘Naoya checked that only Mari’s tree didn’t bare a big fruit by climbing.’
- (7C) Náoya-**sika** [Mári-no kí-ni óokina mi-ga natta no]-o nobotte
 Naoya-TOP Mari-GEN tree-LOC big fruit-NOM was.borne NL -ACC by.climbing
 tasikamé-**nakat**-ta
 check-NEG-PST
 ‘Only Naoya checked Mari’s tree bore a big fruit by climbing it.’

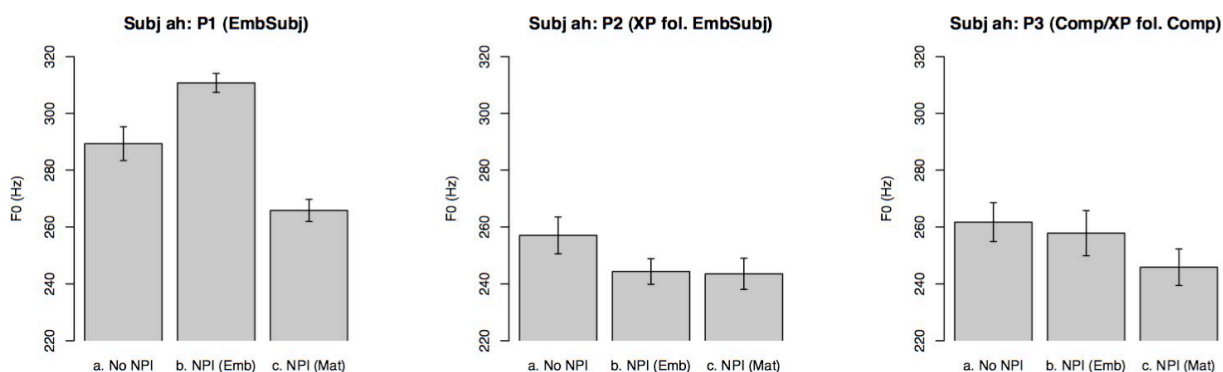
¹² The embedded clause of this sentence is supposed to contain negation. The sentence is nevertheless not excluded from the analysis, since the verbal complex itself is not the target of the measurement.

A.8 *Nomo set*

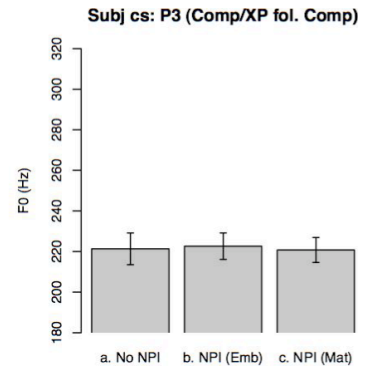
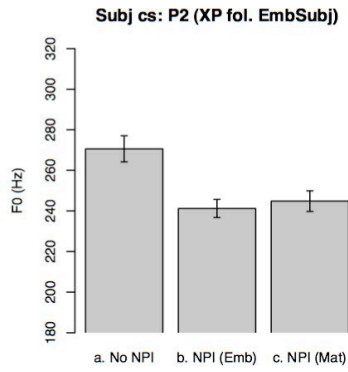
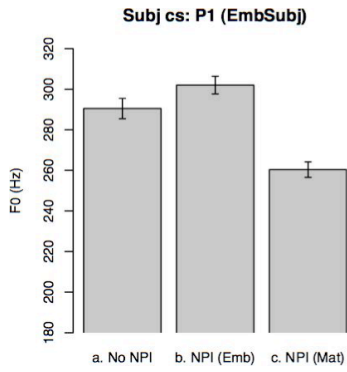
- (8A) áru nyúusu-ga [Nómo-ga Mánii-ni nákkuru-o nagé-nakat-ta to]
 some news-NOM Nomo-NOM Manny-DAT knuckleball-ACC pitch-NEG-PST that
 óokiku hoozita
 widely broadcasted
 ‘Some news program widely broadcasted that Nomo didn’t pitch a knuckleball to Manny.’
- (8B) áru nyúusu-ga [Nómo-**sika** Mánii-ni nákkuru-o nagé-**nakat**-ta to]
 some news-NOM Nomo-SIKA Manny-DAT knuckleball-ACC pitch-NEG-PST that
 óokiku hoozita
 widely broadcasted
 ‘Some news program widely broadcasted that only Nomo pitched a knuckleball to Manny.’
- (8C) áru nyúusu-**sika** [Nómo-ga Mánii-ni nákkuru-o nágeta to] óokiku
 some news-SIKA Nomo-NOM Manny-DAT knuckleball-ACC pitched that widely
 hoozi-**nákat**-ta
 broadcast-NEG-PST
 ‘Only a certain news program widely broadcasted that Nomo pitched a knuckleball to Manny.’

Appendix B Individual Results

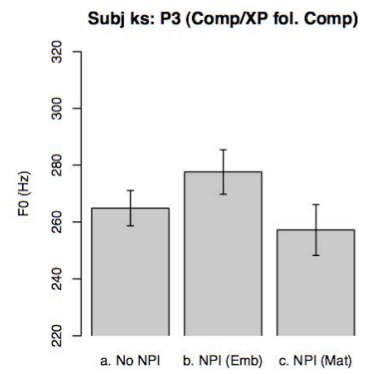
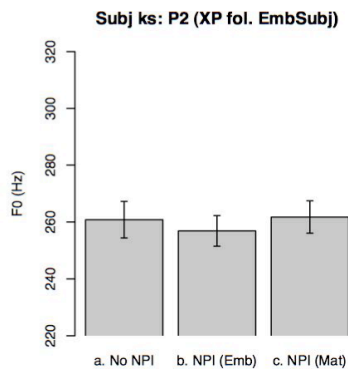
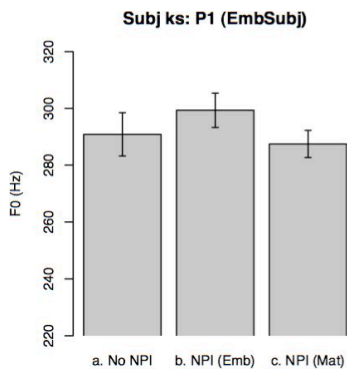
B.1 AH



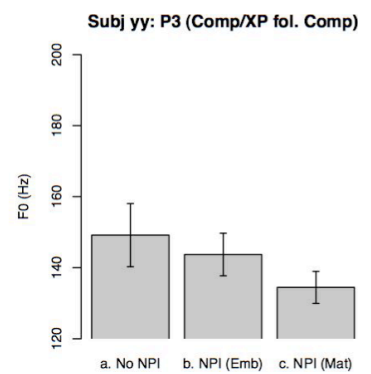
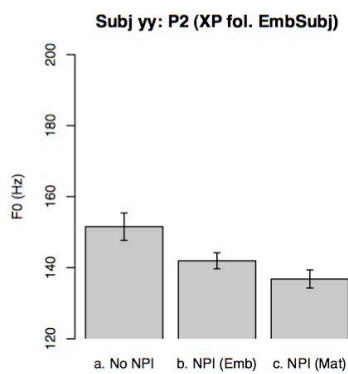
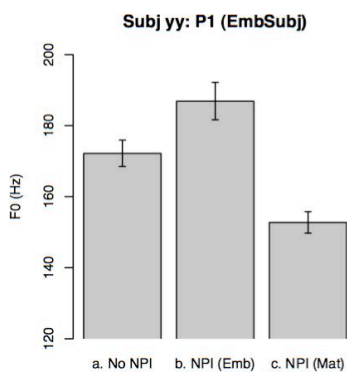
B.2 CS



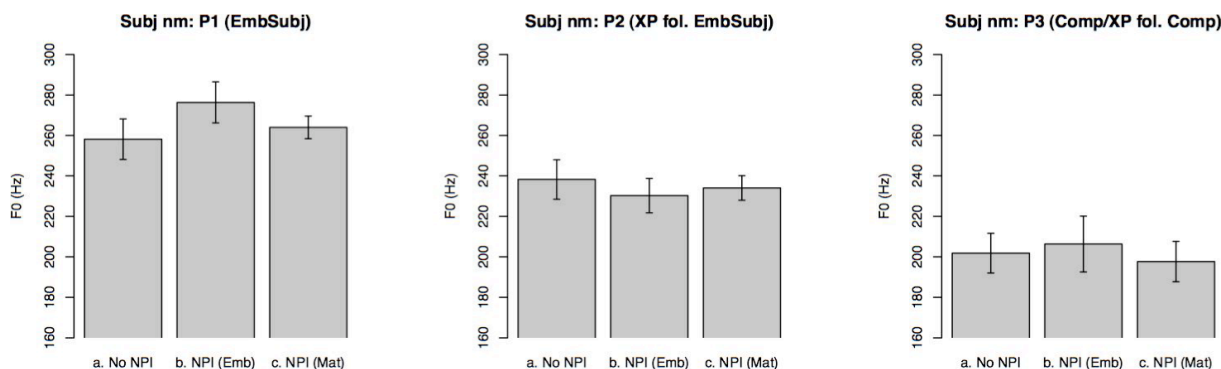
B.3 KS



B.4 YY



B.5 NM (NB: Excluded from the Normalization)



Bibliography

- Berger, Roger L., and Jason C. Hsu. 1996. Bioequivalence trials, intersection-union tests and equivalence confidence sets. *Statistical Science* 11:283–319.
- Boersma, Paul, and David Weenink. 1992–2007. Praat. Computer Software. www.praat.org.
- Deguchi, Masanori, and Yoshihisa Kitagawa. 2002. Prosody and wh-questions. In *NELS 32: Proceedings of the North Eastern Linguistics Society*, ed. Masako Hirotsani, volume 1, 73–92. Amherst, Mass.: GLSA, University of Massachusetts, Amherst.
- Hirotsani, Masako. 2005. Prosody and LF: Processing Japanese wh-questions. Doctoral Dissertation, University of Massachusetts, Amherst.
- Hönig, John M., and Dennis M. Heisey. 2001. The abuse of power: The pervasive fallacy of power calculations for data analysis. *The American Statistician* 55:19–24.
- 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, Mass.: Cascadilla Press.
- Ishihara, Shinichiro. 2003. Intonation and interface conditions. Doctoral Dissertation, Massachusetts Institute of Technology.
- Ishihara, Shinichiro. 2005. Prosody-scope match and mismatch in Tokyo Japanese *Wh*-questions. *English Linguistics* 22:347–379.
- Ishihara, Shinichiro. 2007a. Japanese downstep revisited. Poster presentation at Word Accents and Tones in Sentence Perspective: A symposium in conjunction with the 60th birthday of Professor Gösta Bruce.
- Ishihara, Shinichiro. 2007b. Major phrase, focus intonation, and multiple spell-out (MaP, FI, MSO). *The Linguistic Review* 24:137–167.
- Kubozono, Haruo. 2007. Focus and intonation in Japanese: Does focus trigger pitch reset? This volume.
- Ladusaw, William. 1979. Polarity sensitivity as inherent scope relations. Doctoral Dissertation, University of Texas at Austin.
- Lee, Kisuk, and Satoshi Tomioka. 2001. LF blocking effects are topic effects: *Wh*-questions in Japanese and Korean. Unpublished ms., University of Delaware.
- Maekawa, Kikuo. 1991. Perception of intonation characteristics of WH and non-WH questions in Tokyo Japanese. In *Proceedings of the 22nd International Congress of Phonetic Science*, volume 4, 202–205. Aix-en-Provence, France: Université de Provence.
- Maekawa, Kikuo. 1997. Nihongo gimonsi gimon-bun no intoneesyon [The intonation of the Japanese *wh*-question]. In *Bunpoo to onsei [Grammar and*

- sound*], ed. Onsei Bunpoo Kenkyuu-kai, chapter 3, 45–53. Tokyo: Kurosio Publishers.
- Nagahara, Hiroyuki. 1994. Phonological phrasing in Japanese. Doctoral Dissertation, University of California, Los Angeles.
- Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese tone structure*. Cambridge, Mass.: MIT Press.
- 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, Mass.: GSLA, University of Massachusetts, Amherst.
- Sugahara, Mariko. 2003. Downtrends and post-FOCUS intonation in Japanese. Doctoral Dissertation, University of Massachusetts, Amherst.
- Tomioka, Satoshi. 2004. Pragmatics of LF intervention effects: Wh-interrogatives in Japanese and Korean. Unpublished ms., University of Delaware.
- Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and prominence. Doctoral Dissertation, Massachusetts Institute of Technology.
- Truckenbrodt, Hubert. 2002. Upstep and embedded register levels. *Phonology* 19:77–120.
- Truckenbrodt, Hubert. 2004. Final lowering in non-final position. *Journal of Phonetics* 32:313–348.

Vasishth, Shrvan. 1998. Boolean properties of focus particles and NPIs in Japanese. Paper presented at the Annual Meeting of the Linguistic Society of America, New York City.

Shinichiro Ishihara
University of Potsdam
SFB 632 “Information Structure”
Karl-Liebknecht-Str. 24–25
14476 Golm, Germany
s_i@alum.mit.edu
http://alum.mit.edu/www/s_i/

Intervention Effects in Focus: From a Japanese Point of View*

Satoshi Tomioka

University of Delaware

The most recent trend in the studies of LF intervention effects makes crucial reference to focusing effects on the interveners, and this paper critically examines the representative analyses of the focus-based approach. While each analysis has its own merits and shortcomings, I argue that a pragmatic analysis that does not make appeal to syntactic configurations is better equipped to deal with many of the complex and delicate facts surrounding intervention effects.

Keywords: *Intervention Effect, Alternative Semantics, Wh-interrogatives, Focus, Topic, Post-Focus Reduction*

1 Introduction

Many languages exhibit what have come to be known as ‘LF intervention effects’, in which a certain kind of ‘quantificational’ expression (i.e., ‘intervener’) is prohibited from occupying surface positions that c-command a Wh-phrase. The following are some examples from Japanese.

- (1) ?***Daremo**/**?*Ken-sika** nani-o yom-ana-katta-no?
 Anyone/Ken-except what-ACC read-NEG-PAST-Q
 ‘What did no one read?’

* Many thanks to the following individuals who gave me comments, criticisms, and encouragement: Sigrid Beck, Masao Endo, Paul Hagstrom, Nobuko Hasegawa, Masako Hirotani, Shinichiro Ishihara, Shin-sook Kim, Yoshihisa Kitagawa, Heejeong Ko, Haruo Kubozono, Shige-yuki Kuroda, Chungmin Lee, Shigeru Miyagawa, and Mamoru Saito. Parts of this paper have been presented at *Workshop on Prosody and Information Structure 2* at Universität Potsdam, *Theoretical East Asian Linguistics 3*, and *Linguistic Theory and the Japanese Language Workshop*, both at Harvard University.

- (2) ??**Daremo-ga**/??**Dareka-ga**/???[Ken-**ka** Erika]-ga nani-o yon-da-no
 everyone-NOM/someone- NOM/John-or-Bill- NOM what-ACC read-PAST-Q
 ‘What did everyone/someone/Ken or Erika read?’
- (3) ???Ken-**dake-ga**/???Ken-**mo** nani-o yon-da-no?
 Ken-only- NOM/Ken-also what- ACC read-PAST-Q?
 ‘What did only Ken/Ken also read?’

Although the judgments on these examples are by no means invariable or stable, there is a very clear contrast with the examples shown below, in which the Wh-phrases are scrambled over the offending interveners.

- (4) Scrambled version of (1)
 Nani-o₁ daremo/ Ken-sika t₁ yom-ana-katta-no
- (5) Scrambled version of (2)
 Nani-o₁ daremo-ga/dareka-ga/Ken-ka Erika-ga t₁ yon-da-no
- (6) Scrambled version of (3)
 Nani-o₁ Ken-dake-ga/Ken-mo t₁ yon-da-no

Last twenty years or so have seen a considerable body of research on these intervention effects, including Hoji (1985), who originally discovered the effects in Japanese, S.-W. Kim (1990), Takahashi (1990), Beck (1996), Beck and S.-S. Kim (1997), Tanaka (1997), and Hagstrom (1998), among many others. The majority of the earlier analyses have made crucial use of the following ideas. First, they follow the tradition initiated by Huang (1982) that in-situ Wh-phrases move to Spec of CP at LF in all languages. Second, the intervention effects are claimed to be applicable only at LF so that it affects LF movement but not surface movement. Although the simplicity of this type of analysis is quite appealing, one important question was often left unanswered in

the earlier studies: Which ones are potential interveners, and which ones are not? While we can certainly make a list of interveners, it is surprisingly difficult to name a property that binds all the potential interveners as a natural class. Without knowing the exact nature of interveners, we cannot construct an explanatory analysis of the phenomenon. In this sense, I take it as a welcome change that some of the more recent work on intervention effects (e.g., Beck 2006, Kim 2002, 2005, Tomioka 2007a) takes this question as the starting point.

2 The Role of Focus in Intervention

2.1 Focus Operator ~ as an Intervener

At the beginning of this paper, I mentioned that a certain kind of ‘quantificational’ expression participates in intervention effects. Indeed, the term ‘quantificational’ has been popular as the description of interveners (e.g., Hoji 1985, Beck 1996, Beck and Kim 1997). Although it covers the majority of interveners, the term is misleading in two respects. On the one hand, there are expressions that are quantificational but do not act as interveners. In Japanese, for instance, *subete/zenbu-no NP* ‘all NP’ and *hotondo/daibun-no NP* ‘most NP’ do not seem to induce relevant effects.¹ On the other hand, there are non-quantificational interveners, such as *NP-mo* ‘NP-also/even’.

¹ This statement needs further clarification. Those quantifiers do not show intervention effects when they are topic-marked. With the nominative marker *ga*, they do show the effects. As far as I know, Shin Watanabe is the first one to connect intervention effects with the topic-nominative alternation (Watanabe 1998). His formulation is the following:

- (i) A Japanese wh-question is ill-formed, when a non-wh NP marked with *ga* co-commands a wh-phrase in it. (Watanabe 1998, footnote 7 (ii))

Beck (2006) and Kim (2002, 2005) offer new insight on the identification of potential interveners. Their generalization states that potential interveners are the expressions that come with the focus operator \sim in the sense of Rooth (1992). Thus, the intervention effects are re-formulated as the result of a prohibition against the following structural configuration.

(7) * [Q-Op [YP \sim [[XP]_F [Wh]]] where \sim is Rooth's focus operator.

Building on the same generalization, however, Beck (2006) and Kim (2002, 2005) offer different analyses. For Beck, the intervention effects come about because of the interaction between Rooth's (1992) semantics of \sim and Hamblin's (1973) semantics of Wh-interrogatives, a truly in-situ theory of Wh-questions. The ordinary value of a focused XP is the usual denotation of XP while the focus value of XP is a set of semantic objects of the same type as the denotation of XP. Unlike ordinary expressions, Wh-expressions only have focus values (= the Hamblin set denotations) and lack ordinary values.² Thus, in a non-intervention situation as in (8), the constituent (YP) that is a sister of a Q(uestion) Operator inherits the property of having only one semantic value (i.e., a focus value).

(8) [Q-Op [YP [Wh]]

² It should be noted, however, that a Wh-phrase can be contrasted with another Wh. For instance;

- (i) I already know when Robin got married. I just want to know WHERE she got married.

Intuitively, the focus value of a contrasted Wh should be a set of Hamblin sets, but it is not clear how it should be calculated if a Wh-phrase lacks an ordinary value. Perhaps, one can assume that the focus value of the Wh is already elevated to the ordinary value by the Q-operator in (i), which makes the computation of the focus value possible. The question still remains, however, since the Japanese counterpart of (i) has the Wh in situ as usual.

The role of the Q-Operator is to elevate the focus value of YP to the ordinary value, which results in a set of propositions that constitute possible answers as the ordinary value of the question. Something goes wrong, however, in a configuration like (7) where a focus operator \sim intervenes between a Q-Op and an in-situ Wh. The operator \sim computes the focus value of its complement, but beyond it, the focus value is ‘closed off’, so to speak, and it would be a singleton set of the ordinary value of its sister, as shown in (9).

- (9) $[_{YP} \sim [_{ZP} \dots]]$
 a. The ordinary value of YP: $\llbracket YP \rrbracket^o = \llbracket ZP \rrbracket^o$
 b. The focus value of YP: $\llbracket YP \rrbracket^f = \{\llbracket ZP \rrbracket^o\}$

Recall, however, that a Wh-phrase only has a focus value, and so does all the constituents containing a Wh-phrase prior to its combination with a Q-Op. This means that neither the focus nor the ordinary value is defined at the YP level in (7). Beck argues that this lack of definedness is the source of intervention effects. Of course, there is nothing wrong when a Wh-phrase is scrambled over the \sim operator, as in (10).

- (10) $[Q\text{-Op } [_{Wh}]_i \quad [_{YP} \sim [[_{XP}]_F \quad t_i]]$

In (10), both the focus and the ordinary values of YP are well-defined (with respect to assignment function g).

Although Kim’s (2002, 2005) solution also appeals to focus as the source of intervention effects, it is essentially syntactic. Assuming that a Wh-phrase is inherently focused, she argues that a Q-Op and an in-situ Wh have both a Wh-feature and a focus feature. Furthermore, a Q-Op and an in-situ Wh must be in AGREE relation in the sense of Chomsky (2000). This relation is disturbed

when there is an intervening element which also has a focus feature. The necessary AGREE relation can be established by the scrambling of the Wh-phrase over the offending intervener. (11ab) show how Kim's system works.

- (11) a. * [Q-OP_[WH, F] [F] ~ [Wh_[WH, F]]
 No AGREE !
- b. [Q-Op_[WH, F] [[Wh]_{I[WH, F]} ... [F] ~ t₁ ...]
 AGREE

Although it was motivated by a different generalization, my characterization of interveners in Tomioka (2007a), which originates in Lee and Tomioka (1999), shares with Beck's and Kim's the idea that focus-related nature of interveners is at the core of intervention effects. The main observation is that all of the potential interveners in Japanese and Korean resist the morphological topic marking (i.e., *wa* in Japanese and (*n*)*un* in Korean) while the non-interveners, whether they are quantificational or referential, can have the topic morphology. Based on this generalization, I called potential interveners A(nti)-T(opic) I(tem)s and argued that the anti-topicality of intervener leads to intervention effects through the following steps. First, I assume, following Krifka (2001), that in an ordinary Wh-question, the Wh-phrase is focused while the rest belongs to the background. Second, topics belong to the background (cf. Kuno 1973 among others), but an ATI cannot be in the background by virtue of being Topic-marked. Thus, an ATI must find some other way to be in the background. Third, familiarity in discourse is often associated with a particular prosodic pattern, labeled as 'deaccenting' (Tancredi 1992, Fox 1999), 'phonological reduction' (Rooth 1992b), 'eradication' (Deguchi and Kitagawa 2002), or 'compression' (Hirotani 2004). It generally refers to the reduction of prosodic prominence of various sorts (pitch, duration, stress, etc.). It has been noted by many scholars (e.g., Pierrehumbert and Beckman 1988, Nagahara

1994, Ishihara 2002, 2004 among others) that the location of prosodic reduction in Japanese is intimately related to the location of focus. In particular, focusing in Japanese leads to a prosodic pattern in which high pitch accent is placed on the focused material and pitch accents of the material which linearly follows the focus are reduced. (12) illustrates the correspondence between the prosodic pattern and the syntactic structure.

(12) Syntax: [..... []_{focus}]#

Phonology: [[^{P(itch) P(rominence)}]] ← This part is reduced.

This phonology-syntax correspondence has significant impact on wh-scrambling. Since a wh-phrase is focused, scrambling of it leads to the extension of the reduced portion to the left. When scrambling of a wh-phrase 'jumps over' an intervener in intervention contexts, the intervener is newly placed in the reduction portion, as shown in (14).

(14) Syntactic structure with Wh-scrambling [Wh]₁ [[Intv] ..t₁ ...]

Phonological phrasing [^{PP} Wh] Intv] ← This part is reduced

Since not being in the background means being (part of) focus, (12d) can be restated as 'the source of intervention effects is an ATI being (part of) focus'. In this sense, all of the three analyses treat focus as a key ingredient. Nonetheless,

the notion of focus is not identical among the three authors, and I will raise this issue in the following sub-section.

2.2 Are All Foci the Same?

Both in Beck's and Kim's analyses, the presence of a \sim operator is intimately tied to the intervention effects. Although Beck primarily concerns with focus-sensitive expressions, such as *only*, Kim (2005) mentions, citing the Korean counterpart of the Japanese sentence (13), that an instance of free focus or sentence focus also induces an intervention effect.

- (13) ???[KEN]_F -ga nani-o yon-da-no?
 Ken - NOM what- ACC read-PAST-Q
 'What did KEN read?'

Then, the potential interveners are categorized into two sub-types: (i) expressions that serve as sentence foci, and (ii) focus-sensitive expressions, such as *only*, *even*, *also*. These two types are not the same. While the presence of \sim in the first type depends on the context when the sentence is uttered, the very semantics of the second type requires it no matter where and when it appears. The following example illustrates this point.

- (14) A: Dave only eats MEAT.

B: Oh, no. You got that wrong. (It's) ERIC (*that*) *only eats meat*.

In the last sentence of B, The subject *Eric* is focused, and the rest belongs to the background or the old information. Nonetheless, the meaning of *only* requires the presence of a \sim operator.

This type of example is well-known, but I would like to point out that, in default cases, focus sensitive expressions are also (part of) sentence foci as well. To see this point more clearly, let us consider the standard meaning of focus-sensitive operators (cf. Rooth 1985).

- (15) a. Only ERIC left early.
b. Even ERIC left early.
c. ERIC also left early.
- (16) a. The presupposition of (15a): Eric left early.
The assertion of (15a): No one but Eric left early.
b. The presupposition of (15b): Someone other than Eric left early, and Eric was the least likely person to leave early.
The assertion of (15b): Eric left early.
c. The presupposition of (15b): Someone other than Eric left early.
The assertion of (15b): Eric left early.

The underlined parts of the presuppositions in (16) indicate that the VP meaning in (15a-c) is considered as part of old information. Under the assumption that every sentence must contain some constituent that serves as new information (cf. Vallduvi 1992), this in turn means that focus operators plus their focus-associates constitute new information. Obviously, something more complicated has to be done to deal with cases like (14), where a focus sensitive expression belongs to the background.

At this point, I would like to raise the following question: Do focus sensitive items induce intervention effects simply because of their semantics or because they are often sentence foci as well? The data to be discussed in the next

section indicate that the relevant notion is sentence focus, rather than semantic sensitivity to focus.

3 Matrix Subject Condition in Intervention Effects

One of the key discoveries in Tomioka (2007a) is that intervention effects are the most prominent in Japanese and Korean when the interveners are the matrix subjects. All the sentences in (1)-(3), for instance, have the interveners as the matrix subjects. When this ‘matrix subject’ condition is not met, the effects are either non-existent or significantly weaker.

(17) Embedded Subject Intervenors

- a. Kimi-wa [_{CP} daremo/dareka/Ken-ka Erika/-ga nani-o yon-da
you-TOP everyone/someone/Ken or Erika-NOM what- ACC read-PAST-
-to] omotte-iru-no
-C think-be-Q

‘What do you think that everyone/someone/Ken or Erika read?’

- b. Kimi-wa [_{CP} [Ken-mo/Ken-dake-ga nani-o yon-da-to]
you-TOP Ken-also/Ken-only- NOM what-ACC read-PAST-C
omotte-iru-no
think-be-Q

‘What do you think that Ken also / only Ken read?’

(18) Dative-Marked Indirect Object Intervenors

- a. Ken-wa daremo/dareka/Ken-ka Erika/-ni nani-o mise-ta-no
Ken-TOPEveryone/someone/Ken or Erika-DAT what-ACC show-PAST-Q
‘What did Ken give to everyone?’

- b. Ken-wa Erika-ni-mo/-dake nani-o mise-ta-no
Ken-TOP Erika-DAT-also/only what-ACC show-PAST-Q
‘What did Ken show also/only to Erika?’

(19) Dative/Accusative-Marked ‘Raised’ Object Interveners in Causatives

a. Ken-wa daremo/dareka/Anna-ka Erika/-ni nani-o
 Ken-TOPEveryone/someone/Anna or Erika/-DAT what-ACC
 yom-ase-ta-no
 read-CAUSE-PAST-Q
 ‘What did Ken make everyone/someone/Anna or Erika read?’

b. Ken-wa Erika-dake-(o) doko-ni ik-aset-a-no
 Ken-TOP Erika-or Anna-ACC where-LOC go-CAUSE-PAST-Q
 ‘Where did Ken make Erika or Anna?’

The matrix subject condition in intervention effects has curious correspondence with a particular information structural property of matrix subjects in Japanese.

(20) Obligatory Focus Generalization for Non-topic Subjects

A matrix subject is interpreted as (part of) sentence focus when it is not marked with the topic marker *-wa*.

There are a few facts that exemplify the generalization in (20). Kuroda (1965) observed that a *ga*-marked subject with an individual-level predicate leads to the exhaustive interpretation while no such effects are found with a *wa*-marked subject.³ As is well-known, the exhaustivity implicature is often associated with focusing (cf. Krifka 1993). While Japanese lacks systematic definite marking and a bare common noun can be either definite or indefinite, a bare common noun subject with *ga* is necessarily understood to be indefinite (cf., Portner and Yabushita 1998, Tomioka 2007b). Importantly, neither of these properties is found with embedded subjects or non-subjects (cf. Heycock 1994, Tomioka 2007b).

³ When the predicate is stage-level, the sentence can be interpreted either as the exhaustive or as what Kuno (1973) calls the ‘neutral description’ interpretation. The second meaning is often described as the entire sentence being focused. Thus, even in the second interpretation, the nominative subject is a part of focus.

I do not believe that it is accidental that the matrix subject condition applies both to intervention effects and to obligatory focus interpretation. My interpretation of the correlation is that intervention effects come about when interveners are sentence foci. The presence of a \sim operator that c-commands a Wh-phrase does not necessarily lead to intervention effects.

4 Possible Amendments

Since the pragmatic approach that I advocated in Tomioka (2007a) was tailored specifically for the facts about matrix subjects, it has obvious advantage over its competitors. Although I will not repeat the detailed account that has already been given in Tomioka (2007a), the gist of it is that the obligatory focus interpretation does not apply to ATIs that are not matrix subjects. Thus, they can be more readily in the background even when they are not in the post-focus reduction part.

Despite this obvious advantage, the facts in the previous section do not necessarily refute the two alternatives. Of the two analyses, Kim's analysis is easier to fix. The only thing that needs to be done is to abandon the assumption that the presence of \sim automatically leads to the presence of a focus feature. Once we restrict assigning a focus feature only to sentence foci, a focus sensitive expression that is already in the background does not induce intervention effects. The most straightforward modification that can be made to Beck's analysis is LF movement. When the intervener is not a sentence foci, the in-situ Wh can undergo covert shifting at LF, the result of which is a kind of configuration that does not cause intervention effects. If such a movement can be independently motivated, it will also solve the problem for Kim's analysis as well. Thus, the key to success for this modification is to find a good reason for the contrast of the following kind.

- (21) a. [Wh₁ [~ [INTV_[+FOCUS] t₁]]]
 └── LF movement * ─┘
- b. [Wh₁ [~ [intv_[-FOCUS] t₁]]]
 └── LF movement OK ─┘

LF movement over a focused intervener is prohibited while taking the focus property away from the intervener makes the movement licit. Are there any other instances of LF movement that is susceptible to the focus feature of an intervening material? So far as I know, there aren't any. For instance, quantifier scope seems insensitive to the focus difference.

- (22) a. Ken-wa [Erika-ni-DAKE]_F hotonodo-no hito-o syookai-sita.
 Ken-TOP Erika-DAT-only majority-GEN people-ACC introduce-PAST
 'Ken introduced most people only to ERIKA.'
- b. Ken-wa Erika-ni-dake [HOTONDO]_F -no hito-o syookai-sita.
 Ken-TOP Erika-DAT-only majority-GEN people-ACC introduce-PAST
 'Ken introduced MOST people only to Erika.'
- c. [KEN-ga]_F Erika-ni-dake hotonodo-no hito-o syookai-sita.
 Ken-NOM Erika-DAT-only majority-GEN people-ACC introduce-PAST
 'KEN introduced most people only to Erika.'

With respect to the relative scope of *only* and *most* is concerned, all the three examples in (22) are identical. The *only* > *most* scope reading is overwhelmingly prominent whereas the other reading is absent or considerably weak. It appears that the strength of the reverse scope reading may be a bit different among native speakers, but (i) the *most* > *only* scope becomes instantly available when the direct object is overtly scrambled over the indirect object, which punctuates the weakness of that scope reading in the canonical word order, and (ii) more crucially, we do not find the pattern we are looking for: Not assigning a sentence focus to the focus sensitive expression *Erika-ni-dake*, as in (22b) or (22c), does not seem to encourage the reverse scope interpretation.

If we are not to adopt the LF movement amendment to Beck's theory, perhaps we may consider as an alternative the kind of approach that Beck abandons, namely a selective binding system of focus association (e.g., Wold 1996). The basic idea is that a \sim operator is not as blind as it initially appears. It does not make use of all possible focus values in its scope but is rather associated only with a constituent that is co-indexed with it. Although I acknowledge that Beck's criticisms are valid, I am inclined to think that the selective binding approach has some advantages as well. First, it is intuitively attractive to suppose that a focus sensitive expression in the background keeps the old association intact and does not expand to new association. Otherwise it would not be considered a part of old information. In this sense, it is expected not to play a role in intervention effects. Second, Beck's system already has a selective binding component in the association between a Wh-phrase and a Q-operator. The selectivity derives the well-known Baker ambiguity (Baker 1970). However, it has been noted that the Baker ambiguity itself is sensitive to focusing. As shown below, focusing on the in-situ Wh phrase is a necessary condition for its matrix scope (the *reduced italics* indicate phonological reduction).

- (23) a. WHO *asked who bought what* ? Only the embedded scope for *what*
 b. WHO *asked who bought*WHAT? The matrix scope for *what* possible

The contrast in (23) suggests that the Baker ambiguity is not a simple matter of selective binding. It remains to be seen whether it is technically possible to incorporate selectivity into the semantics of \sim without evoking the problems that Beck points out.

5 More Things to Consider

5.1 Post-Wh Focusing

While the three analyses that have been considered in this paper are to some extent similar in their ways of accounting for the basic intervention facts, they are dramatically different in their interpretations of Wh-scrambling. On the one hand, Beck's and Kim's analyses are still 'structural' in the sense that scrambling creates LF configurations that are legitimate. In Tomioka (2007a), on the other hand, assigning a focused status to non-Wh expressions is generally prohibited, and the improvement effects of scrambling are by-products of giving rise to post-focus phonological reduction. Therefore, the three analyses make different predictions for surface structure like the following.

(24) [...Wh₁ [INTV_[+FOCUS]] t₁]

For Beck and Kim, the structure above would not cause any problem since the Wh-phrase is moved over the ~ operator. For my account, this is expected to be unacceptable. The improvement by scrambling is a result of ensuring the intervener to be a part of the background. By focusing the intervener in that position, however, this improvement effect should be cancelled.

With this difference in mind, let us look at some data in which interveners are placed below Wh-phrases but are nonetheless focused. The results are generally in favor of the pragmatic approach I proposed, but there are some unexpected complications. For many interveners, giving them focused status after Wh-scrambling indeed brings back intervention effects. Here are some relevant examples.

- (25) a. ?*NAni-o [daREMO]_F *yom-ana-katta-no?*
 what-ACC anyone read-NEG-PAST-Q
- b. NAni-o *daremo yom-ana-katta-no?*
 what-ACC anyone read-NEG-PAST-Q
 'What did no one read?'
- (26) a. ???NAni-o [DAREka-ga]_F *yon-da-no?*
 what-ACC someone- NOM read-PAST-Q
- b. NAni-o *dareka-ga yon-da-no?*
 what-ACC someone- NOM read-PAST-Q
 'What did someone read?'

For expressions with focus-sensitive particles, such as *dake* 'only' and *sae* 'even', however, the results are not straightforward. First of all, there is not a uniform pitch accent pattern when such expressions are focused. Either the part that is associated with the particle gets prominence or else the particle itself receives a pitch peak.

- (27) a. ERika-dake-ga / Erika-daKE-ga ki-ta
 ERIKA-only- NOM Erika-ONLY- NOM come-PAST
 'Only Erika came.'
- b. ERika-sae / Erika-SAe ki-ta
 ERIKA-even Erika-EVEN come-PAST
 'Even Erika came.'

Although both pitch patterns are possible, they seem to behave differently in intervention contexts. According to my judgment, putting prosodic prominence on the focus-associates is significantly worse than the other pitch pattern. Not surprisingly, the total reduction is acceptable, just as is the case with other interveners.

- (28) a. NAni-o Erika-daKE-ga / ??ERika-dake-ga *kat-ta-no?*
 what-ACC Erika-ONLY- NOM / ERIKA-only- NOM buy-PAST-Q
- b. NAni-o *Erika-dake-ga kat-ta-no?*
 what-ACC Erika-only- NOM buy-PAST-Q
 'What did only Erika buy?'
- (29) a. NAni-o Erika-SAE / ??ERika-sae *kat-ta-no?*
 what-ACC Erika-EVEN / ERIKA-even buy-PAST-Q
- b. NAni-o *Erika-dake-ga kat-ta-no?*
 what-ACC Erika-only- NOM buy-PAST-Q
 'What did even Erika buy?'

This kind of contrast is unexpected for all the three analyses under consideration but for different reasons. (28a) and (29a) should be acceptable for Beck and Kim, no matter which accent pattern is chosen. Therefore, the degradation with *ERika-dake/sae* is surprising. What is unexpected for my pragmatic account is the fact that focusing on the focus particles is acceptable in the post-Wh positions.

All in all, the facts about post-Wh focusing provide additional support for the pragmatic approach, rather than the structural approach. The prosodic patterns of focus-sensitive particles remain as a puzzle, however, and more careful investigation is called for.

5.2 NPI Interveners

While the grammaticality judgment on intervention effects is notoriously variable and unstable, there are a few things that every native speaker is in agreement about. One is the scrambling effect that we have already discussed. The other concerns the types of interveners. For all speakers, Negative Polarity

Items (NPIs), such as *daremo* ‘anyone’ and *Erika-sika* ‘anyone/no one except Erika’, are the strongest interveners, as noted in Tomioka (2007a). Not only do NPIs induce the strongest intervention effects in the basic cases (e.g., (1ab) are worse than the examples in (2) and (3)), but their effects also persist even in embedding and with non-subjects.

- (30) a. ??Kimi-wa [_{CP} daremo nani-o yom-ana-katta-to] omotteiru-no
 you-TOP anyone what-ACC read-NEG-PAST-comp think-Q
 ‘What do you think that no one read?’
- b. ?? Kimi-wa [_{CP} John-sika nani-o yom-ana-katta-to] omotteiru-no
 you-TOP John-except what-ACC read-NEG-PAST-comp think-Q
 ‘What do you think that no one read?’
- (31) a. ???Ken-wa dare-ni-mo nani-o mise-naka-tta-no
 Ken-TOP who-DAT-also what-ACC show-NEG-PAST-Q
 ‘What didn’t Ken show to anyone?’
- b. ???Ken-wa Erika-ni-sika nani-o mise-naka-tta-no
 Ken-TOP Erika-DAT-exceptwhat-ACC show-NEG-PAST-Q
 ‘What didn’t Ken show to anyone but Erika?’

This special status of NPIs is not accounted for under any existing analyses. In Tomioka (2007a), I suggest that the peculiarity of NPIs requires a ‘hybrid’ approach that has an additional component that is specifically tailored for NPIs. Hirotani (2004) notes the phrasing tendency that NPIs are in the intermediate phrase (or the major phrase) that includes their licensors. Now, consider (31a). The Wh-phrase *nani-o* ‘what-acc’ intervenes between the dative NPI *dare-ni-mo* ‘who-dat-MO’ and the verb-neg complex *mise-naka* ‘show-neg’. Since the Wh-phrase is focused, an intermediate phrase boundary is inserted immediately before the Wh, as shown below.

(32)

...	dare-ni-mo	[_i [NAni-o] _F	mise-naka-tta-no]
		↑ licenser	

The prosodic pattern in (32) goes against the aforementioned phrasing preference, which affects the acceptability negatively. Whether this particular account is correct or not, it exemplifies a strategic proto-type that I believe is needed: A constraint/principle that governs the basic cases of intervention effects is coupled with an additional one that targets NPIs.

6 Conclusion

The main purpose of this paper is to give a kind of progress report on our understanding of intervention effects. I believe that we now have firmer grasp of the nature of intervention effects than before. The influence of focus and/or related informational structural properties is at the core of this phenomenon. As Beck (2006) and Kim (2005) point out, the relevance of focus in intervention effects seems to hold in a number of languages that are not typologically related, which gives support for the overall scheme of things. As an advocate of a pragmatic account based on information structure of Wh-interrogatives, I am naturally more inclined to find a structural account of focus, such as Beck's or Kim's, less attractive than its alternative. I am not overly optimistic, however, that we will or should have one *carte blanche* solution, structural or otherwise, that takes care of all aspects of intervention effects. In this paper, I have given microscopic views of intervention effects from a Japanese point of view, and in light of the complexity and subtlety of the intervention phenomena, I am led to conclude that a focus-based account (of one's choice) should be augmented by

some auxiliary constraint(s) to straighten out the wrinkles that the main account leaves behind.

Reference

- Baker, C. L. 1970: 'Notes on the Description of English Questions: The Role of an Abstract Question Morpheme,' *Foundation of Language* 6: 197-219
- Beck, Sigrid. 1996: 'Quantified structures as barriers for LF movement,' *Natural Language Semantics* 4, 1-56.
- Beck, Sigrid. 2006: 'Intervention effects follow from focus interpretations,' in *Natural Language Semantics*, 14:1-56.
- Beck, Sigrid and Shin-Sook Kim. 1997: 'On Wh- and operator-scope in Korean,' *Journal of East Asian Linguistics* 6, 339-384.
- Chomsky, Noam. 2000: 'Minimalist Inquiries,' in Martin, Roger et al (eds.) *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, 89-155. Cambridge, Mass.: The MIT Press.
- Deguchi, Masanori, and Yoshihisa Kitagawa. 2002: 'Prosody and Wh-questions' in Hirotani, M. (ed.) *The Proceedings of NELS 32*, GLSA, University of Massachusetts at Amherst.
- Hagstrom, Paul. 1998: *Decomposing Questions*, Doctoral dissertation, MIT.
- Heycock, Caroline. 1994: 'Focus projection in Japanese,' *Proceedings of NELS 23*, 157-171, GLSA, University of Massachusetts at Amherst.
- Hirovani, Masako. 2004: Prosody and LF: Processing Japanese Wh-questions, Doctoral dissertation, University of Massachusetts at Amherst.
- Hoji, Hajime 1985: *Logical Form Constraints and Configurational Structures in Japanese*, Doctoral dissertation, University of Washington.
- Huang, James, C-T. 1982: *Logical Relations in Chinese and the Theory of Grammar*, Doctoral dissertation, MIT.

-
- Ishihara, Shinichiro. 2002: 'Invisible but Audible Wh-scope Marking: Wh-constructions and Deaccenting in Japanese,' in Mikkelsen, L., and C. Potts (eds.) *The Proceedings of WCCFL 21*, Somerville: Cascadilla Press, 180-193.
- Ishihara, Shinichiro. 2004: *Intonation and Interface Condition*, Doctoral Dissertation, MIT.
- Kim, Shin-sook. 2002: 'Intervention Effects are Focus Effects,' in N. Akatsuka and S. Strauss (Eds.) *Japanese and Korean Linguistics 10*, Stanford: CSLI Publications, 615-628.
- Kim, Shin-sook. 2005: 'Focus Intervention Effects in Questions,' presented at *Theoretical East Asian Languages 3*, Harvard University.
- Kim, Soo-won. 1991: *Chain Scope and Quantification Structure*, Doctoral dissertation, Brandeis University.
- Krifka, Manfred. 1993: 'Focus, quantification, and dynamic interpretation,' ms. University of Texas at Austin.
- Krifka, Manfred. 2001: 'For a Structured Meaning Account of Questions and Answers,' in Fery Caroline & Wolfgang Sternefeld (eds.), *Audiatur Vox Sapientia. A Festschrift for Arnim von Stechow*, Akademie Verlag (= *studia grammatica* 52), Berlin 2001, 287-319.
- Kuno, Susumu 1973: *The Structure of the Japanese Language*, Cambridge: MIT Press.
- Kuroda, Shigeyuki. 1965: *Generative Grammatical Studies in the Japanese Language*, Doctoral dissertation, MIT.
- Kuroda, Shigeyuki. 1992: *Japanese Syntax and Semantics*, Dordrecht: Kluwer Academic Publishers.
- Lee, Kisuk, and Satoshi Tomioka. 1999. 'LF Intervention Effects are Topic Effects,' presented at *Japanese and Korean Linguistics 10*, UCLA.
- Pierrehumbert, Janet, and Mary Beckman. 1988. *Japanese Tone Structure*, Cambridge, Mass: MIT Press.
- Portner, Paul and Kazuhiko Yabushita. 1998: 'The Semantics and Pragmatics of Topic Phrases,' *Linguistics and Philosophy* 21, 117-157.

-
- Rooth, Mats. 1985. Association with Focus. Doctoral Dissertation, University of Massachusetts, Amherst.
- Takahashi, Daiko. 1990: 'Negative Polarity, Phrase Structure, and the ECP,' *English Linguistics* 7, 129-146.
- Tanaka, Hidekazu. 1997: 'Invisible Movement of *sika-nai* and the Linear Crossing Constraint,' *Journal of East Asian Linguistics* 6: 143-188.
- Tomioka, Satoshi. 2007a: 'Pragmatics of LF Intervention Effects: Wh-interrogatives in Japanese and Korean,' in *Journal of Pragmatics* 39: 1570-1590.
- Tomioka, Satoshi. 2007b: 'Japanese Existential Possession: A Case Study of Pragmatic Disambiguation,' in *Lingua*, 117: 881-903.
- Wold, Dag. 1996: 'Long Distance Selective Binding: The Case of Focus,' in Galloway, T., and J. Spence (eds.) *The Proceedings of SALT 6*, 311-328, Cornell University.

Satoshi Tomioka
University of Delaware
Department of Linguistics
46 E. Delaware Avenue
Newark, DE 19716, U.S.A.
stomioka@udel.edu
<http://www.ling.udel.edu/tomioka/main01.html/>